PONTIFICIAE ACADEMIAE SCIENTIARVM SCRIPTA VARIA

22

SEMAINE D'ETUDE

SUR LE

PROBLEME DES MACROMOLECULES D'INTERET BIOLOGIQUE

AVEC REFERENCE SPECIALE AUX NUCLEOPROTEIDES

(23-31-X-1961)

PONTIFICIA
ACADEMIA
SCIENTIARVM

EX AEDIBVS ACADEMICIS IN CIVITATE VATICANA

MCMLXII

PONTIFICIAE ACADEMIAE SCIENTIARVM SCRIPTA VARIA

22

SEMAINE D'ETUDE

SUR LE

PROBLEME DES MACROMOLECULES D'INTERET BIOLOGIQUE

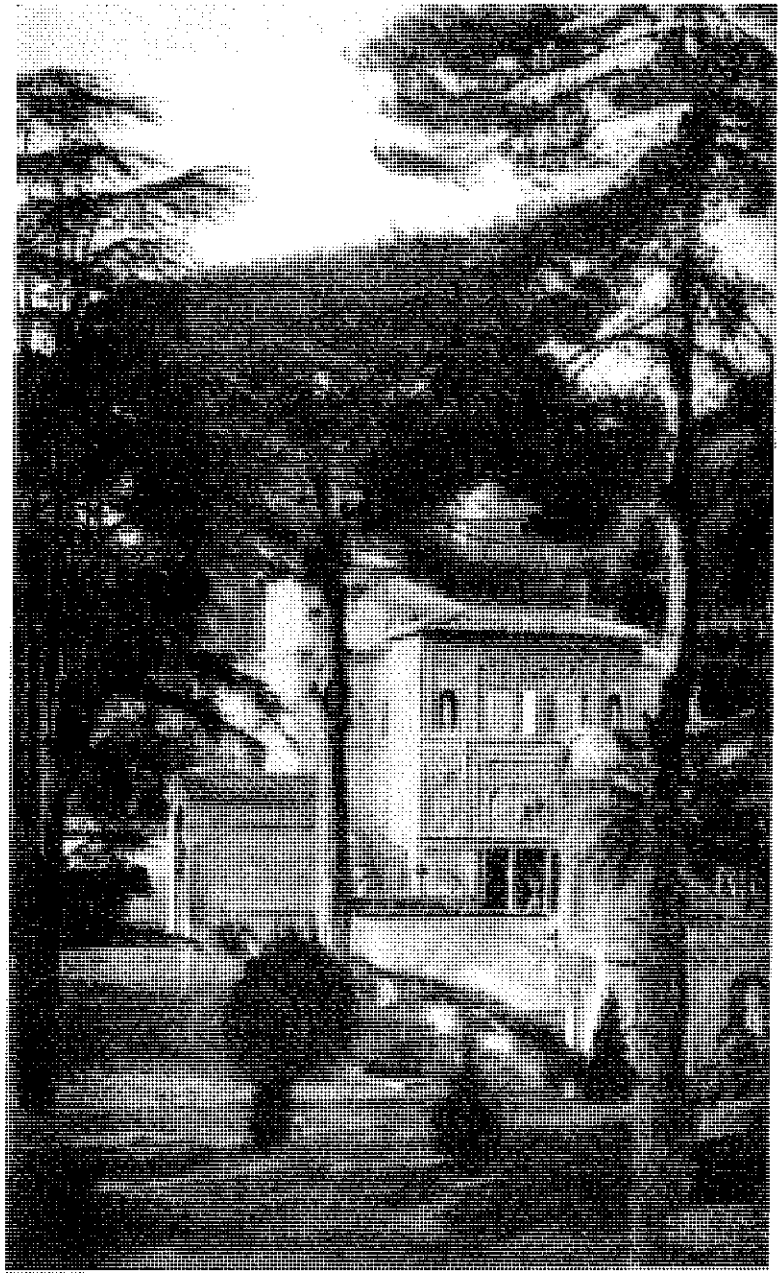
AVEC REFERENCE SPECIALE AUX NUCLEOPROTEIDES

(23-31-X-1961)

PONTIFICIA
ACADEMIA
SCIENTIARVM

EX AEDIBVS ACADEMICIS IN CIVITATE VATICANA

MCMLXII



LE PROBLEME DES MACROMOLECULES D'INTERET BIOLOGIQUE AVEC REFERENCE SPECIALE AUX NUCLEOPROTEIDES

Les brillants résultats atteints par la recherche scientifique au cours de ces dernières années en se servant de techniques très fines, ont permit de déterminer la structure de beaucoup de macromolécules d'intérêt biologique. Parmi ces molécules gigantesques ont été étudiées celles de l'insuline, de la myoglobine, de l'hémoglobine, certaines hormones protéiques et, enfin, les acides nucléiques.

Ces résultats ont contribué d'une façon trè considérable à la connaissance des rapports entre la structure des systèmes moléculaires les plus complexes — tels, par exemple, les virus, les gènes, les chromosomes —, et la fonction biologique de ces systèmes macromoléculaires.

La connaissance de la structure primaire des protéines est subordonnée à la connaissance du poids moléculaire, de la nature des amino-acides qui constituent les chaînes protéiques et de la séquence selon laquelle ces amino-acides se succèdent le long des chaînes. La structure secondaire est déterminée par la forme de la chaîne formée par les différents amino-acides, tandis que la structure tertiaire peut être définie seulement, lorsqu'on connaît aussi la conformation, à savoir la forme définitive acquise par la molécule toute entière à travers des repliements particuliers. Ainsi, par exemple, dans les molécules de la myoglobine et de l'hémoglobine, les catènes polypeptidiques sont repliées d'une manière accentuée autour du groupe central, dit groupe prosthétique, ainsi qu'il résulte des modèles struc-

turels obtenus par KENDREW et PERUTZ respectivement, sur la base d'une étude roentgenologique très approfondie de leur structure.

Il convient de souligner, toutefois, que les résultats obtenus jusqu'ici représentent le fruit des recherches effectuées par des chimiste, des physico-chimistes, des physiciens, des biologistes et des physiologistes qui, grâce à leurs études spécifiques ont contribué à la connaissance des liaisons qui existent entre la structure moléculaire et l'activité biologique.

Presque tous les processus métaboliques sont catalysés par les enzymes qui, comme on le sait, sont constitués de molécules protéiques. A son tour, la synthèse des protéines est contrôlée par les gènes, et il est très probable que chaque gène contrôle la synthèse de chaque enzyme particulier et détermine la structure spécifique de la protéine.

Cependant la partie essentielle du gène n'est pas constituée par des protéines, mais par d'autres macromolécules, c. à d. par des acides nucléiques. Ces derniers pouvent être de deux types différents; l'un peut être l'acide désoxyribonucléique (DNA), et l'autre, l'acide ribonucléique (RNA); ce dernier est typique pour certains petits virus.

Au point de vue chimique, les chromosomes sont composés de DNA et de protéines, tandis que le RNA est présent seulement dans quelques stades du cycle de la reproduction cellulaire et il manque de façon absolue dans les spermatozoïdes.

Par conséquent, on peut exclure que le RNA ait la tâche de transmettre les premières informations héréditaires dans les organismes supérieurs. Pour pouvoir transporter l'information génétique, les acides nucléiques sont disposés de telle façon qu'ils constituent ce qu'on appelle le « chiffre chimique », formé d'une répétition régulière d'une certaine séquence qui est réitérée régulièrement sur toute l'extension de la macromolécule.

Chaque fois qu'une cellule se reproduit, il est nécessaire qu'il existe un « couple » d'informations génétique dont le but

est d'assurer à la « cellule fille » la continuité des caractères propres de la « cellule mère ».

Cependant, de nombreux doutes et beaucoup d'incertitudes empêchent encore l'élucidation de la succession compliquée des évènements aboutissant à la reproduction cellulaire ordonnée ou chaotique.

De toute façon, il existe des présuppositions qui permettent d'espérer que bientôt on pourra connaître à fond la reproduction et la structure du matériel génétique, comme on connaît déjà, du reste, la structure des systèmes moléculaires plus simples et le mécanisme de leur activité biologique.

La présente Semaine d'Etude se proposait donc de faire le point sur l'état actuel des connaissances dans ce domaine et elle poursuivait le triple but:

1) d'examiner les données acquises jusqu'à présent, afin d'établir les relations éventuelles entre la structure chimique et spatiale des macromolécules d'intérêt biologique, d'un côté, et les fonction spécifiques, dont elles sont chargées dans les êtres vivants, de l'autre;

2) de jeter les bases des recherches ultérieures aptes à élucider le mécanisme compliqué de l'interaction entre les différentes molécules d'intérêt biologique, afin de pouvoir, entre autres, trouver des moyens appropriés pour la lutte contre les maladies moléculaires et les affections virales;

3) d'attribuer au problème des systèmes macromoléculaires d'intérêt biologique la place qu'il mérite dans le plus vaste problème de la vie, sans abandonner, cependant, le domaine de la recherche expérimentale, en évitant ainsi les extrapolations faciles, ingénues et sans fondement qui exercent une si grande influence sur la formulation des théories concernant l'origine de la vie.

PIETRO SALVIUCCI
Chancelier de l'Académie

LA SEMAINE D'ETUDE
SUR LE PROBLEME DES MACROMOLECULES
D'INTERET BIOLOGIQUE
AVEC REFERENCE SPECIALE
AUX NUCLEOPOTEIDES

Le but des « Semaines d'Etude » de l'Académie Pontificale des Sciences a été ainsi défini par son premier Président, S.E. le Rév.me Père AGOSTINO GEMELLI O.F.M.:

« Tandis qu'on fixait, après sa fondation, les travaux de l'Académie, un problème se présenta bien vite avec évidence: les sciences posent chaque jour des problèmes nouveaux qui donnent lieu d'ordinaire à divers essais de solution, souvent contradictoires. Il arrive ainsi constamment que parmi les représentants les plus autorisés d'une science et, en particulier, entre ceux qui se sont consacrés à l'étude d'une même question, on rencontre des opinions opposées. De pareilles divergences se maintiennent parfois pendant de longues périodes et constituent à la fois une grave difficulté pour l'enseignement des sciences et fréquemment aussi un obstacle considérable à leur développement. D'ailleurs, l'expérience montre que les méthodes actuellement pratiquées dans la discussion des problèmes scientifiques, n'ont qu'une efficacité limitée au point de vue de l'établissement d'une unité de doctrine. Il serait hautement souhaitable de promouvoir tout ce qui pourrait favoriser une entente sur les points en discussion.

« Un tel procédé semble devoir être particulièrement utile sous ce rapport: savoir établir des contacts personnels prolongés entre quelques représentants d'opinions différentes au sujet d'une question déterminée ».

Dans ce but, l'Académie Pontificale des Sciences a réalisé une nouvelle « Semaine d'Etude » ayant pour titre: « Le Problème des Macromolécules d'intérêt biologique avec référence spéciale aux nucléoprotéides » (1).

Bien que ces derniers temps un travail intense ait été fourni sur les divers aspects de ce problème, il restait cependant quelques questions de détail à résoudre, et de nouvelles questions s'étaient de plus posées pendant ces dernières années.

Etant donné qu'on n'avait pas encore provoqué un débat approfondi à ce sujet et que le moment semblait propice pour le faire, l'Académie Pontificale des Sciences s'est proposée de réunir un nombre restreint de savants spécialistes de la question. Son but était de recueillir, au cours d'une discussion approfondie, les synthèses des nombreuses recherches effectuées dans ce domaine; de formuler clairement l'état des différents problèmes qui s'y rapportent; et par là de pouvoir fixer les directives de recherche les plus logiques, les plus persuasives et les plus prometteuses, étant donné l'état actuel de la science.

A cet effet ont été invités par l'Académie des experts qualifiés en chimie, en physico-chimie, en physique, en biologie et en phy-

(1) Cette « Semaine d'Etude » sur « Le Problème des Populations Stellaires » est la quatrième de la série.

La première « Semaine d'Etude » a eu lieu du 6 au 13 juin 1949; elle a été dédiée au « Problème biologique du Cancer », et a été présidée par l'Académicien Pontifical S. E. PIETRO RONDONI, Professeur de Pathologie Générale et expérimentale à l'Université de Milan; y ont participé personnellement 15 savants tandis que 3 autres ont envoyé des mémoires. Les comptes-rendus de la « Semaine d'Etude » ont été publiés dans le 7-ème volume des « scripta Varia » de l'Académie; ils représentent un volume de 364 pages.

La deuxième « Semaine d'Etude » a eu lieu du 19 au 26 novembre 1951; elle a été dédiée au « Problème des Microséismes », et a été presidée par l'Académicien Pontifical S. E. FRANCESCO VERCELLI, Directeur de l'Institut Thalassographique et de l'Observatoire Géophysique de Trieste; y ont participé personnellement 15 savants tandis que 4 autres ont envoyé des mémoires. Les comptes-rendus de la « Semaine d'Etude » ont été publiés dans le 12-ème volume des « Scripta Varia » de l'Académie; ils forment un volume de 466 pages.

La troisième « Semaine d'Etude » a eu lieu du 24 avril au 2 mai 1955;

siologic qui, grâce à leurs études spécifiques ont contribué à la connaissance des liaisons qui existent entre la structure moléculaire et l'activité biologique.

La présidence de cette « Semaine d'Etude » sur « Le Problème des Macromolécules d'intérêt biologique avec référence spéciale aux nucléoprotéides » a été confiée par le Président de l'Académie Pontificale des Sciences, S.E. le Rév.me Monseigneur GEORGES LEMAÎTRE, à l'Academicien Pontifical S.E. ARNE TISELIUS, Professeur de Biochimie à l'Université de Uppsala, et l'organisation générale au Chancelier de l'Académie Pontificale des Sciences Dr. PIETRO SALVIUCCI.

Ont été invités à la réunion les savants suivants:

Prof. Dr. CHRISTIAN BOHMER ANFINSEN, Directeur du « Laboratory of Cellulary Physiology and Metabolism, National Heart Institute ». - *Bethesda, Md.* (U.S.A.).

Prof. Dr. HUBERT CHANTRENNE, Professeur à l'Université libre de Bruxelles et directeur du Laboratoire de Chimie biologique. - *Bruxelles* (Belgique).

---

elle a été dédiée au « Problème des Oligoéléments dans la vie végétale et animale », et a été présidee par l'Académicien Pontifical, S. E. JOSÉ MARIA ALBAREDA HERRERA, Directeur de l'Institut de Pédologie et de Physiologie végétale de l'Université de Madrid, Secrétaire Général du Conseil Supérieur des Recherches Scientifiques d'Espagne; y ont participé personnellement 19 savants tandis qu'un autre a envoyé un mémoire. Les comptes-rendus de la « Semaine d'Etude » ont été publiés dans le 14-ème volume des « Scripta Varia » de l'Académie; ils forment un volume de 630 pages.

La quatrième « Semaine d'Etude » a en lieu du 20 au 28 mai 1957; elle a été dédiée au « Problème des Populations stellaires », et a été présidée par l'Académicien Pontifical Surnuméraire le Rév.me Père DANIEL J. K. O'CONNELL, Directeur de la « Specola Vaticana » de Castelgandolfo; y ont participé personellement 21 savants. Les comptes-rendus de la « Semaine d'Etude » ont été publiés dans le 16-ème volume des « Scripta Varia » de l'Académie; ils forment un volume de 615 pages.

L'organisation générale de chaque « Semaine d'Etude » à été confiée au Chancelier de l'Académie Pontificale des Sciences, le Dr. PIETRO SALVIUCCI. Toutes les réunions ont été tenues au Siège de l'Académie à la « Casina di Pio IV » dans les Jardins du Vatican.

Prof. Dr. ERWIN CHARGAFF, Directeur du Laboratoire de Chimie cellulaire et professeur de Biochimie à la Columbia University. « Cell Chemistry Laboratory, Department of Biochemistry, Columbia University ». - *New York, N.Y.* (U.S.A.).

S.E. Prof. PIETER JOSEF WILLIAM DEBYE, Académicien Pontifical, président du « Department of Chemistry » de la « Cornell University ». - *Ithaca, N.Y.* (U.S.A.).

S.E. Prof. GEORGE CHARLES DE HEVESY, Académicien Pontifical, professeur de Biochimie à l'Université. - *Stockholm* (Suède).

Prof. Dr. HEINZ L. FRANKEL-CONRAT, Chercheur virologiste au Laboratoire des Virus de la « California University » et professeur de virologie à la Section de Virologie. - *Berkeley, Calif.* (U.S.A.).

S.E. Prof. Dr. GIORDANO GIACOMELLO, Académicien Pontifical, professeur de Chimie pharmaceutique et directeur de l'Institut de Chimie pharmaceutique et toxicologique à l'Université. - *Roma* (Italie).

Prof. Dr. FRANÇOIS JACOB, Chef du Service de génétique microbienne de l'Institut Pasteur. - *Paris* (France).

Prof. Dr. EPHRAIM KATCHALSKI, Directeur du Service de biophysique au « Weizman Institute of Science ». - *Rehovoth* (Israël).

Prof. Dr. JOHN C. KENDREW, Vice-directeur du Groupe de biologie moléculaire du Conseil des recherches médicales près le « Cavendish Laboratory ». - *Cambridge* (Grande-Bretagne).

Prof. Dr. INGVAR LINDGVIST, Professeur chercheur adjoint à la « Uppsala Universitet Kemiska Institutionen ». - *Uppsala* (Suède).

Prof. Dr. FRITZ LIPMANN, Professeur de Chimie biologique et membre du « Rockefeller Institute ». - *New York, N.Y.* (U.S.A.).

Prof. Dr. ALFONSO MARIA LIQUORI, Directeur de l'Institut de Chimie physique de l'Université ». - *Napoli* (Italie).

Prof. Dr. MATTHEW MESELSON, Professeur de Biologie moléculaire à la « Harvard University ». - *Cambridge, Mass.* (U.S.A.).

S.E. Prof. Dr. PAUL SAN-ICHIRO MIZUSHIMA, Académicien Pontifical, directeur de l'Institut des recherches scientifiques « Yawata » de l'Université de Tokyo et professeur émérite de Chimie à la même Université. - *Tokyo* (Japon).

Prof. Dr. MAX FERDINAND PERUTZ, Directeur au « Medical Research Council Laboratory for Molecular Biology ». - *Cambridge* (Grande-Bretagne).

Prof. Dr. PAUL PUTZEYS, Professeur de Biochimie à l'Université catholique. - *Louvain* (Belgique).

Prof. Dr. SILVIO RANZI, Directeur de l'Institut de Zoologie de l'Université de Milan, professeur de Zoologie à la même Université. - *Milano* (Italie).

Prof. Dr. ALEXANDER RICH, Professeur de Physique biologique au « Massachusetts Institute of Technology ». - *Cambridge, Mass.* (U.S.A.).

Prof. Dr. ALESSANDRO ROSSI-FANELLI, Directeur de l'Institut de Chimie biologique de l'Université. - *Roma* (Italie).

Prof. Dr. MIGUEL RUBIO HUERTOS, Chercheur scientifique à l'Institut de microbiologie « Jaime Ferrán » du « Consejo Superior de Investigaciones Cientificas ». - *Madrid* (Espagne).

Prof. Dr. GERHARD SCHRAMM, Directeur du « Max-Planck-Institut für Virusforschung » de Tubingen et professeur à la même Université. - *Tubingen* (Allemagne).

Prof. Dr. NORIS SILIPRANDI, Directeur de l'Institut de Chimie biologique de l'Université de Padoue. - *Padova* (Italie).

Prof. Dr. SOL SPIEGELMAN, Professeur de Microbiologie à l'Université d'Illinois. - *Urbana, Ill.* (U.S.A.).

Prof. Dr. HUGO THEORELL, Professeur de Chimie et directeur du « Biokemiska Avdelningen, Medicinska Nobelinstitutet ». - *Stockholm* (Suède).

S.E. Prof. ARNE TISELIUS, Académicien Pontifical, Professeur de Biochimie à l'Université. - *Uppsala* (Suède).

Prof. Dr. HANS TUPPY, Professeur de Biochimie à l'Université. - *Wien* Autriche).

Prof. Dr. PAUL ZAMECNICK, Professeur d'Oncologie et directeur du « Warren Laboratories » à la « Harward University ». - *Boston, Mass.* (U.S.A.).

Le Président de la « Semaine d'Etude » a fait appel à son collègue l'Académicien S.E. GIORDANO GIACOMELLO pour l'organisation scientifique de la réunion à laquelle ont participé comme Secrétaires scientifiques: le Dr. CLAUDIO BOTRÉ en qualité de Chef du Groupe, le Dr. FRANCA ASCOLI, le Prof. GIOVANNI GIOVANNOZZI et le Prof. ALESSANDRO VACIAGO. Secrétaire du Groupe: Mlle ALPHONSINE GRUBER.

Le « Règlement des Semaines d'Etude » prescrivant que le nombre des Participants doit être rigoureusement limité, a malheureusement empêché d'inviter d'autres illustres savants.

Ont aussi participé à la réunion: en qualité d'interprète et chef de Secrétariat Mme VALENTINE PRÉOBRAJENSKI; en qualité de sténographes polyglottes de séance Mlles HERMIONE MONTANÈR et PAMELA SUTTON; en qualité de sténo-dactylographes polyglottes chargées des Procès-verbaux: Mlle VALERIA CRAJA, Mlle JOSÉPHINE LUCAS et Mme PAULETTE ROSSALDI; en qualité de technicien pour l'enregistrement et la projection, Mr MAURO ERCOLE, assisté par des opérateurs de Radio-Vatican. Le Bureau de Presse était confié au Dr. FRANCESCO SALVIUCCI, coadjuteur du Chancelier de l'Académie.

Le Comité de Réception pour les Dames dirigé par Mme HÉLÈNE LOTTI, était composé de la Comtesse KARINA CALVI DI COENZO. Mlle GIOVANNA LOTTI, Mme PAOLINA PUNZI.

Les séances de la « Semaine d'Etude » se tenaient deux fois par jour, le matin de 9 h. 30 à 12 h. 30 et l'après-midi de 16 h. à 19 h.; chaque séance était présidée par l'un des Participants à la réunion et les discussions des differents rapports se déroulèrent groupées selon l'affinité des sujets.

La réussite de la « Semaine d'Etude » a pleinement satisfait les illustres congressistes qui ont voulu envoyer les télégrammes suivantes:

« S.E. GEORGES LEMAÎTRE, Président de l'Académie Pontificale des Sciences. Cité du Vatican. — Les participants à la Semaine d'Etude sur les macromolécules d'intérêt biologique tiennent à exprimer au Président de l'Académie Pontificale des Sciences leur sincère et profonde gratitude pour les inoubliables journées fécondes en importants résultats scientifiques passées au cours de la Semaine d'Etude dans l'accueillante atmosphère de la Casina Pio IV dans un climat de parfaite liberté et indépendance. — ANFINSEN, CHANTRENNE, CHARGAFF, DEBYE, DE HEVESY, FRAENKEL-CONRAT, GIACOMELLO, JACOB, KATCHALSKI, KENDREW, LINDQVIST, LIPMAN, LIQUORI, MESELSON, MIZUSHIMA, PERUTZ, PUTZEYS, RANZI, RICH, ROSSI-FANELLI, RUBIO-HUERTOS, SCHRAMM, SILIPRANDI, SPIEGELMAN, THEORELL, TISELIUS, TUPPY, ZAMECNIK ».

« DR. PIETRO SALVIUCCI, Chancelier de l'Académie Pontificale des Sciences. Cité du Vatican. — Les Participants de la Semaine d'Etude sur les Macromolécules d'intérêt biologique remercient chaleureusement le Chancelier de l'Académie Pontificale des Sciences pour la parfaite hospitalité et les conditions idéales de travail qui leur ont été offertes grâce auxquelles ils espèrent avoir accompli un travail utile à la science et à l'humanité tout entière. — ANFINSEN, CHANTRENNE, CHARGAFF, DEBYE, DE HEVESY, FRAENKEL-CONRAT, GIACOMELLO, JACOB, KATCHALSKI, KENDREW, LINDQVIST, LIPMAN, LIQUORI, MESELSON, MIZUSHIMA, PERUTZ, PUTZEYS, RANZI, RICH,

ROSSI-FANELLI, RUBIO-HUERTOS, SCHRAMM, SILIPRANDI, SPIEGELMAN, THEORELL, TISELIUS, TUPPY, ZAMECNIK ».

A la fin de leurs travaux, les Participants à la « Semaine d'Etude » ont tenu à exprimer au Saint Père leur profonde gratitude et leur très sincère admiration pour cette manifestation scientifique si réussie, en envoyant à l'Auguste Pontife, animateur et mécène de l'Académie, le télégramme suivant:

« Sa Sainteté le Souverain Pontife JEAN XXIII. Cité du Vatican. — Les participants de la Semaine d'étude sur les macromolécules d'intérêt biologique, profondément attristés de n'avoir pu exprimer au Souverain Pontife leurs sentiments de profond respect, souhaitent un prompt rétablissement à Sa Sainteté. Ils prient Sa Sainteté de daigner accepter l'expression de leur admiration et de leur gratitude pour les conditions idéales de quiétude et d'indépendance qui leur ont consenti, grâce à Sa Munificence, de travailler et d'avoir en pleine liberté des échanges de vues hautement profitables au progrès de la science et de l'humanité. — ANFINSEN, CHANTRENNE, CHARGAFF, DEBYE, DE HEVESY, FRAENKEL-CONRAT, GIACOMELLO, JACOB, KATCHALSKI, KENDREW, LINDQVIST, LIPMAN, LIQUORI, MESELSON, MIZUSHIMA, PERUTZ, PUTZEYS, RANZI, RICH, ROSSI-FANELLI, RUBIO-HUERTOS, SCHRAMM, SILIPRANDI, SPIEGELMAN, THEORELL, TISELIUS, TUPPY, ZAMECNIK ».

A ce télégramme d'hommage et de remerciement, le Saint Père a daigné répondre par le message suivant, envoyé au Président de la « Semaine d'Etude » S.E. l'Académicien Pontifical ARNE TISELIUS et signé par Son Eminence le Cardinal GIOVANNI AMLETO CICOGNANI, Secrétaire d'Etat de Sa Sainteté.

« *Très touché sentiments exprimés par Participants Semaine Etude sur Macromolécules Intérêt Biologique Souverain Pontife renouvelle tout coeur remerciements pour message delicat et vifs regrets n'avoir pu accueillir personnellement illustres semainiers stop Se*

réjouissant échanges profitables de cette récente rencontre Sa Sainteté felicite volontiers Ses savants hôtes et encourage leurs labeurs en vue développement science véritable ».

Dans les pages qui suivent, après le « Règlement des Semaines d'Etude », sont imprimés les rapports originaux présentés à la Réunion, et les discussions qui les ont suivis.

Les « Conclusions » de la « Semaine d'Etude » se trouvent à la fin du présent volume.

Pendant la Semaine, les travaux scientifiques se poursuivirent sans interruption sauf le matin du jéudi 26 pour visiter les Musées Pio-Clementino, Chiaramonti, Etrusque, Egyptien; le Braccio Nuovo; les Galeries des tapisseries, des cartes geographiques; les Stances et lle Loggia de Raphaël, la Chapelle de Fra Angelico, la Chapelle Sixtine, l'Appartament Borgia et la Pinacothèque Vaticane avec l'assistance du Prof. Comm. FILIPPO MAGI et du Dr. DEOCLECIO REDIG DE CAMPOS de la Direction Générale des Monuments de la Cité du Vatican.

Ils ont aussi visité la Bibliothèque Apostolique Vaticane et les Archives Secrètes Vaticanes sous la conduite des Rév.mes Préfets Mons. MARTINO GIUSTI et Père Abbé ANSELMO ALBAREDA, O.S.B.; et la Station de Radio-Vatican qui leur fut présentée par le Directeur, le Rév.me P. ANTONIO STEFANIZZI S.I.

Au cours de la même journée de jeudi 26, les Participants et leurs épouses dinèrent a Castelgandolfo, où les savants étaient attendus par le Dr. Gr. Uff. EMILIO BONOMELLI, Directeur des Villas

Pontificales qui leur a fait visiter en détail le Palais des Papes, les ruines de la villa romaine de Domitien, l'ex-Palais des Barberini, et les autres dépendances.

L'après-midi les Dames visitèrent les Jardins de l'enceinte extra-territoriale suburbaine du Vatican, tandis que les savants poursuivirent leurs travaux dans une salle du Palais Barberini.

Le samedi 28 octobre il y avait exactement vingt-cinq années que ce grand mécène des hautes études que fut le Souverain Pontife Pie XI (Achille Ratti) décida de fonder l'Académie Pontificale des Sciences avec le but précis d'honorer la science pure, où qu'elle se trouve, d'en assurer la liberté et d'en favoriser les recherches qui constituent la base indispensable du progrès des sciences appliquées.

Et c'est, en effet, par la date du 28 octobre 1936 qui précède la signature du grand Pontife, que termine le document officiel, à savoir le *Motu Proprio* « *In multis solaciis* », au moyen duquel il créa l'Académie Pontificale des Sciences, placée sous la haute et directe protection du Souverain Pontife et composée de soixante-dix Académiciens Pontificaux appartenant à quelque Nation et confession religieuse que ce soit et choisis parmi les spécialistes les plus insignes en sciences mathématiques et expérimentales du monde entier.

C'est pourquoi, le Souverain Pontife avait choisi cette journée pour recevoir tous les Académiciens Pontificaux présents à Rome ainsi que les Participants à la Semaine d'Etude en une Audience Solennelle au Palais Apostolique du Vatican en présence des Cardinaux se trouvant à Rome et du Corps Diplomatique accrédité près le Saint-Siège.

Après l'Audience Pontificale aurait dû avoir lieu à la Casina de Pie IV une Séance Extraordinaire pour commémorer le vingt-cinquième anniversaire de la fondation de l'Académie.

Malheureusemente, au dernier moment, une légère indisposition empêcha le Souverain Pontife de réaliser son projet et faute

d'Audience, Il eut l'amabilité de faire envoyer le texte du discours qu'Il s'était promis de prononcer et qui fut lu par le Président de l'Académie et suivi avec une attention recueillie par tous les assistants (voir à la pag. XXXI).

Ensuite eut lieu la Séance Extraordinaire pour commémorer le vingt-cinquième anniversaire de la fondation de l'Academie.

L'orateur officiel fut l'Academicien S.E. Sir HUGH STOTT TAYLOR, un des premièrs soixante-dix savants choisis pour cette fois là personellement par le Pape Fondateur.

Suivirent les différents discours des Académiciens S.E. GIAMBATTISTA BONINO, S.E. OTTO HAHN, S.E. EDUARDO CRUZ COKE et S.E. PAULO-SAN-ICHIRO MIZUSHIMA.

Le Président de la « Semaine d'Etude » S.E. ARNE TISELIUS voulut ensuite exprimer de vive voix toute le gratitude et les remerciements des Participants a cette réunion scientifique qui fut si bien réussie.

Le lundi 30 octobre le Souverain Pontife, qui s'était entretemps rétabli, voulu, par un geste de paternelle bienveillance,, recevoir les savants qui demeuraient encore à Rome.

La rencontre eut lieu en une salle de l'Appartement Pontifical au Palais du Vatican.

Ce fut une Audience certes hors de l'ordinaire par l'atmosphèere d'intime recueillement dans laquelle le Sain Père a tenu à la conduire quasi comme une compensation en l'absence de la plus grande solennité qui était en programme.

Le Présidente LEMAITRE à cette occasion s'adressa au Saint Père en ces termes:

« Très Saint Père, laissez-moi vous exprimer la joie qui rempli nos coeurs en voyant combien rapidement, et, nous l'espérons, combien complètement Vous avez pu Vous remettre de l'indisposition qui nous a privé de l'Audience Solennelle que Vous nous aviez réservée.

« Vous avez bien voulu me permettre de prononcer les augustes paroles que Vous aviez préparées à cette occasion. Elles ont été un précieux encouragement pour notre Académie, en ce jour de son vingtcinquième anniversaire. Elles resteront longtemps dans notre mémoire comme un gage de paternelle bonté et de sollicitude pour notre Académie Pontificare des Sciences que Vous continuez à combler avec munificence suivant l'exemple de vos illustres prédécesseurs le Pape Pie XI, fondateur, dont Vous avez donné le nom à la médaille d'or que Vous nous avez donné de décerner au Professeur Woodward, et le Pape Pie XII d'immortelle mémoire qui l'a conduite à son plein développement.

« Je puis Vous dire que malgré la grande déception qui nous a privé de Votre auguste présence, notre Session Plénière, et plus encore la Semaine d'Etude sur les ''Macromolécules d'intérêt biologique'' que Vous nous avez donné d'organiser a été un plein succès. Je suis l'interprète de tous les ''Participants'' en Vous exprimant leur profonde reconnaissance.

« Les quelques privilégiés qui ont pu prolonger leur séjour à Rome et sont ici présents ont été particulièrement sensibles au bonheur qu'ils Vous doivent de pouvoir Vous exprimer eux-mêmes leurs sentiments de gratitude et de filial attachement ».

Le Souverain Pontife répondit alors aux paroles du Président LEMAITRE en Se félicitant des manifestations scientifiques de Son Académie qui venaient de se dérouler et Se plut à raconter, par des expressions simples mais efficaces, certains épisodes qui lui etaient arrivés et dont Il rattachait familièrement le souvenir aux activités des savants présents.

Le Saint-Père S'entretint ensuite avec le Président LEMAITRE, les

Académiciens Pontificaux, le Chancelier SALVIUCCI et les savants « Participants » à la « Semaine d'Etude » trouvant pour chacun d'aimable paroles de félicitations et de souhaits.

L'assistance exprima enfin ses remerciements au Saint-Père pour le don inespéré qu'Il avait daigné leur faire et — avec le regret pour les Collègues déjà partis — renouvela au Souverain Pontife sa reconnaissance émue et sa profonde gratitude.

Enfin, le soir di même samedi 28 octobre, un diner d'adieu a été offert, selon le contûme, aux savants participants a la « Semaine d'Etudes ».

LE DISCOURS DU SAINT-PERE

Messieurs,

Nous sommes heureux de recevoir aujourd'hui pour la première fois le nouveau Président et les membres de votre illustre et docte assemblée. Ce jour anniversaire de Notre Election se trouve être aussi celui de la consécration épiscopale de Pie XI, le sage fondateur, ou pour parler plus exactement, le restaurateur de cette Académie qui s'honore du titre d'Académie Pontificale. Notre contentement est d'autant plus grand que vos rangs se sont ouverts récemment à d'autres personnalités de plusieurs pays, que recommandent leurs hauts mérites et l'étendue de leurs connaissances scientifiques.

Répondant à Notre invitation cordiale, vous vous êtes rassemblés depuis plusieurs jours pour tenir une session plénière ainsi qu'une nouvelle semaine d'étude sur la strutture des macromolécules d'intérêt biologique.

Laissez-Nous vous remercier de tout coeur de l'empressement avec lequel, malgré vos nombreuses occupations, vous avez répondu à cette invitation, et vous dire la fierté qu'éprouve l'Eglise de voir réunie dans la Cité Vaticane pour de savants échanges de vue une pareille élite de chercheurs.

Vous constituez vraiment, Messieurs, par la diversité de vos origines et par la variété de vos compétences, comme un reflet fidèle du monde savant contemporain et vous attestez l'accord complet qui a toujours existé entre l'Eglise et la véritable science.

Ce n'est pas, vous le savez, par souci de fidélité à des traditions humanistes héritées de la Renaissance, que l'Eglise vous accueille. C'est dans la conscience d'accomplir là une partie de sa mission constante de mère et d'éducatrice. Partout où elle s'est implantée, elle a toujours imprimé un élan remarquable au développement de la culture intellectuelle.

Tel était d'ailleurs le noble dessein que poursuivait Notre Prédécesseur Pie XI en fondant, voici exactement un quart de siècle, l'Académie Pontificale des Sciences. Ce but l'avait conduit à insérer dans le texte du motu proprio *de l'institution la lumineuse affirmation du premier Concile du Vatican sur les rapports entre la foi et la raison, qu'il Nous plaît de rappeler ici: « Non seulement la foi et la raison ne peuvent jamais s'opposer, mais elles s'apportent une aide réciproque ». Et le même Pontife en concluait: « C'est Notre souhait ardent et Notre ferme espérance que par cet Institut, qui est à la fois le Nôtre et le leur, les Académiciens Pontificaux contribuent toujours davantage et mieux au progrès des sciences. Nous ne leur demandons rien d'autre, car en ce dessein généreux et ce noble labeur consiste le service, en faveur de la vérité, que Nous attendons de leur part »* (1).

(1) Motu Proprio *in Multis Solaciis*, 28 oct. 1936; A.A.S. 28, p. 421.

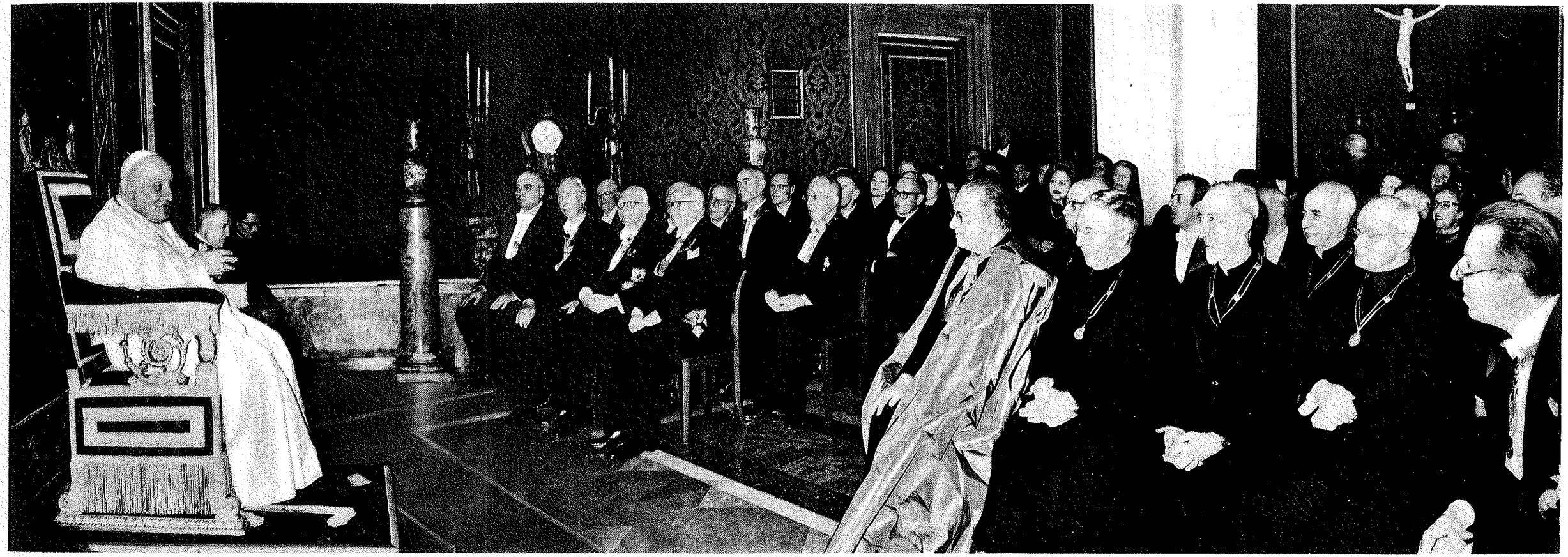
Cette attente est aussi la Nôtre, soyez-en bien persuadés. Vous savez tous l'importance que Nous attachons personnellement au travail intellectuel et à l'investigation scientifique. Nous avons toujours eu à coeur d'employer les loisirs que Nous laissaient Nos diverses fonctions à des recherches d'ordre pastoral et historique. Aussi est-ce avec joie que Nous avons salué en son temps la création de votre Académie. Dans cette perspective, Nous tenons à évoquer aujourd'hui la mémoire de son premier Président, Agostino Gemelli, qui s'est montré fidèle à l'admirable programme tracé par votre fondateur: sa vie a honoré l'Eglise et la science. Il Nous est aussi agréable, d'autre part, d'apprécier à sa valeur l'heureux choix que votre illustre compagnie a fait en conférant la médaille d'or de Pie XI au Professeur Robert Burns Woodward, de l'Université de Harvard, dont l'activité scientifique dans le domaine chimique est pleine de réalisations et de promesses.

En effet, bien loin de redouter les découvertes les plus audacieuses des hommes, l'Eglise estime au contraire que tout progrès dans la possession de la vérité appelle un épanouissement de la personne humaine, constitue un acheminement vers la Vérité première ainsi qu'une glorification de l'oeuvre créatrice de Dieu. La Sainte Ecriture revient souvent sur ces importantes pensées et sans doute vous arrive-t-il parfois. Messieurs, de laisser chanter en vous, dans l'enthousiasme de la recherche et de la découverte, le magnifique cantique que rapporte le livre de Daniel: « Benedicite omnia opera Domini Domino » (1).

(1) Dan. 3, 57.

A la suite des trois enfants d'Israël saisis d'admiration devant les merveilles de la nature, comment ne pas appeler les anges, les astres et les éléments, les animaux, les plantes et les minéraux, les hommes les plus saints et les plus écoutés de Dieu à interpréter nos sentiments de louange au Créateur?

Dans l'assurance que vous collaborez de toutes vos forces à cette entreprise du savoir et de la louange, Nous invoquons volontiers sur vos travaux et vos personnes, en gage de Notre paternelle bienveillance, une large effusion des divines bénédictions.



LES "SEMAINES D'ETUDE"
ET LEUR REGLEMENT

Lorsque l'Académie Pontificale des Sciences fut fondée par le Souverain Pontife Pie XI, de vénérée mémoire, par son « Motu Proprio » du 28 octobre 1936 « In multis solaciis », cette initiative suscita dans les milieux scientifiques un mouvement général de sympathie et d'admiration. Cette institution unique au monde, qui groupait en une même assemblée des représentants de toutes les Nations civilisée, était appelée, en effet, à de hautes destinées dans le développement de la pensée scientifique.

D'autre part, cette oeuvre de coopération fut accueillie avec un véritable soulagement par tous ceux que plongeait dans le désarroi le plus profond la période qui suivit la guerre 1914-18. On voyait, en effet, s'altérer profondément les caractères d'objectivité et de desintéressement propres au travail scientifique, et s'affirmer même une tendance à asservir la science à des fins pragmatiques.

Tout au contraire, dans l'immortel « Motu Proprio » du 28 octobre 1936, le Pape Pie XI proclamait solennellement la dignité de

---

A general movement of sympathy and admiration was aroused in scientific circles when, in 1936, the Pontifical Academy of Science was founded by His Holiness Pope Pius XI, of venerable memory, by means of his « Motu Proprio » of October 28, « In Multis solaciis ». This institution, the only one of its kind in the world, which brought the representatives of all civilized nations into touch with each other, was, in fact called upon to play a leading role in the development of scientific thought.

This work of cooperation was, moreover, welcomed with a sense of real relief by all those who were plunged in a deep state of confusion in the period following the 1914-18 war.

Signs of drastic changes were, in fact, discernible in the objective and disinterested nature of scientific work and even a tendency to make science subject to pragmatic aims.

la recherche de la vérité pour elle-même (*), et, élevant sa pensée au-dessus de toute préoccupation utilitaire, affirmait qu'il ne demandait rien d'autre aux nouveaux « Académiciens Pontificaux » que de se consacrer, avec une ferveur toujours plus grande, au progrès de la science et, par là, au culte de la vérité: « C'est Notre souhait ardent et Notre ferme espérance: que par cet Institut, à la fois Notre et leur, les "Académiciens Pontificaux" contribuent toujours plus et mieux au progrès des sciences. Nous ne leur demandons pas autre chose; car en ce dessein généreux et ce noble labeur consiste le service, qu'en faveur de la vérité, nous attendons de leur part ».

La consécration pratique de cette idée, par la nomination d'un certain nombre de non-catholiques parmi les nouveaux Académiciens Pontificaux a fait une profonde impression sur beaucoup d'esprits, comme l'ont montré les réactions de la presse internationale de l'époque et de nombreux témoignages individuels d'hommes de science et des plus grands savants du monde.

(*) « Nobis autem in votis expectationeque est, fore ut "Pontificii Academici" vel per hoc Nostrum suumque studiorum Institutum, ad scientiarum progressionem fovendam amplius excelsiusque procedant; ac nihil praeterea aliud petimus, quandoquidem hoc eximio praeclaroque labore famulatus ille nititur servientium veritati, quem ab iisdem postulamus ».

In his immortal « Motu Proprio » of October 28, 1936, Pope Pius XI, on the contrary, solemnly proclaimed the dignity of the search for truth for its own salce (*) and, raising his thoughts above all preoccupations of a utilitarian nature, asserted that all he asked of the new « Pontifical Academy » and its members was that they should dedicate themselves, with increasing fervour, to the furthering of the progress of science and, consequently, to the cult of truth: « It is Our ardent wish and firm hope that, by means of this Institute, which is both Ours and theirs, the "Pontifical Academicians" will contribute to an increasingly great extent to the progress of science. We ask nothing more than that from them because the service in favour of truth that We expect from them consists in this generous intention and noble work ».

By including a certain number of non-Catholics amongst the new Pontifical Academicians, the practical application of this idea made a deep impression on many persons, as is proved by the reaction of the interna-

Beaucoup de préjugés à l'égard de l'Eglise ont été fortement ébranlés par ce geste du Souverain Pontife qui a obligé à reconnaître la place éminente réservée aux valeurs purement intellectuelles dans l'Eglise Catholique.

Pour toutes ces raisons, la fondation de l'Académie Pontificale des Sciences à été hautement appréciée dans le monde scientifique et y a fait naître de grands espoirs quant aux possibilités d'action d'une institution si opportune.

Le Saint-Père Pie XII, qui avait collaboré avec son Prédécesseur au projet et à la fondation de l'Académie et qui l'avait représenté comme Légat personnel lors de l'inauguration solennelle, ne s'est pas borné à maintenir à son égard ses sentiments de haute estime par sa présence à de solennelles séances académiques, où il daigna prononcer ses discours d'une haute portée scientifique; il a tenu en outre à lui donner un nouveau témoignage de son auguste satisfaction en accordant à ses membres le titre d'Excellence par le Bref Apostolique du 25 novembre 1940.

* * *

Les sciences posent chaque jour des problèmes nouveaux qui donnent lieu d'ordinaire à divers essais de solution, souvent contra-

---

tional press of the time and by the innumerable individual tributes paid by scientists and by the greatest scholars of the world.

Many prejudices against the Church were very deeply shaken by this gesture on the part of the Sovereign Pontiff, since it called attention to the lofty place reserved for purely intellectual values in the Catholic Church.

For all these reasons, the foundation of the Pontifical Academy of Science was greatly appreciated by the scientific world and aroused high hopes as to the prospects open to such a timely institution.

His Holiness Pope Pius XII, who had helped his predecessor to draw up the plan and to found the Academy, and who had represented Him as His personal Legate at the time of its solemn inauguration, did not confine himself to the expression of lofty sentiments when attending solemn academic gatherings, where he deigned to make speeches of great scientific importance, but he also afforded proof of his august satisfaction by granting the title of Excellency to the members of the Academy, by an Apostolic Brief of November 25, 1940.

dictoires. Il arrive ainsi constamment que parmi les représentants les plus autorisés d'une science, et en particulier parmi ceux qui se sont consacrés à l'étude d'une même question, on rencontre des opinions opposées. Pareilles divergences se maintiennent parfois durant de longues périodes et constituent à la fois une grave difficulté pour l'enseignement des sciences et fréquemment aussi un obstacle considérable à leur développement.

Par ailleurs, l'expérience montre que les méthodes actuellement pratiquées dans la discussion des problèmes scientifiques n'ont qu'une efficacité limitée au point de vue de l'établissement d'une unité de doctrine.

Il serait dès lors hautement souhaitable de promouvoir tout ce qui pourrait favoriser un accord sur les points en discussion.

Un procédé semble devoir être particulièrement utile sous ce rapport: à savoir, l'établissement de contacts personnels prolongés entre quelques représentants d'opinions différentes au sujet d'une question déterminée.

En effet, le contact personnel entre hommes de science constitue, sans aucun doute, le moyen le plus efficace de résoudre les controverses scientifiques.

Dans ce but, l'Académie Pontificale des Sciences a décidé d'organiser de pareilles rencontres scientifiques. L'organisation de ces ren-

---

* * *

Every day science raises new problems, which usually give rise to various, and often contradictory, solutions. Consequently it often happens that amongst the most authoritative representatives of a given branch of science, and particularly amongst those who are engaged in studying the same question, one meets with contrasting opinions. Divergences of this kind often exist over long periods of time and are a serious obstacle not only to the teaching of science but also to its development.

Experience shows, moreover, that the methods at present in use in the discussion of scientific problems have only a limited efficacy in so far as concerns doctrinal unity.

It would, therefore, be highly desirable if everything that could favour agreement on controversial points were to be promoted.

contres qu'on a appelées « Semaines d'Etude » a été établie de la manière suivante:

RÈGLEMENT DES SEMAINES D'ÉTUDE

1. - L'Académie invite quelques illustres savants, parmi ceux qui, ayant étudié spécialement une question déterminée, sont arrivés à des conclusions différentes, à se rencontrer à Rome, à son siège, la « Casina di Pio IV », à l'intérieur de l'Etat de la Cité du Vatican, afin d'y procéder en commun, en dehors de toute autre préoccupation, un examen général de toutes les données du problème.

2. - Le but essentiel de ces discussions est de chercher à formuler de façon précise les raisons qui sont à la base de la divergence des opinions. Les savants conviés aux réunions s'engageraient d'avance à concentrer leurs efforts dans cette direction.

3. - Un examen critique de ces raisons aboutira soit à un accord sur une solution déterminée, soit à la constatation qu'à l'état actuel des connaissances, il est impossible d'établir une unité de doctrine au sujet du problème envisagé.

One process that would seem to be particularly useful from this point of view would be the establishment of prolonged personal contacts between some of the representatives of different trends of thought on a given subject.

Personal contacts amongst scientists are, in fact, the most efficacious means of solving scientific controversies.

With this aim in mind, the Pontifical Academy of Science decided to organize scientific meetings of this description. These meetings, known as « Study Weeks », were planned on the following lines:

STANDING RULES FOR « STUDY WEEKS »

1. - The Academy invites a number of illustrious scholars — comprising those who have especially studied a given question and have arrived at different conclusions — to meet in Rome at its headquarters, the « Casina di Pio IV », situated in the Vatican City, so as to make a joint examination, free from all other preoccupations, of all data concerning the problem.

2. - The chief aim of these discussions is to endeavour to formulate precisely the reasons which are at the root of the differences of opinion.

Dans ce dernier cas, les savants invités auront pour tâche:

a) de préciser les motifs pour lesquels un accord s'avère présentement irréalisable;

b) de définir le genre de recherches qu'il serait souhaitable d'entreprendre en vue de résoudre la question.

4. - L'invitation ne sera adressée par l'Académie qu'à un très petit nombre de représentants de chaque science: ceux-ci seront choisis parmi les personnalités étrangères à l'Académie, auxquels se joindront, dans la discussion, les Académiciens versés dans la même discipline. Cette invitation, de plus, ne se rapportera qu'à l'étude d'une question déterminée, pour chaque science.

5. - Les discussions auront un caractère strictement privé; elles prendront la forme de conversations particulières, sans autre assistance que celle de quelques membres de l'Académie Pontificale des Sciences particulièrement compétents dans la matière.

Des interprètes polyglottes, des sténographes, des rapporteurs, etc., seront mis à la disposition des savants réunis.

6. - Les « Conclusions » des discussions seront publiées sous la

---

The scholars invited to these meetings undertake in advance to concentrate their efforts on this.

3. A critical examination of these reasons should lead, either to agreement on a given solution or else to the conclusion that, on the basis of the information actually available, it is impossible to establish doctrinal unity on the problem envisaged.

In the latter event the scholars concerned will be called upon:

a) to define the reasons why agreement appears to be impossible for the present;

b) to specify the kind of research work it would be desirable to undertake with a view to solving the problem.

4. - The invitation will be addressed by the Academy to only a small number of representatives of each branch of science: these will be selected from amongst those who are not connected with the Academy. They will be joined during the discussions by Academicians versed in the same discipline. This invitation, moreover, will apply only to the study of one precise problem in each branch of science.

forme d'une « Note Collective Finale » (à laquelle pourront éventuellement être jointes des annotations individuelles), mentionant :

a) les points sur lesquels un accord aurait été réalisé;

b) les points sur lesquels un accord n'aurait pas paru réalisable;

c) les raisons pour lesquelles l'accord n'aurait pu être réalisé;

d) des suggestions relatives aux recherches paraissant les plus aptes à résoudre les difficultés.

7. - Les « Conclusions » seront aussitôt imprimées et communiquées, par le soins de l'Académie Pontificale des Sciences, à tous les centres scientifiques qu'elles seraient de nature à intéresser.

8. - Tous les frais de voyage et de séjour à Rome des personnalités invitées seront à la charge de l'Académie Pontificale des Sciences. L'hospitalité sera assurée dans l'un des principaux hôtels de Rome.

L'Académie se fera un plaisir d'offrir la même hospitalité aux épouses des savants invités, à l'exclusion toutefois des frais de voyage.

---

5. - The debates will be strictly private and will take the form of personal talks, in the presence only of a few members of the Pontifical Academy of Science with special knowledge of the subject under discussion.

Polyglot interpreters, stenographers, reporters, etc. will be placed at the disposal of the participants.

6. - The « Conclusions » arrived at will be published in the form of a « Collective Note » (to which may eventually be added individual notes) mentioning:

a) the points on which agreement was reached;

b) the points on which it was impossible to reach agreement;

c) the reasons why it was not possible to reach agreement;

d) suggestions regarding the research work which appears most suitable for arriving at a solution of the difficulties.

7. - The « Conclusions » reached will be immediately printed and transmitted, by the Pontifical Academy of Science, to all the scientific centres which might be interested therein.

8. - All travelling expenses, and accommodation in one of the best hotels in Rome, of the persons invited to the meetings will be borne by the Pontifical Academy of Science.

The Academy will be pleased to offer similar accommodation to the wives of the scholars who are invited, but not their travelling expenses.

TRAVAUX SCIENTIFIQUES

INTRODUCTION

This study week organized by invitation of the Pontificia Academia Scientiarum » has as subject of the discussion: *Macromolecules of biological interest with special reference to nuclear proteins.*

This field is to day, no doubt, one of the most important domains in the whole of science.

The reason is, I believe, that at this section of the rapidly advancing scientific frontier, we have perhaps come closer to an understanding of some of the fundamental processes of life than in any other field.

During this week we have discussed how the giant molecules of, for example, proteins and nucleic acids which govern life, are built up, that is to say their structure in detail, how they look and how they behave in the living organism, how they are formed and how they interact.

Some of these substances are of fundamental importance for reproduction and for inheritance.

Others may be dangerous, contagious agents which may even destroy life.

It is rather interesting that one particular useful material in this work is found among the so-called bacteriophage and

among the viruses which, although they may destroy life, are able to give us perhaps the most useful information to day in this field.

I shall not bore you with technical details about these interesting phenomena here.

I shall have to point out, however, that in these fields there is at present a terrific concentration of intellectual effort all over the world.

It is therefore particularly gratifying that the Pontifical Academy of Sciences at this time has called upon a selected group of specialists from many different countries to discuss their problems under the auspicies of the Pontifical Academy.

This week has been extremely stimulating to us all, and I can only say that we agree in many, perhaps in most things.

We should also, according to the rules of these study weeks, report where we disagree, and why.

Of course, we disagree in many points.

If we did not disagree, we should not be scientists, and especially not scientists working at the frontiers.

I think we are also supposed to recommend how this disagreement should be resolved, how we should come to an agreement. And there the reply is very simple: just more experiments.

It now remains for me, to express our deeply felt thanks to His Holiness the Pope, to the Academy and to all those who have helped to organize this meeting, for their most generous aid and hospitality as well as for the initiative for this important meeting and for all the work that has been done to make this week a unique experience.

ARNE TISELIUS
Pontifical Academician

FACTORS INFLUENCING THE FORMATION AND MAINTENANCE OF THE SECONDARY AND TERTIARY STRUCTURE OF PROTEINS

CHRISTIAN B. ANFINSEN

Laboratory of Cellular Physiology and Metabolism, National Heart Institute, Bethesda, Md. (U.S.A.)

It has now been well documented, both theoretically and by direct experimentation, that polypeptide chains do not automatically assume specific three-dimensional structures in solution in the absence of stabilizing intrachain cross-links. Thermodynamic calculations made by John SCHELLMAN [1] indicate that the helical and random chain forms exist with nearly equal probability under conditions of aqueous solution at neutral pH and moderate ionic strength. Considerations of this kind make it clear that the exactly defined geometry of proteins in solution is determined by an equally exactly defined set of interactions which may be either covalent, non-covalent, or both.

In general terms, covalent cross-linkages endow protein molecules with stability toward denaturing conditions. Classical examples of such very stable proteins are ribonuclease and egg-white lysozyme, both of which contain a number of disulfide bridges and are devoid of free SH groups. (The presence of free SH groups in a protein which is otherwise highly cross linked introduces a special case of instability, probably due to the pos-

sibilities for disulfide interchange. Egg albumin is a good example of an easily denatured protein in this category.)

Covalent cross-linkages are not necessarily a prerequisite for intrinsic stability to denaturing conditions. Another lysozyme, namely that from bacteriophage T4, contains no disulfide bridges and is apparently a single chain dependent entirely upon its *non*covalent interactions for the determination of its three-dimensional configuration [2]. After unfolding in strong urea solution, for example, this molecule can find its way back to the native, catalytically active configuration upon removal of the denaturing agent.

I would like to consider in this brief lecture some of the biological principles that may underly the elegance of « design » of proteins and some of the structural factors that lead to what appears to be an invariably correct geometry. My interest in this area stems from experiments that we have been carrying out over the past four or five years on the « predetermination » of secondary and tertiary structure in ribonuclease and, more recently, on several other proteins [3]. These studies have employed as a major tool the process of complete reduction of disulfide bridges followed by reoxidation of the resulting sulfhydryl groups under a variety of conditions. For orientation, Figure 1 shows the covalent structure of ribonuclease, which contains four disulfide bridges. If, as appears more and more likely in the light of contemporary studies on protein biosynthesis, the polypeptide chain is formed as a random coil, possibly oriented on a template surface and subsequently cross-linked by pairing of half-cystine residues, one would guess that biosynthesis involves the conjugation of cysteine rather than cystine.

In a typical experiment, the protein is reduced by incubation in 8 M urea in the presence of a considerable excess of mercaptoethanol. After acidification to pH 3, the reduced protein is isolated by passage through a Sephadex column [4]. Experiments may now be carried out on this material to de-

[1] *Anfinsen* - pag. 2

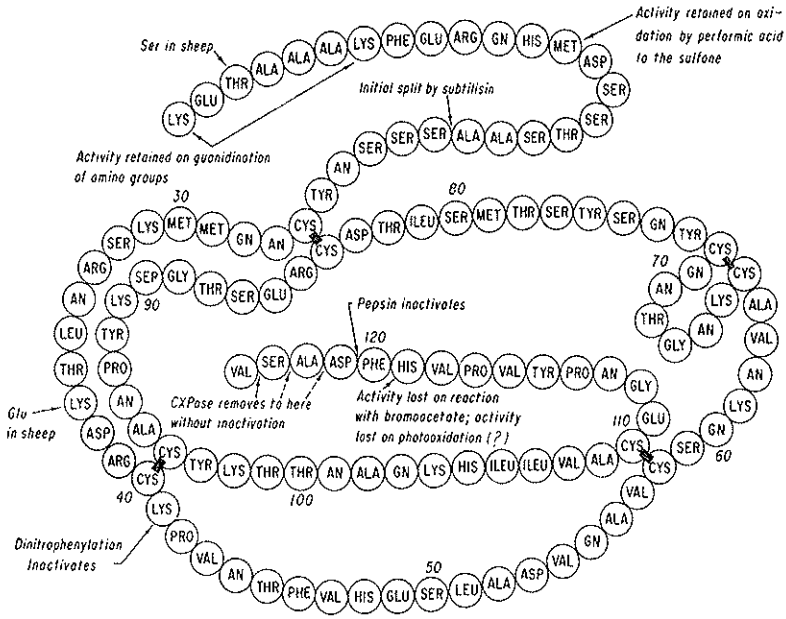


FIGURE 1 — The covalent structure of bovine pancreatic ribonuclease. The studies leading to this structure are described in the following articles: C.H.W. HIRS, S. MOORE and W.H. STEIN, *J. Biol. Chem.*, *235*, 633 (1960);
D.H. SPACKMANN, W.H. STEIN and S. MOORE, *J. Biol. Chem.*, *235*, 648 (1960);
R.R. REDFIELD and C. B. ANFINSEN, *J. Biol. Chem.*, *221*, 385 (1956);
C. B. ANFINSEN and F. H. WHITE, Jr., in *The Enzymes*, Second Edition (P.D. Boyer, ed.), Vol. 5, pp. 95-147, New York, Academic Press, 1961;
D.G. SMYTH, W.H. STEIN, and S. MOORE, *J. Biol. Chem.*, *237*, 845 (1962);
J.T. POTTS, A. BERGER, J.P. COOKE and C. B. ANFINSEN, *J. Biol. Chem.*, *237*, 1851 (1962).

termine, first of all, the adequacy of the « genetic information » coded into the primary sequence for the determination of proper pairing of SH groups and, secondly, the nature of the environmental factors which can disorient this pairing. As shown in Table I, essentially complete re-formation of the native enzyme molecule can be achieved under conditions of proper acidity

TABLE I

OXIDATION OF REDUCED RNASE OF VARIOUS PROTEIN CONCENTRATIONS

The solution of reduced protein was adjusted to pH 8,5 and allowed to stand without stirring at room temperature for 20 hours, at which time the proportion of soluble protein and the enzyme activity of this fraction were estimated.

Concentration of reduced protein	% Yield of soluble protein	Activity of soluble fraction; % of equivalent concentration of native RNase	% Regeneration of activity
mg/ml			
7.0	27	31	8
4.8	42	70	29
2.3	87	75	65
0.9	100	77	77
0.35	100	94	94

and protein concentration [4]. In more concentrated solutions, only a fraction of the reoxidized mixture is soluble. The insoluble material probably contains disulfide-linked trimers, as suggested by ultracentrifuge studies in strong urea solutions. Some dimers are present in the soluble fraction but in amounts that approach zero with the use of more and more dilute protein solutions during the oxidation process. Dr. Fred WHITE [5] has examined a number of physical properties of the re-formed native protein, and preliminary X-ray crystallographic studies have been made on this material by BELLO and HARKER [6]. The protein isolated from the reoxidation mixtures appears to be native in all respects and shows full catalytic activity.

We have recently examined some of the physical events occurring during various stages of reoxidation [7] and some of these are summarized in Figure 2. Optical rotatory properties

[1] *Anfinsen* - pag. 4

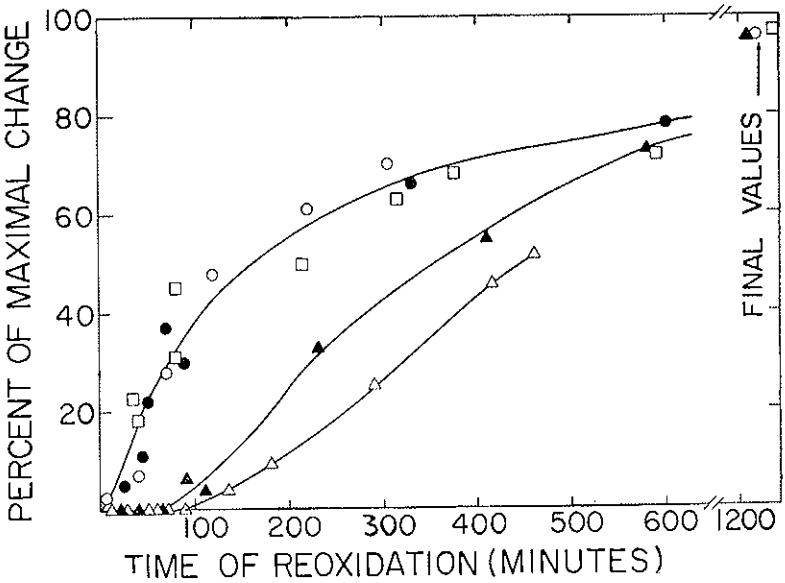


FIGURE 2 — Changes, during the oxidation of reduced ribonuclease, in SH groups as followed by titration with *p*-chloromercuribenzoate (●) and by reaction with radioactive iodoacetate (0), in optical rotation (□), and in enzymatic activity as measured against ribonucleic acid (▲) and against uridylic-2′,3′-cyclic phosphate (△).

and the disappearance of free SH groups follow a smooth first-order-like curve in spite of the fact that no activity appears during the considerable lag phase shown in the figure. Dr. WHITE and I have recently studied, in a preliminary way, the pairing of half-cystine residues at various times during the reoxidation. It is possible to state with reasonable certainty that the absence of activity during the early period is attributable to the relatively random pairing of SH groups. These « incorrectly formed » bridges, then, must undergo disulfide interchange with the eventual formation of the correct set of disulfide bouds.

The likelihood of this rearrangement is made quite high by experiments carried out on the reactivation of certain derivat-

ives of ribonuclease formed by oxidation of the reduced chain under adverse conditions [8]. If reduced ribonuclease is allowed to reoxidize in air in the presence of urea or guanidine, as shown in Table II, orienting interactions are disturbed, and

TABLE II

EFFECT OF DENATURANTS ON OXIDATION OF REDUCED RNASE

Compound	Concentration (molar)	Activity % of native
Urea	0.25	100
Urea	1.0	92
Urea	2.0	40
Urea	4.0	25
Urea	6.0	10
Urea	8.0	< 1
Guanidinium chloride	1.0	88
Guanidinium chloride	2.0	25
Guanidinium chloride	3.0	4
Guanidinium chloride	4.0	< 1
KCl	1.0	101
KCl	3.0	94

the native configuration is not produced. Such derivatives, possessing only the enzymatic activity to be expected for a completely random mixture of half-cystine pairings (about 1%; 105 possible combinations of SH groups are theoretically possible), have physical properties remarkably similar to those of native ribonuclease, except for some not unexpected changes in optical rotatory properties and viscosity. The spectra of these materials are anomalous and show ultraviolet light absorption with a maximum at 276 rather than 277.5 mμ as is

[1] *Anfinsen* - pag. 6

observed for proteins containing tyrosine residues with anomalous resonance characteristics.

When these abnormal ribonucleases are dissolved in buffer solution to which relatively small amounts of sulphydryl reagents have been added, a slow rearrangement occurs with the regeneration of nearly full enzymatic activity, as shown in Table III [8]. It would appear that nature has evolved the

TABLE III

REARRANGEMENT OF INACTIVE RNASE DERIVATIVES IN MERCAPTOETHANOL

Material	Initial activity %	Activity after 24 hours incubation at pH 8.0 %	Activity after 24 hours incubation at pH 8.0 in presence of 100-fold molar excess mercaptoethanol
RNase oxidized at pH 5	7	7	83
RNase oxidized in urea	< 1	< 1	79
RNase oxidized in guanidinium chloride .	< 1	< 1	55

most probable geometric form, selecting a specific amino-acid sequence which predetermines the tertiary interactions required for activity. This statement is perhaps a truism. It is, nevertheless, reassuring to find that the structure occurring in nature is, indeed, the one that forms spontaneously. It may not be necessary to search for varieties of genetic information beyond that determining primary structure. The formation of three-dimensional structure may be an automatic process.

We have been able to obtain some evidence on the nature of the interactions that direct the refolding of reduced ribonuclease by carrying out reoxidations under conditions of low

[1] *Anfinsen* - pag. 7

pH, in the presence of various analogues of tyrosine, and following substitution of epsilon amino groups with polyamino-acid side chains of various lengths. Table IV shows a list of

TABLE IV

EFFECT OF TYROSINE ANALOGUES AND OTHER COMPOUNDS ON REGAIN OF ENZYMATIC ACTIVITY (*)

Compound	Concentration (molar)	Final activity of oxidized RNase- % of native
Phenol	0.1	< 1
p-OH-phenylacetic acid	0.1	< 1
Catechol	0.1	< 1
p-cresol	0.1	< 1
p-chlorophenol	0.05	< 1
l-naphthol	0.04	< 1
Tyrosine-glutamic acid copolymer 1:1	1×10^{-5}	< 1
Polyaspartic acid	5×10^{-4}	67
Benzoic acid	0.2	89
p-OH-benzoic acid	0.5	76
Phenylacetic acid	0.2	60
Phenylalanine	0.17	104
Cyclohexanol	0.2	88
Benzyl alcohol	0.2	39
Methanol	1.0	88
Dioxane	1.0	100
p-nitrophenol	0.1	45
Aniline	0.1	67
Bovine serum albumin	1%	74
Gelatine	1%	67

(*) [Reduced RNase] = 1.4×10^{-5} M.

[1] *Anfinsen* - pag. 8

substances that were effective as inhibitors of correct refolding [8]. The common characteristic of these inhibitors is the presence of a phenolic hydroxyl group and the absence of an ionizable group on the ring. Many of these substances bear a strong structural resemblance to tyrosine, which is, of course, precisely the reason they were chosen since native ribonuclease appears to depend, in part, on tyrosine interactions for the integrity of its three-dimensional form. As I mentioned a few moments ago, various inactive derivatives are characterized by an absorption spectrum more like that of free tyrosine than of proteins in which anomalously titrating tyrosine residues exist.

A particularly interesting modification of ribonuclease has been the derivative produced upon treatment of the native enzyme with the N-carboxyanhydrides (the Leuchs anhydrides) of alanine and of tyrosine [9]. In the case of alanine, it appears that 8 of the 11 amino groups of ribonuclease may be substituted and that even in the most highly substituted derivative, in which the 7 polyalanine side chains average 6-8 alanine residues per chain, activity against RNA is still 65% of native activity, and activity against synthetic uridine 2′, 3′-cyclic phosphate is approximately 20% of normal. The polyalaninated enzyme may be separated from polyalanine itself and from native ribonuclease by purification on a phosphorylated cellulose column as shown in Figure 3. The preparation, whose purification is shown in this figure, was fully active against RNA and showed a shifted, « active type » spectrum with a maximum at 277.5 mμ. Following complete reduction of its disulfide bridges, more than 75% of its activity was restored upon oxidation in air. These experiments indicate that the epsilon amino groups, themselves, are not prerequisites for normal interaction and suggest that the critical feature about these groups is the positive charges which they bear and which may be moved out from the surface of the protein by a distance corresponding to at least four or five alanine residues without scrambling of the sequential information content.

[1] *Anfinsen* - pag. 9

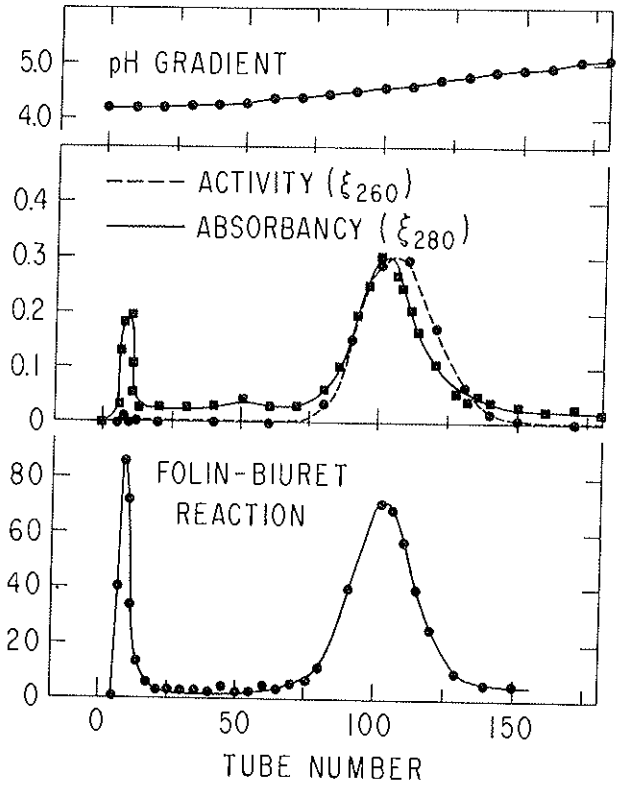


FIGURE 3 — *Isolation of polyalaninated ribonuclease.* After reaction of native ribonuclease with the N-carboxyanhydride of alanine (weight ratio 1:2) at pH 7 in 30% dioxane; 70% 0.05 M phosphate buffer, 5°C, the mixture was dialyzed against water and lyophilized. The dry product was subjected to chromatography on phosphorylated cellulose (7 mE capacity/gram of ion exchanger) using the buffer gradient system - 0.02 M citrate buffer, pH 4.2; 0.2 M citrate buffer pH 4.2; 0.2 M sodium citrate pH 7.4. With this system, polyalanine and native ribonuclease are well separated from the polyalaninated derivative.

The three amino groups that are unreactive to anhydride treatment are presumably buried within the fabric of the enzyme. Their identity is not yet definitely established although it seems nearly certain at the moment that one of the three

[1] *Anfinsen* - pag. 10

may be identified as residue 41. This residue is, incidentally, the lysine which is dinitrophenylated in the inactice mono-DNP ribonuclease prepared by Werner HIRS and his colleagues [10].

Dr. WHITE and Dr. Charles EPSTEIN have examined the reversible reduction and reoxidation of other proteins in an attempt to ascertain whether the primary sequence is sufficient for refolding as a general principle. Efforts with several proteins, including trypsin, chymotrypsin, trypsinogen, and chymotrypsinogen, have been generally unsuccessful because of the marked solubility of the reduced proteins. Dr. WHITE has, however, been able to regenerate a large fraction of the activity of reduced lysozyme, as has also been demonstrated recently by Dr. ISEMURA and his colleagues at Osaka University [11]. Dr. EPSTEIN has had some success with trypsin by attaching the protein to carboxymethyl cellulose through the azide derivative of the ion exchanger to prevent its aggregation. Insoluble columns may be prepared which very rapidly hydrolyze the substrate benzoylarginineamide. If such columns are exposed to mercaptoethanol in 8 M urea for several hours, they become completely inactive, due presumably to complete reduction of SS bridges. After washing the column with 0.1 M acetic acid, the SH groups may be oxidized with, as yet, a modest regeneration of activity by the passage of neutral buffer solutions, saturated with air, through the material. Experiments are in progress to determine the exact extent of reduction and the catalytic and physical properties of the re-formed enzyme after oxidation. It would seem, at any rate, that, with several catalytically active proteins, that portion of the tertiary structure necessary for function is regeneratable on the basis of the « instructions » inherent in the primary sequence alone.

What use can we make in the practical sense of this phenomena? One of the more obvious applications is the utilization of the automatic refolding process as part of the chain of events in the synthesis of protein molecules from smaller peptide fragments. Although our experiments along these lines

are very preliminary, Dr. Robert GOLDBERGER and I have recently begun studies on the reversible blocking of functional groups in ribonuclease, including the amino and carboxyl groups, and the SH groups after reduction. Encouraging degrees of reversibility have been achieved, and it has also been possible to fragment the blocked, reduced polypeptide chain at a limited number of points by controlled trypsin digestion. I shall not outline the details of the various steps we have examined but might suggest that this is a topic that would be of possible general interest in our subsequent discussions. There is, at least, some place for optimism in looking ahead to the total synthesis of an enzyme molecule and to the study of the mechanism of enzyme action through the systematic organic modification of restricted positions of the polypeptide chain.

[1] *Anfinsen* - pag. 12

REFERENCES

[1] J.A. SCHELLMAN, *Compt. rend. trav. lab. Carlsberg. Ser. chim.*, Vol. *30*, No. 21-27, 363 (1958).

[2] W.J. DREYER, M. CRUMPTON, G. STREISINGER and C. B. ANFINSEN (unpublished).
See also C. B. ANFINSEN, *Fed. Proceedings, 20,* 634 (1961).

[3] M. SELA, F.H. WHITE Jr., and C. B. ANFINSEN, *Science, 125,* 691 (1957).
F. H. WHITE Jr. and C. B. ANFINSEN, *Ann. N.Y. Acad. Sci., 81,* 515 (1959).
F. H. WHITE Jr., *J. Biol. Chem., 235,* 383 (1960).
E. HABER and C. B. ANFINSEN, *J. Biol. Chem., 236,* 422 (1961).

[4] C. B. ANFINSEN and E. HABER, *J. Biol. Chem., 236,* 1361 (1961).

[5] F. H. WHITE Jr., *J. Biol. Chem., 236,* 1353 (1961).

[6] J. BELLO, D. HARKER and E. DE JARNETTE, *J. Biol. Chem., 236,* 1358 (1961).

[7] C. B. ANFINSEN, E. HABER, M. SELA and F. H. WHITE Jr., *Proc. Nat. Acad. Sci., 47,* 1309 (1961).

[8] E. HABER and C. B. ANFINSEN, *J. Biol. Chem., 237,* 1839 (1962).

[9] M. SELA and C. B. ANFINSEN, Symposium on Poly-α Amino Acids, June, 1961. University of Wisconsin Press, in press.
C. B. ANFINSEN, M. SELA and J.P. COOKE, *J. Biol. Chem., 237,* 1825 (1962).

[10] C.H.W. HIRS and M. HALMAN, *Brookhaven Symposia in Biology,* No *13,* 131 (1960).

[11] T. ISEMURA, T. TAKAGI, Y. MAEDA and K. IMAI, *Biochem. Biophys. Res. Comm.*, *5,* 373 (1961).

DISCUSSION

THEORELL

May I add early examples to Dr. ANFINSEN's collection of unfolding and spontaneous recoiling reactions. In 1934 I dialyzed the « hold yellow enzyme » against 0.02 N HCl; the FMN went out, and when I neutralized the colourless solution from the interior of the cellophane bag, the protein precipitated in a « denatured » shape. When instead of neutralizing the solution, it was dialyzed against water, no precipitation occurred, and the protein could now be recombined with FMN. Sedimentation constant and activity were identical with those before the separation. This shows that, in agreement with Dr. ANFINSEN's views, spontaneous recoiling can occur even in a molecule as large as 103,000. By the way, the spontaneous rearrangement of ternary protein structures is one of the rare examples of reactions in biological materials not needing enzymatic catalysis. You will have a very hard time to find many examples of this kind.

RICH

There are two interpretations possible for the fact that you get 10% activity back when you have about 50% of the coiling. The one is that 10% of the molecules have renatured fully and they are therefore active; the other is that a larger percentage of the molecules have refolded partially and you measure some sort of partial acti-

vity. I wonder if it is experimentally possible to differentiate these alternatives.

The other question is: Can you isolate by hydrolysis some of these postulated incorrectly crosslinked -S-S- bridges and show that the wrong cysteines are joined?

ANFINSEN

It has been possible to show that under the conditions of oxidation employed intermolecular -S-S- bridges are predominately formed and that these rearrange through disulfide interchange to yield the native, fully active enzyme. When oxidations are permitted to occur at lower concentrations of reduced protein, such incorrect bridging can be avoided almost completely and theoretical yields of native enzyme are obtained in 30 minutes or less without a lag period.

FRAENKEL-CONRAT

I must confess that until recently I entertained a very small element of doubt in how far we could interpret these elegant experiments as proving that the conformation of proteins is completely and entirely dictated by the amino acid sequence. From our experiments on renaturation of the TMV protein we got the impression that this was possible from 8 M urea solution where the -SH group remains somewhat masked, but not from guanidine salt solutions, where it is free. We therefore wondered whether retention of some folding was after all a prerequisite for perfect refolding. The complete absence of interaction and retention of a minuscule focus of structure in the denatured protein would indeed be difficult to prove. Recent experiments, however, have shown us that renaturation of TMV protein is possible even from guanidine solution if the -SH group is protected by the addition of mercaptoethanol, and thus no experimental basis exists for the doubts expressed above.

[1] *Anfinsen* - pag. 16

SCHRAMM

I feel the renaturation is not so difficult if the protein is pure and if the concentration is so low that intramolecular reactions are preferred.

In the case of TMV the serological properties of the native protein and the denatured protein were studied. Are similar studies carried out with ribonuclease?

ANFINSEN

The native protein and the material produced upon reoxidation of reduced ribonuclease are serologically identical. It is difficult to study the « denatured » protein in this way since denaturation, including the sort resulting from incorrect -S-S- bridging, is easily reversible and the preparation of antibodies to denatured RNase might, therefore, be extremely difficult. It is true, however, that ribonuclease with « scrambled » disulfide bridges is not cross-reactive with anti-native ribonuclease.

SCHRAMM

I have another question. Is renaturation generally possible after substitution of some minor groups in the ribonuclease?

ANFINSEN

Renaturation appears to be generally possible after substitution if the substituted protein is active. Such observations suggest that the active center may be a major force in the direction of proper folding.

KATCHALSKI

Your interesting studies, Dr. ANFINSEN, indicate that the reformation of the three dimensional structure of ribonuclease from the corresponding randomly coiled, reduced polypeptide, is made possible by a continuous conformational fluctuation in the intermediate macromolecular structures formed. In this connection it is pertinent

[1] *Anfinsen* - pag. 17

to note that chemical and physical chemical data suggest that the shape of some proteins may change reversibly under the suitable conditions. In the case of bovine serum albumin, for example, we found (1) that the volume of the molecule increases at pH values above 8.0 or below 3.0. The « molecular swelling » is accompanied by the exposure of some of the -S-S- groups of the protein, which become susceptible to reduction. The swelling is reversible in the pH range of 2.0 to 10.5. At neutral pH the albumin seems to attain the most favorable and most compact conformation. It should be noted, however, that irreversible denaturation is known to occur in so many cases, that it is perhaps premature to draw any generalizations as to the factors determining the stability of the three dimensional structure of a protein molecule.

KATCHALSKI

Two additional comments. Polypeptidyl enzymes were useful in your studies. We have recently completed the preparation of polytyrosyltrypsin by initiating the polymerization of N-carhoxy-*L*-tyrosine anhydride with trypsin (2). The modified enzyme differs markedly in its solubility characteristics from trypsin. Its suhstrate specificity, however, resembles closely that of the parent enzyme. No loss in enzymatic activity occurred during the modification procedure adopted.

While studying the restoration of the enzymatic activity of ribonuclease, you recorded the corresponding changes in optical rotation. In this connection it is worth mentioning that from the data so far accumulated on the optical rotatory properties of poly-α-amino acids (3) it can be concluded that the changes in optical rotation

(1) E. KATCHALSKI, G.S. BENJAMIN and V. GROSS, *J. Am. Chem. Soc.*, 79, 4096, 1957.

(2) A.N. GLAZER, A. BAR-ELI and E. KATCHALSKI, *J. Biol. Chem.*, 237, 1832 (1962).

(3) E. KATCHALSKI and I. STEINBERG, *Annual Rev. Physical Chem.*, 12, 433, 1961.

reflect the corresponding conformational changes in relatively short helical regions. Alterations in the hydrodynamic parameters, on the other hand, measure the change in the overall shape of the molecule.

ANFINSEN

I might only add that the rapid increase in optical rotation of reduced ribonuclease during the early stages of its oxidation (during which incorrect and intermolecular -S-S- bridges are being formed) indicated that perhaps completely random covalent cross-linkages may be sufficient to permit the formation of helical structures.

LIPMANN

Do copper or other heavy metals catalyze reoxidation? Is mercaptoethanol oxidized and are cross links to mercaptoethanol formed that then rearrange, or is it a re-reduction-opening of -S-S- and reaggregation of ZSH---S-S- linkages?

ANFINSEN

I believe that the remarks I made in answer to Dr. RICH's question would cover this point.

[1] *Anfinsen* - pag. 19

ON THE MODE OF ACTION OF AZAGUANINE ON PROTEIN SYNTHESIS

H. CHANTRENNE

Laboratoire de Chimie Biologique, Université Libre de Bruxelles, Bruxelles (Belgique)

The growth of a variety of viruses, bacteria, and of animal and plant cells is inhibited by certain structural analogues of purines and pyrimidines (see MANDEL, 1959). For instance, when 8-azaguanine is added to a culture of *Bacillus cereus*, the rate of growth as estimated from turbidity measurements is decreased, and it was at first believed that azaguanine exerts a general inhibitory action on the complex process of growth. Upon closer examination, the effect of the guanine analogue was found to be much more specific. The most dramatic action is on protein synthesis: ten minutes after the addition of azaguanine, the rate of protein synthesis is reduced to about 10% of the normal figure, and formation of active enzymes is completely blocked. On the contrary, the rate of synthesis of RNA (in mg per ml of bacterial suspension) is not much decreased; it can actually be increased when the analogue is added to bacteria in the late growth phase. The synthesis of DNA is practically not touched when the concentration of azaguanine is low (e.g. 4 μg/ml) although sufficient to block enzyme synthesis completely. If the concentration of analogue is increased to 40 μg/ml, the rate of DNA synthesis eventually diminishes, but DNA synthesis continues for many hours. Fi-

nally, the synthesis of cell wall material, including the peptides it contains is completely insensitive to azaguanine. These results, that we reported a few years ago (CHANTRENNE, 1958; CHANTRENNE and DEVREUX, 1958, 1960), have now been confirmed by others (RICHMOND, 1959; MANDEL and ALTMAN, 1960; ROODYN and MANDEL, 1960; OTAKA, 1960).

Azaguanine thus specifically inhibits the formation of proteins, without doing much harm to the production of the other macromolecules. In this respect, the effects of azaguanine on *B. cereus* very much resemble those of chloramphenicol upon *E. coli*.

CHANGES IN RNA CAUSED BY AZAGUANINE

Azaguanine does not prevent RNA formation, but the RNA made in its presence is abnormal in several respects: part of the guanine is replaced by azaguanine, the amount of non sedimentable RNA is greatly increased and the RNA which contains azaguanine is more labile in acid medium than normal RNA. Thus azaguanine changes the composition, the distribution and the properties of RNA at the same time it inhibits the synthesis of proteins. (CHANTRENNE and DEVREUX, 1958; OTAKA, 1960).

A first question to be asked is whether the effect of the analogue on protein synthesis is linked with the incorporation of azaganine into nucleic acids. It is clear that the inhibitory action of azaguanine depends on the incorporation of the analogue into nucleotidic material, for azaguanine resistant mutants derived from a sensitive strain differ from the original strain by the loss of the enzyme which reacts guanine and azaguanine with phosphoribosyl-pyrophosphate forming guanylic acid or its analogue. (BROCKMAN et al., 1959, 1961).

Among the nucleotidic compounds into which azaguanine is incorporated, special attention must be paid to azaguano-

sinetriphosphate (Aza GTP), since this substance plays a part in the passage of amino-acids from transfer RNA to the nascent polypeptide on the ribosome. Recent research by Roy et al. (1961) showed that in rat liver preparations, Aza GTP does not interfere with this process: it can even reactivate preparations lacking GTP.

The almost complete inhibition of amino-acid incorporation into protein material therefore cannot be explained by a failure of the system at the level of GTP. What matters is apparently the incorporation of azaguanine into polynucleotides. For instance, *Tetrahymena geleii* requires uracil for continued growth; when deprived of uracil, it stops making RNA but can continue to increase its mass. (Heinrich et al., 1952). Under such conditions, azaguanine exerts no effect on protein synthesis. On the contrary, if the medium contains uracil, RNA synthesis is ohserved, and protein synthesis is inhibited.

Nature of the RNA into which azaguanine is incorporated

In *Bacillus cereus*, it takes only ten minutes for the inhibition of protein synthesis by azaguanine to be completely expressed. The RNA made in 10 minutes amounts to a 10% increase of the total RNA present at the time of addition of the analogue. If the inhibition of protein synthesis is due to the formation of RNA in which azaguanine is substituted for part of the guanine, it must be concluded that the appearance of a relatively small amount of abnormal RNA is enough to reduce all protein synthesis to a very low rate, and that further production of abnormal RNA does not make things worse.

The turnover of ribosomal RNA is very low in bacteria (Davern and Meselson, 1960); it is clear therefore that most of the ribosomal RNA present at the time of addition of the analogue persists when protein synthesis stops.

Research by P. CAMERMAN in our laboratory indeed showed that azaguanine is incorporated only into RNA molecules that are assembled after the analogue was added; it is not incorporated by any kind of exchange process into old RNA molecules. Unless it is assumed that the active life of ribosomes is short, and that most of the ribosomes of a bacterium are useless, it must be concluded that the inhibition of protein synthesis by azaguanine is not due primarily to an action on ribosomal RNA. Other RNA fractions must be involved.

Fractionation of bacterial extracts by centrifugation 10 minutes after adding azaguanine — i.e. when the inhibition of protein synthesis is completely expressed — indeed shows that non sedimentable RNA contains much more azaguanine per mg than does sedimentable RNA.

This observation and the facts reported above led to the conclusion that the general inhibition of protein synthesis caused by azaguanine is due to its incorporation into a soluble RNA fraction which is fairly rapidly renewed (CHANTRENNE, (1958).

In order to gain information on the nature of the RNA which is replaced during the first ten minutes of action of azaguanine, we used the method of VOLKIN and ASTRACHAN (1957); that is the study of the distribution of ^{32}P among the 4 nucleoside - 2'(3')-phosphate obtained by alkaline hydrolysis. This distribution closely reflects the composition of the RNA made during P incorporation.

The results of such experiments differed considerably according to the stage of growth of the bacteria. With bacteria in the logarithmic phase of growth, the composition of the RNA made during the first 10 minutes after azaguanine addition had the same composition as the RNA made in the control, except for the replacement of about 30% of the guanine by azaguanine (Fig. 1 *a* and *b*), in agreement with previous observations by MANDEL and MARKHAM, 1958.

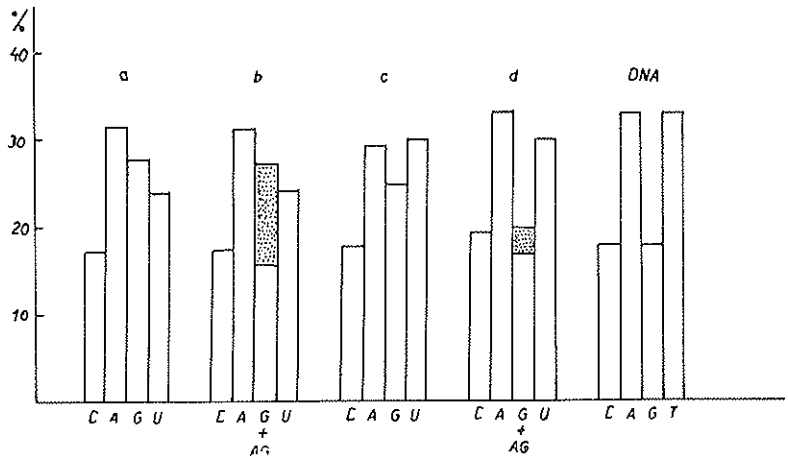


FIG. 1 — Base composition of the RNA made in *B. cereus* as determined by the method of VOLKIN and ASTRACHAN. Phosphate incorporated for 10 minutes.

a and *b*: exponentially growing bacteria,
a: control
b: with azaguanine;
c and *d*: bacteria at the end of their growth phase,
c: control
d: with azaguanine.

The results are different when the bacteria are taken during the phase of diminishing growth rate. For instance, when bacteria are taken at a time when their mean generation time is 180 min. (i.e. three times that corresponding to the maximal rate of logarithmic growth) and transferred into fresh medium, the composition of RNA which is made approaches that of a complementary RNA with equal amounts of A and U on one hand, and of G and C on the other. This phenomenon is similar to that reported by HAYASHI and SPIEGELMAN (1961).

If azaguanine is added to such bacteria, it again replaces part of the guanine in this particular kind of RNA. In several experiments, we observed that azaguanine can even shift the

composition of RNA from the normal one to that of complementary RNA, as if the analogue would favour the formation of complementary RNA and depress that of the normal RNA assortment (Fig. 1 *c* and *d*). This result, however, depends on the state of the bacteria in a way that we were not able to solve yet.

In any case, our results indicate that azaguanine can be incorporated into RNAs which have a base composition similar to that of the DNA of the bacterium. Rapid renewal and a composition similar to that of DNA are two properties of the informational RNA or messenger RNA of bacteria, the significance of which is discussed in the reports by Dr. JACOB and Dr. SPIEGELMAN in the present book. Another feature of informational RNA is that it can form hybrids with the corresponding DNA. (HALL and SPIEGELMANN, 1961). Recently, OTAKA et al. (1961) reported that azaguanine is incorporated into an RNA fraction which is found in association with DNA in bacterial extracts. There is every reason to conclude that azaguanine is rapidly incorporated into informational RNA. The same conclusion was also reached by GROS and NAONO (1961) for another analogue, 5-fluorouracil.

The inhibition of protein synthesis by azaguanine may be due to structural modifications on the informational RNA which cannot accomplish its normal function.

It should at present be possible to test such an RNA directly in view of the recent results reported by NIRENBERG and MATTHAEI (1961). Such experiments would indicate directly whether the transcription into polypeptides is prevented by the presence of azaguanine in messenger RNA.

EFFECTS OF AZAGUANINE ON THE SYNTHESIS OF INDIVIDUAL PROTEINS

The experiments considered above dealt with the inhibition of protein synthesis in general. In *B. cereus,* the inhibition

is drastic: it amounts to 90% under our experimental conditions for the bulk of the protein material and it is complete for the enzymes studied (catalase, penicillinase, respiratory system).

With other types of cells, a less general inhibition by azaguanine can be observed: for instance, in *S. aureus* and in plant tissues, the analogue can inhibit the synthesis of certain enzymes selectively — without affecting the formation of other enzymes (CREASER, 1956; HEYES, 1959). A comparable phenomenon is observed with *B. cereus* during the restoration of protein synthesis after inhibition by azaguanine (CHANTRENNE and DEVREUX, 1960). Guanosine can indeed completely release the inhibition but the restoration of protein synthesis becomes more and more sluggish as the period of azaguanine inbibition increases. A striking fact is that the synthesis of penicillinase is not restored at the same time as that of the average protein material (Fig. 2); a lag of more than two hours can be observed between the recovery of protein synthesis and the recovery of penicillinase production. This result is obtained with strains of *B. cereus* which are inducible as well as with penicillinase constitutive mutants, and it is immaterial whether the inducer (penicillin G) is added before azaguanine, together with it, or together with guanosine.

Catalase formation, on the contrary, is restored rather more rapidly than the average protein synthesis (Fig. 3).

Beside the general inhibition that it exerts on the synthesis of all proteins, azaguanine thus causes other lesions which do not affect equally the synthesis of all enzymes. These secondary lesions develop rather slowly and they are not rapidly cured.

When an enzyme activity fails to increase, this does not necesseraly mean that the synthesis of the enzyme protein is completely arrested. It may be that certain enzyme activities fail to increase because the protein forming system makes mistakes, and produces slightly modified proteins which are

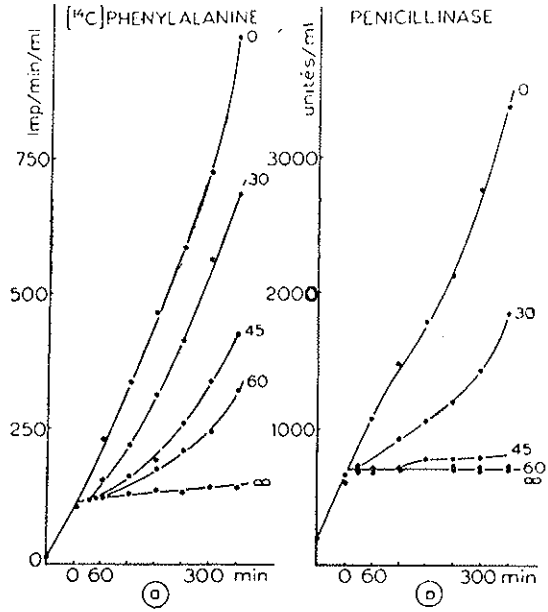


FIG. 2 — Restoration of the synthesis of protein and penicillinase caused by guanosine added after various times of action (0, 30, 45, 60 minutes) of azaguanine. Note that when added 45 or 60 minutes after azaguanine, guanosine causes a good recovery of phenylalanine incorporation into protein material but little recovery of penicillinase production.

devoid of enzyme activity. In *E. coli*, there is good evidence that the yield in β-galactosidase activity is reduced by 5 F-uracil or 2-thiouracil simply because the enzyme made in the presence of the analogue is slightly abnormal (NAONO and GROS, 1960; HAMERS and HAMERS CASTERMAN, 1961). The catalase made during restoration after azaguanine inhibition may also be slightly abnormal. It does not differ from the normal enzyme in its affinity for two typical inhibitors, hydroxylamine and azide, but it is more thermolabile than normal catalase (H. CHANTRENNE, 1961).

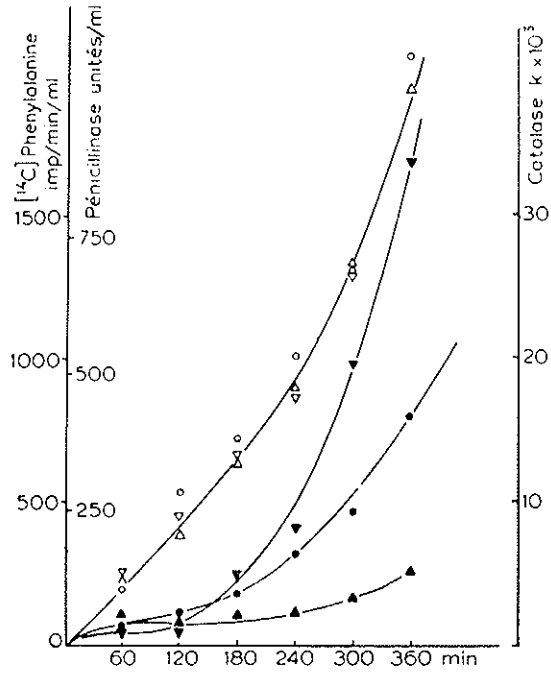


FIG. 3 — Restoration of the synthesis of penicillinase, catalase and total protein material caused by the addition of guanosine 45 minutes after azaguanine.

White dots: (control) azaguanine and guanosine added simultaneously at time O.

Black dots: azaguanine added at time O, guanosine 45 minutes later.

○ ● : ^{14}C incorporated into protein.

△ ▲ : penicillinase activity.

▽ ▼ : catalase activity.

Our search for abnormal proteins during restoration concentrated on penicillinase. Serious technical difficulties were encountered in our attempts at fractionating the proteins secreted by *B. cereus*. — among which penicillinase is found — due mainly to the production of proteolytic enzymes by the bacterium. In the present state of our research, conducted

in collaboration with J.F. PECHERE and J. ZANEN, we have not detected the formation of any protein which would cross react with penicillinase anti serum. It would seem at present that during the period of recovery the synthesis of penicillinase remains inhibited for a rather long time, and that either no abnormal penicillinase-like protein is made, or if such a protein is made, that it is different enough from penicillinase to have lost both the enzyme activity and the serological properties of normal penicillinase.

It is difficult under these conditions to establish whether or not an abnormal protein controlled by the structural gene of penicillinase is made during the phase of the restoration period when no recognizable penicillinase appears. The use of a general fractionation method, like electrophoresis, and of an inducible mutant will possibility help solving this matter.

If no protein depending on the structural gene of penicillinase is produced during the restoration of protein synthesis while other proteins are being formed, a disparity in the control of synthesis of individual proteins would be indicated; further research will be necessary to establish whether the observations reported can be accounted for by the current theory of the control of protein synthesis, or whether the disparity among protein synthesizing systems is real and raises new problems.

The formation of well defined abnormal proteins under the action of F-uracil and possibly of azaguanine, leaves other questions unanswered. It is easy to understand that replacement of a normal base by an analogue in messenger RNA or in any part of the reading mechanism might result in the production of abnormal proteins because the information becomes difficult to read, or because the reading machine is damaged and likely to make mistakes.

It is clear however that *each* substitution of an analogue for its normal counterpart does not result in a mistake in the amino acid sequence, otherwise one would observe the formation of a variety of nonsense polypeptides, not the production

of rather homogeneous and slightly abnormal proteins. One wonders then under which circumstances a mistake is made in the assembly of the polypeptide.

It is not known whether replacement of guanine by azaguanine is completely random; the probability of substitution might depend on the nature of the neighbouring nucleotides. Another puzzling observation which was reported for F-uracil is that the analogue apparently causes the substitution of certain residues of a given amino-acid species by another, but not the substitution of all the residues of that species. A degenerate code would afford an explanation, for it may be that substitution in one only of the various coding sequences corresponding to an amino acid would result in a mistake. A redundant code could explain that in spite of a high percentage of replacement, few mistakes are made.

Although the nature of the coding principle and the code itself are about to be solved by direct experiments (NIRENBERG and MATTHAEI, 1961), it is our opinion that analysis of the effects of analogues may afford valuable information about the nature and properties of the reading mechanism.

REFERENCES

BROCKMAN R.W., DEBAVADI C.S., STUTTS P. and HUTCHINSON D.S., *J. Biol. Chem., 236,* 1471 (1961).

BROCKMAN R.W., SPARKS C., SIMPSON M.S. and SKIPPER H.E., *Biochem. Pharmacology, 2,* 77 (1959).

CHANTRENNE H., *In Biological Structure and Function.* (T.W. Goodwin and O. Lindberg, Eds.). Academic Press London, vol. I, 281 (1961).

CHANTRENNE H., *Rec. Trav. chim. Pays-Bas, 77,* 586 (1958).

CHANTRENNE H. and DEVREUX S., *Exp. Cell. Res.* Suppl. 6, 152 (1958).

— *Nature, Lond., 181,* 1737 (1958).

— *Biochim, Biophys. Acta, 39,* 486 (1960).

— *Biochim. Biophys. Acta, 41,* 239 (1960).

CREASER E.H., *Biochem. J., 64,* 539 (1956).

DAVERN C.I. and MESELSON M., *J. Mol. Biol., 2,* 153 (1960).

GROS F. and NAONO S., *In Protein Biosynthesis.* (R.J. Harris, Ed.). Academic Press London, p. 195 (1961).

HALL B.D. and SPIEGELMAN S., *Proc. Natl. Acad. Sci.,* U.S., *47,* 137 (1961).

HAMERS R. and HAMERS-CASTERMAN C., *J. Mol. Biol., 3,* 166 (1961).

HAYASHI M. and SPIEGELMAN S., *Proc. Natl. Acad. Sci. U.S., 47,* 1564 (1961).

HEINRICH M.R., DEWEY V.C., PARKS R.E. and KIDDER G.W., *J. Biol. Chem., 197,* 199 (1952).

HEYES J.K., *Symp. Soc. Exptl. Biol., 13,* 365 (1959).

MANDEL H.G., *Pharmacol. Rev., 11,* 743 (1959).

MANDEL H.G. and ALTMAN R.L., *J. Biol. Chem., 235,* 2029 (1960).

MANDEL H.G. and MARKHAM R., *Biochem. J., 69,* 297 (1958).

NAONO S. and GROS F., *Compt. Rend. Acad. Sc., 250,* 3527 (1960).

NIRENBERG M.W. and MATTHAEI J.H., *Proc. Natl. Acad. Sci. U.S., 47,* 1588 (1961).

OTAKA E., *Exp. Cell. Res., 21,* 229 (1960).

OTAKA E., OSAWA S. and OOTA Y., *J. Mol. Biol., 3,* 693 (1961).

RICHMOND M.H., *Biochim. Biophys. Acta, 34,* 325 (1959).

ROODYN D.B. and MANDEL H.G., *J. Biol. Chem., 235,* 2036 (1960).

ROY J.K., KVAM D.C., DALL J.L. and PARKS R.E., *J. Biol. Chem., 236,* 1158 (1961).

VOLKIN E. and ASTRACHAN L., in *The Chemical Basis of Heredity.* The John's Press, Baltimore, p. 686 (1957).

DISCUSSION

THEORELL

Have you made, or intended to make, attempts to crystallize and analyze the catalase formed with and without azaguanine? Determinations of catalase activity in impure extracts may easily be misleading, but there probably should be no particular difficulties in isolating catalase in a crystalline shape, and to make analyses of the total amino acid composition. I think very interesting differences could be revealed which would not necessarily interfere with the activity, and thus escape observation by activity measurements.

CHANTRENNE

No we did not try to purify catalase; our main effort is on the isolation of penicillinase and on the search for modified penicillinases.

SPIEGELMAN

It would be awfully interesting to arrance a condition in which one gets the analogue into the transfer RNA rather than in informational (I think that really can be arranged) and then to see whether that has an effect on the production of abnormal protein. It may be that we have to use « in vitro » systems to accomplish this. It is relatively easy to separate complementary RNA from transfer RNA on the HERSHEY column, i.e. kieselguhr coated with methylalbumin. This column does recognize base composition as was indicated in the very early work of CHARGAFF.

CHANTRENNE

I might mention that OSAWA and OGATA did some experiments in the direction you suggest. They showed that transfer RNA in which 25% of the guanine is replaced by azaguanine still accepts amino acids as does normal transfer RNA, with no apparent change in specificity.

SPIEGELMAN

This is not decisive. Because when you isolate s-RNA by the usual procedure, it is heavily contaminated with complementary RNA.

CHANTRENNE

This may very well be. OGATA's experiments at least indicate that one of the functions of transfer RNA is not changed by azaguanine, namely the capacity to accept amino acids.

THE PROBLEM OF THE PRIMARY STRUCTURE OF THE DEOXYRIBONUCLEIC ACIDS

ERWIN CHARGAFF

Cell Chemistry Laboratory, Department of Biochemistry, Columbia University, New York, N.Y. (U.S.A.)

The first requirement for an understanding of the primary structure of a high polymer is the knowledge of the number and nature of its constituents and of the manner in which these are linked. In the case of those nucleic acids that have so far been studied, this may be said to have been more or less accomplished in the course of the last 15 years [1]. It must, however, be admitted that this statement can apply only to the relative, but not to the absolute, number of monomeric constituents. In almost no case has a clear agreement been reached concerning molecular weight and acceptable criteria of integrity and homogeneity of a nucleic acid. This is not surprising; quite apart from the great difficulty of handling a family of substances of very high molecular weight and closely similar physical properties, it must be assumed that the more intimately related a compound is to the central life processes, the more precarious it will be to decide in what form and association it acts in the living cell.

Of the isolated and chemically investigated deoxyribonucleates (DNA) the statement can be made that they are built of chains of four or five principal deoxyribonucleotides that are linked by 5′:3′ phosphodiester bridges. The existence

of minor constituents cannot be excluded; nor are generalizations permissible when dealing with nucleic acids that have not yet been investigated in detail.

If the important rôle currently assigned to DNA as genetic determinants is accepted, it stands to reason that the "information" coded into DNA must find its chemical expression in the sequence of purines and pyrimidines on the sugar phosphate backbone. There may, of course, exist also other aspects of expression, e.g., physical or, better, geometrical; but the shape in which a nucleic acid functions in the cell — for instance, in the form of a nucleoprotein — is completely obscure. The remaining decisive problem in the chemical definition of the primary structure of a deoxyribonucleic acid is without doubt the development of methods for the determination of nucleotide sequence. The conceptual definition of what is meant by a nucleic acid molecule also remains to be achieved.

Only four or five letters, consequently, make up the "alphabet" of a DNA; but the words thus spelled are truly gigantic: the estimates vary from 10,000 to 200,000 nucleotides per chain; and even more enormous is the number of possible isomers of any given total composition. The baffled cryptographer is faced with a most intractable task. He may hope, while laboriously developing the trends of nucleotide arrangement characterizing different DNA specimens, that other, more convenient and less ambiguous, scanning procedures will be discovered. He may fear that even the successful perusal of an entire DNA text will prove less informative than is expected at present.

Whereas in most other respects the nucleic acids are infinitely more refractory to sequence determination than are the proteins, they stand out through one fortunate feature, namely, the differences in the stability of the glycosidic linkages of their component nucleotides, so that the bonds attaching the purines to the sugar are much more easily cleaved by acid than are those holding the pyrimidines. We made early use

of this method of differentiation in our studies on the production [2] and degradation [3] of what we called apurinic acid.

It is on this early work that the more refined procedures of the analysis of the distribution of solitary pyrimidines [4] and of pyrimidine clusters or bunched pyrimidines [5] are based. For purposes of definition it may be said that a pyrimidine nucleotide flanked, within the polynucleotide chain, only by purine nucleotides will be referred to as « solitary ». Pyrimidine nucleotides occurring in stretches of two or more, flanked on both sides by purine nucleotides, are designated as « bunched » or, for long tracts, as « clusters ». Solitary pyrimidine nucleotides can be determined with relative ease, since they will be released, through a series of elimination reactions, in the form of the corresponding nucleoside 3',5'-diphosphates by mild acid treatment [4]. Similar procedures lead to the isolation of the various pyrimidine oligonucleotide clusters, consisting of sections of different length, each carrying phosphomonoester groups on both sides at positions 3' and 5', respectively. The pertinent literature can be found in the papers quoted; the main conclusions from our studies have been reviewed recently [6].

Such gradual and controlled degradations can teach us a great deal about certain features of the nucleotide arrangement in DNA. The general conclusions of our work on tendencies of pyrimidine distribution may be summarized as follows. 1) The method of pyrimidine distribution analysis permits the distinction of deoxyribonucleic acid specimens that are indistinguishable by the analysis of their purine and pyrimidine composition. Specimens having identical total nucleotide contents, though derived from different organisms, differ widely in most sequence characteristics that can be developed by the distribution procedure. - 2) A large proportion — in many cases 60% or more — of the pyrimidines occurs as tracts of three or more pyrimidines in a row. - 3) The nucleotide arran-

gement in most deoxyribonucleic acids examined so far is not random. This may be taken as an indication that a deoxyribonucleic acid chain could represent a meaningful text. It must, however, be understood that it is only through the intelligence of the reader, or hearer, that an array of letters, or sounds, can acquire significance.

The convenient, if somewhat misleading, slogan « biological information » is another example of the motorized anthropomorphism that disfigures much of our biological reasoning at present. This concept postulates the existence of a chain of information in the cell which is thought to be represented by a series of signals, as it were, emanating from a specific deoxyribonucleic acid and transmitted by a specific ribonucleic acid messenger of matching composition to the unspecific ribonucleoproteins of the ribosomes, thus instructing the latter to proceed to the manufacture of specific proteins. Our ignorance of the first causes thus is concealed by a terminological scaffold which, while lacking the beauty of an allegory, retains its vagueness.

Sequential specificity of high polymers of cellular origin is, as we have learned in the recent past, one of the principal marks of the specificity of the cell. Now that sequence studies on proteins are being carried out so widely, we are becoming familiar with the statement that, in a particular peptide sequence, a given amino acid A is « replaced » by another one, B. This is an easy statement to make; but what does replacement mean in biological terms? We have attempted to explore this question in the field of the nucleic acids in which the replacement of one constituent occupying a definite position in the chain by another must be at least as significant, if the rôle currently assigned to the nucleic acids is correct. In this field, statements have been even more sweeping than in the case of the proteins; and this thanks to the pairing principles which since their first announcement [7] have become well established. Just as, in the total molar composition of a

deoxyribonucleic acid, adenine equals thymine and guanine equals cytosine, we know that guanine also equals the sum of cytosine and 5-methylcytosine in polymers containg the latter pyrimidine. Similarly, in ribonucleic acids some of the uracil can be « replaced » by 5-fluorouracil [8]. We should, therefore, be inclined to say that one 6-amino pyrimidine can replace another; and the same would hold for 6-keto compounds. An example for such a conclusion will be found in Table I, taken from a previous paper [4].

TABLE I — *Composition of deoxyribonucleic acids of rye germ and wheat germ*

	Rye germ	Wheat germ
	Moles/100 g-atoms P	
Adenine	27.7	28.1
Guanine	22.6	21.8
Cytosine	16.4	16.8
Methylcytosine	5.9	5.9
Thymine	27.4	27.4
	Molar ratios	
Adenine + thymine to guanine + cytosine + methylcytosine	1.23	1.25
Purines to pyrimidines	1.01	1.0
Adenine to thymine	1.01	1.03
Guanine to cytosine + methylcytosine	1.01	0.96
Cytosine to methylcytosine	2.78	2.85

One could have surmised — and many have done so — that this « replacement », for instance, of cytosine by 5-methylcytosine, means that any cytosine molecule within the chain has an equal chance of being substituted by a methyl-

cytosine molecule. But is this really the case? Even our first findings served to throw doubt on such a generalization. It is possible to fractionate deoxyribonucleic acids through the fractional dissociation of the nucleoproteins or of artificially prepared histone nucleates or similar salts with other basic proteins [9]. When such preparations from wheat germ or rye germ were examined, they were found to differ with respect to the relative proportions of cytosine that had been « replaced » by 5-methylcytosine (Table II, taken from previous publications [4, 10]). It is quite clear that in the process of the assembly of the polynucleotide chain a discrimination between the two 6-amino pyrimidines must have taken place.

Even if it were assumed that the methylation of part of the cytosine occurs after the formation of the polymer — and there have been no indications of such a reaction so far — this would only serve to localize such a predetermined and specific discrimination in the enzyme performing this hypothetical and rather improbable step.

Even more surprising were the results of distribution studies. The relative abundance of solitary pyrimidine nucleotides in the total deoxyribonucleic acid of rye germ and in seven fractions isolated therefrom is shown in Table III (taken from a previous paper [4]). It will be seen that methylcytosine has a much higher tendency to occur as a solitary unit interjacent between purines than has the other 6-amino pyrimidine cytosine. It resembles, in this respect, thymine. It is clear that one cannot speak of a random replacement of cytosine by its methyl derivative; much rather could one conclude that it is the two 5-methyl pyrimidines that go together.

In conclusion, a few words should be said about the recent attempts of my laboratory to develop a method for the separation of all pyrimidine clusters that occur in a given deoxyribonucleic acid interspersed between two or more purine nucleotides. In this manner, a much more complete account can be given of the array of pyrimidine oligonucleotides. We

TABLE II — *Uneven distribution of 5-methylcytosine in DNA fractions*

Source of DNA	Methylcytosine as mole % of cytosine + methylcytosine in fractions								
	I	II	III	IV	V	VI	VII	VIII	IX
Wheat germ. Prep. 1 .	28.6	28.4	26.7	23.2	24.3	23.5	24.1	23.0	
Wheat germ. Prep. 2 .	27.6	26.9	26.5	25.8	24.0	23.4	23.3	22.0	20.9
Wheat germ. Prep. 3 .	30.7	26.5	25.7	25.8	25.7	26.6			
Rye germ	28.6	28.3	26.9	25.9	26.6	24.7	22.9		

TABLE III — *Pyrimidine distribution analysis of the total deoxyribonucleic acid of rye germ and of its fractions*

Mole % of total constituent in DNA	Total DNA	Fr. I	Fr. II	Fr. III	Fr. IV	Fr. V	Fr. VI	Fr. VII
Solitary thymine	18	14	14	17	17	14	19	15
Solitary methylcytosine . . .	18	12	18	17	20	16	22	18
Solitary cytosine	11	7	11	10	12	9	13	11

have investigated several nucleic acid specimens by this procedure; as an example our preliminary results on calf thymus deoxyribonucleic acid are reproduced in Table IV taken from a recent paper [5]. It may be of interest to list the various pyrimidine oligonucleotides which were isolated, in collaboration with Dr. J. H. SPENCER, from the nucleic acids of calf thymus and rye germ. From the compilation in Table V it will be seen that even pentathymidylic acid occurs in both specimens, whereas the largest exclusive oligonucleotide of cytosine is the trinucleotide.

TABLE IV — *Distribution of pyrimidine nucleotide clusters in calf thymus DNA* (*)

Fraction No.	Number of pyrimidine nucleotides in fragment n	Total phosphorus in fraction μ moles	Total pyrimidine phosphorus in fraction μ moles	Proportion of DNA phosphorus %
I	1	7.51	3.75	10.1
II	2	5.53	3.69	9.9
III	3	4.34	3.25	8.7
IV	4	3.57	2.86	7.7
V	5	2.49	2.08	5.6
VI	6	1.79	1.54	4.1
VII	7	1.40	1.22	3.3
VIII	8	0.86	0.77	2.1
IX	**	1.98	1.81	4.9

(*) The hydrolysate subjected to fractionation contained 37.31 μ moles of DNA phosphorus of which 5.79 μ moles had been released as inorganic P, recovered in Fr. 1. The total recovery amounted to 98.6% of the DNA phosphorus.

(**) This fraction comprises the nucleotide clusters Py_np_{n+1} in which n equals 9 or more. The average chain length of this fraction corresponded to $n \backsim 11$. The values referring to this fraction in the last two columns are based on this estimate.

TABLE V — *Pyrimidine nucleoside 3', 5'-phosphates isolated from DNA* (*)

	RYE GERM	CALF THYMUS
Mononucleotides	T	T
	C	C
	M	M
Dinucleotides	T_2	T_2
	CT	CT
	C_2	C_2
	M_2	
	MT	
Trinucleotides	T_3	T_3
		T_2C
	TC_2	TC_2
	C_3	C_3
	C_2M	
	T_2M	
Tetranucleotides	T_4	T_4
	T_3C	T_3C
	T_2C_2	T_2C_2
		TC_3
	T_3M	
Pentanucleotides	T_5	T_5
	T_4C	T_4C
	T_3C_2	T_3C_2
		T_2C_3
	T_4M	

(*) Positional isomers are not distinguished in this list. T designates thymine; C, cytosine; M, 5-methylcytosine.

[3] *Chargaff* - pag. 10

REFERENCES

[1] CHARGAFF E. and DAVIDSON J.N. (editors), *The Nucleic Acids*, Academic Press, New York, vols. I, II (1955); III (1960).

[2] TAMM C., HODES M.E. and CHARGAFF E., *J. Biol. Chem.*, *195*, 49 (1952).

[3] TAMM C., SHAPIRO H.S. and CHARGAFF E., *J. Biol. Chem.*, *203*, 673 (1953).

[4] SHAPIRO H.S. and CHARGAFF E., *Biochim. Biophys. Acta*, *23*, 451 (1957); *26*, 596, 608 (1957); *39*, 62, 68 (1960).

[5] SPENCER J.H. and CHARGAFF E., *Biochim. Biophys. Acta*, *51*, 209 (1961), and in press.

[6] CHARGAFF E., in *Biological Structure and Function* (IUB/IUBS Symposium, Stockholm, 1960). Academic Press, New York, p. 67.

[7] — *Experientia*, *6*, 201 (1950).

[8] HOROWITZ J. and CHARGAFF E., *Nature*, *184*, 1213 (1959).

[9] CHARGAFF E., CRAMPTON C.F. and LIPSHITZ R., *Nature*, *172*, 289 (1953).

[10] LIPSHITZ R. and CHAGAFF E., *Biochim. Biophys. Acta*, *19*, 256 (1956).

DISCUSSION

FRAENKEL-CONRAT

What is the status of hydrazinolysis as a method to obtain apyrimidinic acid?

CHARGAFF

Unfortunately, the methods described heretofore for the preparation of apyrimidinic acid are not very satisfactory. In our hands, at any rate, the specimens so obtained often contained still traces of pyrimidines and they are, moreover, depolymerized to a large extent. It is, therefore, questionable whether they can be utilized for sequence studies. We are, however, working quite actively on this problem since it would be of obvious interest to be able to compare our previously determined pyrimidine sequences with the purine sequences from the same DNA. Within certain limits, it should thus be possible to arrive, for the first time, at a chemical decision on the existence of base-pairing in all regions of a polynucleotide sequence.

SCHRAMM

It is very surprising that after incorporation of the bromouracil such great changes in the total sequence occur. A number of solitary pyrimidines are changed. Are there any possibilities to explain this fact?

I remember that FREESE has investigated the reason for this mutation. He thinks that bromouracil can replace to a certain extent also cytosine, not only thymine. It may be to a very small extent. But it is enough to explain the mutations.

CHARGAFF

Further investigations on the effect of the incorporation of bromouracil into DNA have failed to confirm our original conclusions as to a severe distortion of the pyrimidine arrangement in *E. coli* DNA.

We are now preparing for publication a report on quite extensive studies with two different mutants that have led us to correct our previous observations. We have also found bromouracil to be an excellent tool for the study of the mode of replication of bacterial DNA. A report of this work is also being prepared for publication. It is, however, possible that the brominated analogue disturbs what might be called the « micro-sequence », namely, in certain predetermined spots. There must after all be a reason for the occurrence of preferred locations of mutability arising from the incorporation of bromouracil.

JACOB

What is the net increase in DNA content when you replace thymine by bromouracil?

CHARGAFF

I cannot provide quantitative data to answer your question.

In *E.* coli 15 t⁻arg⁻ it is possible to get several replication cycles with bromouracil in the entire absence of thymine.

CRITICAL OPALESCENCE AND THE RANGE OF MOLECULAR INTERACTION

P. DEBYE

Department of Chemistry Cornell University - Ithaca, N.Y. (U.S.A.)

When molecules have dimensions large enough to be comparable to the wavelength of visible light, observation of the dependence on the angle of scattering of the intensity of the light scattered by dilute solutions can be used to determine those dimensions experimentally [1]. The scattered intensity of polarised light observed in a plane perpendicular to the electrical vibration decreases with increasing scattering angle. This dissymmetry effect has been used to show that in the case of coiling molecules the average distance between ends of the molecular chain is very nearly proportional to the square root of the molecular weight. In this way the analogy between the flexible chain and the mathematical problem of the random flight is established. Polystyrene of a molecular weight of one million dissolved in benzene has an average end-to-end distance of about 1100 Å. The practical limit for measurements of these dimensions lies at about 300 Å where the angular dissymmetry becomes too small to be measured with some accuracy.

In 1950 B. ZIMM [2] observed angular dissymmetry of the light scattered by mixtures of two liquids consisting of small molecules, provided the experiment was carried out in the vicinity of the critical mixing point. It is well known that

the theory of SCATCHARD-HILDEBRAND predicts for such mixtures that the curve representing the osmotic pressure of say liquid B in A as a function of the concentration does exhibit a critical point that is an inflexion point with a horizontal tangent (parallel to the concentration axis) for a definite critical temperature. The abscissa of this point defines the critical concentration. Above the critical temperature the two liquids are miscible in all proportions; below this temperature separation in two phases occurs.

In the vicinity of this critical point the osmotic work necessary to separate the mixture in two components becomes very small, zero in the critical point itself. If now a light scattering experiment is performed, say at the critical concentration but above the critical temperature, it is observed that the scattered light becomes very strong. The nearer the temperature comes to the critical temperatnre the higher the intensity. This is readily understood since the osmotic work concerned in separating the two liquids becomes very small, zero in the critical point itself. The thermal motion therefore can produce strong fluctuations of the concentration and consequently of the refractive index. This has been known for a long time. What was added to this as a result of the observation of ZIMM was that the increasing intensity is coupled with increasing angular dissymmetry, although the molecules in question are small (carbon tetrachloride and perfluorated methylcyclohexane).

It is to be expected that something else (not the molecular dimensions) has become comparable to the wavelength of visible light in the vicinity of the critical point. What this is becomes clear when we try to visualize the geometrical structure of the instantaneous distribution of the refractive index in the mixture. If we plot this index as an ordinate along a line AB of length r in the liquid we observe a curve representing fluctuations around an average valne of the index which are characterized by two featnres. The one feature is their amplitude, the other

feature is the distance over which a given amplitude persists before it changes say from a positive value to a negative value. This second feature can be measured exactly by introducing the so called correlation or autocorrelation function [3]. Take for a given distance r of the points A and B the product of the fluctuations Δ_A and Δ_B and establish the average of this product. For zero distance this product becomes equal to the average value of Δ^2. Dividing $Av <\Delta_A \Delta_B>$ by $Av <\Delta^2>$ we obtain a number which will depend on the distance $AB = r$. This represents the correlation function $C(r)$ which according to this definition will start with 1 for $r = 0$. We can predict that $C(r)$ will tend to 0 for large values of r, since under these circumstances what will happeu in B will be totally independent of what happens in A.

The theory, based on MAXWELL's equations, now predicts the following connection between the correlation function and the intensity distribution. If $C(r)$ is known as a function of r the scattered intensity I can be calculated as a function of the scattering angle Θ, which appropriately can be measured by $s = 2 \sin \frac{\Theta}{2}$. Also the opposite is true. If the intensity I is known as a function of S, the correlation function C can be calculated as a function of the distauce r. The qualitative effect is such that a slow decrease of C with increasing r acts the same, as far as scattering is concerned, as a big molecule in a diluted solution.

Mathematically it is appropriate to define the square of a persistence-length L by the second moment of the correlation function. If this is done it can be said that the angular dissymmetry of the scattered intensity will be the more pronounced the bigger the persistence-length is. From the rate of decrease of $I(s)$ plotted as a function of s^2 for small angles an exactly defined value of L^2 can be derived. In the case of ZIMM it turned out that at a temperature distance of 0.02° from the critical temperature the persistence-length L is 1500 Å.

[4] *Debye* - pag. 3

The next question [4] to be considered was how the persistance-length L depends on the temperature-distance from the critical temperature T_c. ZIMM's experiments show that for experiments performed at the critical concentration the relation is

$$L^2 = \frac{l^2}{\frac{T}{T_c} - 1}$$

in which l is a constant length much smaller than L which in ZIMM's case turns out to be $l = 14.7$ Å.

This relation generates at once the desire of a theoretical derivation. This can be performed if we start with an interpretation of EINSTEIN's theory of light scattering introduced by BRILLOUIN [5] and apply it to a background provided by a molecular theory of the equation of state. The most simple case, which can be considered is that of the effect of density-fluctuations in a gas of one component in the vicinity of the critical point, commonly considered in this case as characterised by an inflection-point with horizontal tangent of the critical isotherm. Considering the light scattered under a definite angle Θ with the direction of the primary light, BRILLOUIN's calculations show that if a FOURIER analysis is made of the instantaneous density-fluctuations there is only one component of this analysis which determines the intensity for the angle. This component represents a wave with a front making the angles $\Theta/2$ with the primary and the secondary direction. This front acts so to say as a mirror on which the primary light is reflected. Of components with such a front there are a very large number contained in the analysis each characterised by a wavelength Λ which goes from $\Lambda = \infty$ to a wavelength of the order of the mutual distance of neighboring molecules. Only one component however is important, namely that component which has a wavelength Λ which is related to the wavelength λ of the light by BRAGG's relation, making

[4] *Debye* - pag. 4

$$\frac{r}{\Lambda} = \frac{s}{\lambda}$$

The intensity of the light scattered in the direction Θ now is proportional to the square of the amplitude of this FOURIER-component.

Measuring the intensity of the scattered light in going from $\Theta = 0$ to $\Theta = 180°$ therefore amounts to measuring the square of the amplitudes of partial FOURIER-waves; their sum representing the total thermal density-fluctuation. The wavelength Λ of these partial waves starts with ∞ for $\Theta = 0$ and ends with $\Lambda = \frac{\lambda}{2}$ for $\Theta = 180°$. Obviously this merely covers a small part of the total thermal spectrum of the fluctuations (which extends to much smaller wavelengths) at least as long as only visible light is considered.

In his original paper EINSTEIN [6] implicitly calculated the energy connected with one of the thermal waves from the compressibility of the liquid. He concluded in this way that this energy was independent of the wavelength. Applying the equipartition principle he then came to the conclusion that the amplitudes of all the different waves to be considered are the same in thermal equilibrium and independent of their wavelength. This leads unavoidably to the conclusion that no angular dissymmetry of scattering should exist.

There is no question about the fact that under the prevailing circumstances the equipartition principle will hold. What has to be investigated is whether the energy of a wave of a given amplitude is indeed wholly independent of the wavelength. The vicinity of the critical point is characterised by the fact that (in our example of the density fluctuations) the compressibility becomes very large, infinite in the critical point itself. The calculated energy for a fluctuation of a given amplitude a will become the smaller the nearer we are to the critical point. In this region it will therefore be necessary to investigate whether there exist no other reasons for an energy additional

to the compressional energy. In any molecular theory of the equation of state as for instance the classical theory of VAN DER WAALS the potential energy of a molecule with respect to its surroundings has to be introduced. Commonly this energy is calculated from the molecular interaction energy between two molecules by supposing that the variation of the density around the central molecule can be neglected over distances which are of the order of the range of the molecular forces. Strictly this is of course not quite right [7]. The density around the central molecule can be developed in powers of the coordinates around this point. If this is done and only the first terms of this development are retained it turns out that not only the average density but also the local variation of the density is involved in calculating the potential energy of the central molecule. If this is followed through it ultimately turns out that for a fluctuation represented by a wave of wavelength Λ the energy is still proportional to the square of its amplitude a. The factor however now contains an additional term which is proportional to the square of the ratio: range of molecular forces l to the wavelength Λ. Since as we have seen this wavelength obeys BRAGG's law and varies with the angle, equipartition now demands that the square of the amplitude of the special thermal disturbance responsible for the scattered light at the angle Θ decrease with increasing angle. This effect will be the more important the higher the value of the energy-correction due to the local density variations is as compared to the amount of compressional work. This shows why a special effect has to be expected in the vicinity of the critical point where this compressional work becomes very small.

The same kind of reasoning can be applied to concentration fluctuations in mixtures. The only difference is that instead of one range-constant l we will have another range-constant which is composed of three constants measuring the interactions AA, BB, and AB in a case of a mixture of two components A and B. In the most simple representation of the theory,

[7] *Debye* - pag. 6

along the lines of VAN DER WAALS for a simple gas the value of l^2 is the second moment of the curve representing the interaction-energy as a function of the distance of the two interacting molecules [8].

Ultimately the prediction is that if a plot is made representing the reciprocal of the scattered intensity as a function of s^2 a series of slanting straight lines (all parallel to each other) is obtained. The slope of these lines is a measure of l. In ZIMM's case the value of l so obtained was $l = 14.7$ Å, in the case of a mixture of methanol and cyclohexane the experimental value of l was 12.7 Å. The remarkable fact is, that by measuring the angular intensity-distribution of the scattered light in the vicinity of the critical point it is possible to measure molecular interaction-ranges of the order of 10 Å with visible light of a wavelength of the order of 3000 Å.

After these interconnections were established it seemed interesting to apply the method to coiling polymers [9]. In a solution the average space occupied by such a molecule is mostly filled with the solvent. So for a polystyrene molecule of one million molecular weight, which in a benzene solution occupies approximately a spherical space with a diameter equal to a measured average end-to-end-distance of 1100 Å the average density in this space due to the polymer is only $2.4 \cdot 10^{-3}$ g/cc. The centers of gravity of two polymer-molecules which are approaching each other will get much nearer than the end-to-end-distance. It seemed interesting to determine experimentally with the new method what the range of interaction for such molecules would be and to compare such ranges with the size for different molecular weight. This has been done for polystyrene in cyclohexane ranging in molecular weight from 69,000 to 2,820,000. Similar measurements were also made for a few samples in ethyl-cyclohexane. The somewhat surprising result was that the interaction range was a good deal smaller than the expected and measured coil-dimensions. The values of l for the different molecular weights men-

tioned above varied from 24 Å for the lowest molecular weight to 66 Å for the highest. In order to see whether this surprising result was real a second method for the determination of l was conceived. Light passing through a solution is lost solely as a result of scattering. This scattering due to thermal fluctuations can be rather large in the vicinity of the critical point. Apparent absorption coefficients α of the order of 1 cm^{-1} have been observed. Theoretically such apparent absorption-coefficients can be calculated by integrating the scattered intensity over all angles of a sphere. The formula which is obtained in this way will show α as dependent on the wavelength λ of the light, since the angular distribution depends on λ. The coefficient will theoretically become infinite for the critical temperature T_c. If the results of such a calculation are plotted for a given wavelength λ in the form $1/\alpha$ as a function of the temperature T a curve is obtained which is a sloping straight line for larger values of $T - T_c$. Approaching the critical temperature this straight line curves down and reaches the T-axis at $T = T_c$. To different wavelengths λ belong different curves. All end at $1/\alpha = 0$ for $T = T_c$ but their straight parts have different slopes. If these straight parts are extrapolated they hit the T-axis in different points depending on the wavelength. In this way the impression is created as if the critical temperature T_c depends on the color of the light by which it is observed. An experimental investigation [10] with a spectro-photometer of a polystyrene solution in cyclohexane with a polymer of a molecular weight of 147,000 yielded a series of straight lines for $1/\alpha$ for wavelengths ranging from 3,600 Å to 4,800 Å. Extrapolation of these straight lines leads indeed to different values of T_c. Now the theory predicts that if those values of T_c^* determined in this way are plotted themselves as a function of $1/\lambda^2$ a straight sloping line should result and the slope of this straight line should be a measure for the values of l. Such a line was indeed obtained and the value of l derived from its slope was $l = 29.1$ Å. For the same polymer l had

[4] *Debye* - pag. 8

been measured previously by observing the angular dissymmetry of the scattered radiation and from this the value $l = 28.4$ Å had been derived. No doubt seems to be left of the fact that indeed the polymer-molecules are peuetrating in each other to a large extent before appreciable interaction occurs.

It will have been noticed in the previous part that the theory so far put forward as an explanation of the observed dissymmetries is only an approximation. In the case of density-fluctuations for instance the density around a central molecule was represented by a development in powers of the coordinates around that molecule and only the first terms of this development were retained. This approximate calculation is of course not necessary and a final result can be reached without making use of such a development. In this way one is led to the observation that it should be possible to derive for instance the interaction energy between two molecules as a function of their mutual distance, from another curve which represents the scattered intensity as a function of the scattering angle observed under appropriate circumstances. So far the observed curves of this kind have been very simple. The reciprocal intensity plotted as a function of s^2 was represented by a straight line. What we are looking for are deviations of this straight line. Presumably they should occur if we can make observations of the angular intensity curve near enough to the critical temperature. Unfortunately this has not been possible for polymer solutions up to now. The available polymers which have a nominal molecular weight are as a matter of fact mixtures of a range of molecular weights. It was supposed that the molecular weight distribution curve of the polymers which were available to us would be rather sharp, since their quotient: weight average to number average molecular weight was near 1. So for instance for a sample, mentioned before of a nominal molecular weight of 147,000 this quotient was given as 1.04. Thinking this over it appears very soon however that the nearness of this quotient to unity is a very

bad measure for the sharpness of the distribution curve. In the special case of the quotient 1.04 it is easily shown that half the width of the distribution curve at an ordinate which is 50% of the ordinate at its maximum is equal to 30% of 147,000, the molecular weight at the maximum. The existence of this relatively broad distribution prevents an approach nearer than 0.7° to the apparent critical temperature of the mixture.

Samples with sharper distribution curves have kindly been made available to us in the meantime by Dr. BREITENBACH in Vienna and Dr. WENGER of the Mellon Institute in Pittsburgh. However they are still not sharp enough for our purpose.

Under these circumstances it seems that another approach to the case of polymer solutions is more promising, at least for the time being. What is measured primarily is the correlation-curve connecting fluctuations in different points of the solution. To perform this measurement an interference-effect is used. Generally in such an effect the scattered intensity distribution depends on a variable of the form $L\,s/\lambda$ in which L is a length determined by the special circumstances. Therefore it is possible to observe the more details of this curve the larger the range of values s/λ is, which can be covered. The angular variable s can only range from 0 to 2. This indicates that our goal might be reached by substituting X-rays for visible light. From the already known magnitude of the persistence-lengths L in practical cases it is clear that these X-ray measurements will have to be made at small angles. Such measurements have been started. At this moment not enough work has been done to say more than that the results confirm in a qualitative way our expectations that for larger values of s/λ deviations from the straight-line representation of $^1/I$ occur indeed.

[4] *Debye* - pag. 10

REFERENCES

[1] P. DEBYE, *J. Appl. Phys., 15,* 338 (1944); *J. Phys. Colloid Chem., 51,* 18 (1947); « Les Grosses Molécules en Solution », *Hommage National à Paul Langevin et Jean Perrin,* Collège de France (1948); « Estructura de alto Polimeros Estudiada por Metodos Opticos », *Annales de la Real Sociedad Española de Fisica y Quimica* B, *46,* 343 (1950); « The Structure of Polymers », *Science* (A.A.A.S. The Present State of Physics), 169 (1954).

[2] B. H. ZIMM, *J. Phys. Colloid Chem., 54,* 1300 (1950). See also R. FÜRTH and C. L. WILLIAMS, *Proc. Roy. Soc.* (London) A *224,* 104 (1954).

[3] The importance of correlation was stressed not long after the classica' publication of A. EINSTEIN, *Ann. Physik, 33,* 1275 (1910) by F. ZERNIKE, *Thesis,* Amsterdam; *Arch. Néerl.* III A, *4,* 74 (1917), *Proc. Acad. Sci. Amsterdam, 17,* 793 (1916); L. S. ORNSTEIN and F. ZERNIKE, *Physik. Z., 19,* 134 (1918), *27,* 761 (1926).

[4] P. DEBYE in *Non-Crystalline Solids* (Conferences, Alfred, N.Y., Sept. 3-5, 1958), John Wiley and Sons, New York, p. 1; *J. Chem. Phys., 31,* 680 (1950).

[5] L. BRILLOUIN, *Ann. de Phys., 17,* 88 (1922).

[6] A. EINSTEIN, loc. cit.

[7] The effect of a fluctuation gradient on the scattering already stressed by Yves Roccard, *J. de Phys., 4,* 165 (1933).

[8] P. DEBYE, *J. Chem. Phys., 31,* 680 (1959).

[9] P. DEBYE, H. COLL and D. WOERMANN, *J. Chem. Phys., 32,* 939 (1960), *33,* 1746 (1960). See also P. DEBYE, *Makromol. Chem., 35,* A, 1 (1960); *Zeitschrift für Physik, 156,* 256 (1959); P. DEBYE, B. CHU and D. WOERMANN, *J. Chem. Phys., 36,* 1803 (1962).

[10] P. DEBYE, B. CHU and D. WOERMANN, *J. Chem. Phys., 36,* 1803 (1962).

DISCUSSION

ANFINSEN

What about the possibility that the globular proteins might be unwound?

DEBYE

The method should work independent of the special molecular structure.

ANFINSEN

How high must the molecular weight be?

DEBYE

There is no lower limit. Ordinary small molecules are also appropriate.

PUTZEYS

The molecular field of the solvent molecules, such as benzene, manifests itself by a nearly complete depolarization of the scattered light. Is there any effect on the depolarization in this case?

DEBYE

This has not been investigated. However I do not expect anything special.

SCHRAMM

What kind of molecular forces act at these distances?

DEBYE

All forces which are responsible for instance for liquefaction and which are sometimes called VAN DER WAALS forces.

KATCHALSKI

Does the theory developed apply also to native rigid macromolecules which possess the shape of elongated ellipsoids of revolution?

DEBYE

The theory applies to any kind of molecules and any kind of interaction. For rigid molecules I do not expect the obtainable information to go much beyond information about size or form. It seems more interesting at least to me and at this moment to do experiments about coiling molecules especially since it was found for polystyrene molecules that the range of interaction is a good deal smaller than the size, indicating mutual penetration of the molecules.

IRON TRANSPORT RATE IN THE NEOPLASTIC ORGANISM

G. DE HEVESY

Institutionen för kemi och biokemi, Kungl. Universitetet i Stockholm, Stockholm (Sverige)

A quarter of a century ago HEILMEYER and PLOETNER have drawn attention to the diagnostic importance of the determination of the plasma iron concentration. Following up the above suggestion an immense amount of useful information was since obtained. Such informations can be enlarged by determining not alone the concentration of the plasma iron but also the rate at which the iron atoms present in the plasma are cleared.

Under physiological conditions about 90 p.c. of the iron atoms of the plasma find their way into the bone marrow and are replaced by iron atoms coming from the iron depots. The plasma iron clearance rate supplies thus some information on the rate of haemopoiesis going on in the bone marrow.

To determine the clearance rate of plasma iron atoms we have to label it with the radioactive Fe^{59} and follow the exodus of the radioactive atoms from the circulation. The labelling is carried out by incubating the plasma sample with minute amounts (1 μg or less per ml) of with Fe^{59} labelled iron citrate for about 20 minutes, as first done by HUFF and associates. The iron atoms added get combined with β_1-globulin molecules present in the plasma. It is bound to β_1-globulin that all

plasma iron is present under physiological conditions. At intervals of 20 minutes blood samples are taken, their plasma isolated and the activity of a known plasma volume compared with that of a known aliquote of the incubated sample. From the rate of decrease of the radioactivity of the sample and the colorimetrically determined iron content of the plasma the milligrams of iron can be calculated which left the circulation in the course of an hour, which amounts under physiological conditions to about 1.4 mg.

Is the haemopoiesis accelerated, as it is the case in the height where anoxia prevails, the flow rate of plasma iron is spectacularly increased as demonstrated by John LAWRENCE and his colleagues who compared the plasma iron transport rate in students on sea level and in the Andes mountains.

In cancer and a number of other diseases the life cycle of erythrocytes is often shortened, haemoglobin is leaving the circulation at an accelerated rate and correspondingly anoxia prevails. This induces, presumably through an increased haemopoetin supply, an increased rate of haemopoiesis. Such an increase necessitates a more copious iron supply to the bone marrow, a more rapid transport rate of plasma iron.

In advanced cancer the increased rate of haemopoiesis does mostly not suffice to fully compensate the erythrocyte (haemoglobin) loss due to the shortening of the life cycle of the erythrocytes and correspondingly anaemia sets in. In non-advanced cancer almost full compensation may take place. This was the case for patients of the Radiumhemmet of Stockholm suffering from invasive carcinoma of the uterine cervix, the iron clearance of which was studied in our laboratory prior to and 1 year after radiation therapy.

As seen from Fig. 1 stating the mean value of 27 cases taken from a paper of my collaborator LOCKNER the only haematological data which significantly differ prior to and after therapy are plasma iron transport rate expressed as half-disappearance

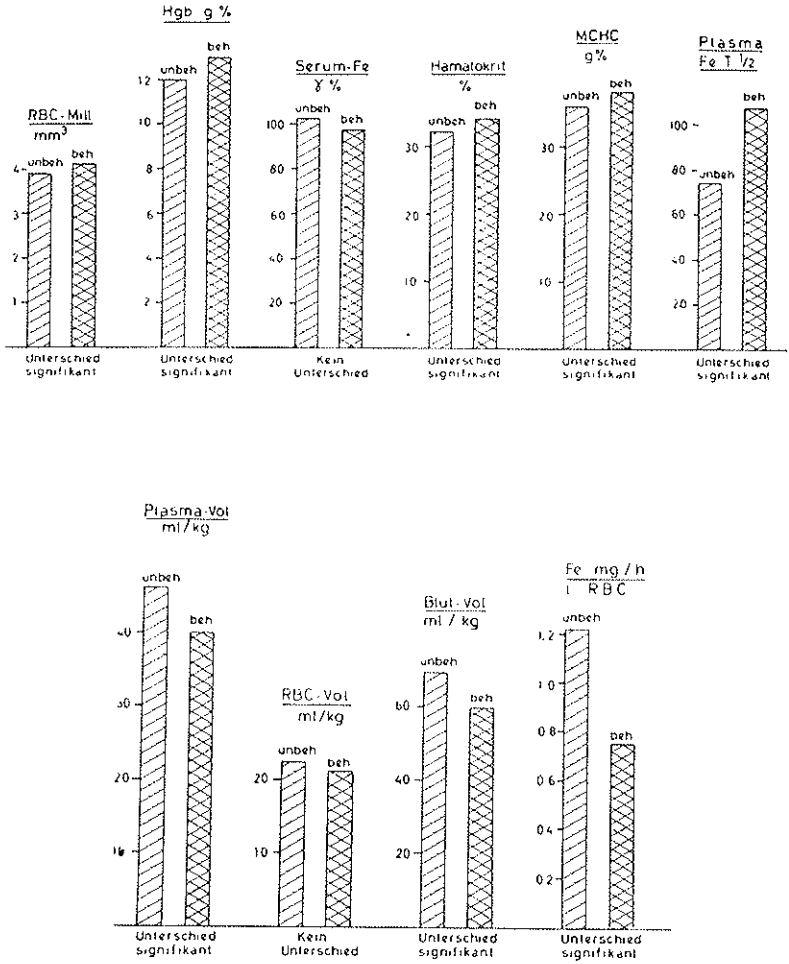


FIG. 1 — Haematological data of 27 patients suffering from invasive carcinoma of the uterine cervix before and one year after radiation therapy.

time (T½) of the labelled plasma iron or as mg. iron leaving per hour the plasma, furthermore the plasma volume.

After successful treatment the plasma iron transport rate was reduced in the average with about 40%. The mean value of the mg iron/hr/l red blood corpuscles was before treatment 1.22 and after successful treatment 0.75 ($p<0.001$). No decrease of the transport rate took place when the treatment was not successful. In some cases the transport rate was found even considerably increased in patients with radio-incurable or recurrent carcinoma. In three patients the iron turnover increased several weeks before the recurrent carcinoma could be diagnosed clinically.

It is not always easy to decide if a plasma iron clearance rate is pathological or not in view of the fact that the clearance rate in healthy subjects is showing quite a large variation. This difficulty did not arise in our investigations as we based our conclusions on results obtained by investigating the same person prior to and after therapy. From the fact that the clearance rate was reduced after successful therapy we could conclude that it was enhanced prior to treatment. The patients investigated have not shown a significant difference in their haemoglobin content or body weight prior to and after therapy and thus from the difference in the iron transport rate before and after treatment some conclusion could be drawn as to the effectiveness of the therapy.

From the rate of disappearance of Cr^{51} labelled red corpuscles we can conclude only on the decrease of the number of viable erythrocytes present in the circulation with time, while even the formation of short living erythrocytes or reticulocytes, furthermore ineffective erythropoiesis going on in the bone marrow is indicated by the iron clearance rate figures. The measurement of plasma iron trasport rate leads therefore to a further going information than does the measurement of the life cycle of erythrocytes.

[5] *De Hevesy* - pag. 4

DISCUSSION

SPIEGELMAN

Your explanation for the observation of increased clearance observed in carcinomatous patients would suggest that other pathological condition might also lead to the same result. Have any been found?

DE HEVESY

In almost every case in which the life cycle of the erythrocytes is shortened, we find an increased clearance rate of plasma iron. The shortening of the life cycle of the erythrocytes leads to anoxia which in turn produces an accelerated haemopoiesis, an accelerated haemoglobin formation, an accelerated plasma iron transport rate.

In advanced cancer and several other diseases the accelerated haemoglobin formation does mostly not suffice to compensate the loss of the haemoglobin content due to the shortened life cycle of the erythrocytes and correspondingly anaemia sets in.

In non-advanced cancer, the haemoglobin loss due to life cycle shortening of the erythrocytes may be fully compensated by an increased rate of haemopoiesis which necessitates a correspondingly increased rate of plasma iron transport rate. This was the case in the investigation described.

After a successful therapy the transport rate normalised.

We do not draw conclusions from a change in the clearance rate alone but from the difference in the clearance rate observed prior to and after therapy.

LIPMANN

There is no direct relation between this comment and Professor DE HEVESY's lecture, but nevertheless I would like to mention what Dr. BRAUN of the Rockefeller Institute told me about iron transport through plant membranes. He finds that inositol deficiency abolishes iron transport, which is promptly resumed on inositol addition. This reminds of HODKIN's demonstration that in animal systems the phosphate turnover of inositol-phosphate is connected with iron transport, such as transport of potassium (K^+) and sodium (Na^+) ions. Is anything similar known for iron transport in mammalian cells?

DE HEVESY

We have extensive information how different substances influence extracellular iron transport and also the incorporation of iron into haemin inside the cell, but not much is known on the mechanism of cellular intrusion of iron.

STRUCTURAL AND FUNCTIONAL RELATION BETWEEN THE NUCLEIC ACID AND PROTEIN OF VIRUSES

H. FRAENKEL-CONRAT
Virus Laboratory, University of California, Berkeley, Calif. (U.S.A.)

All viruses consist primarily, if not exclusively, of protein and nucleic acid. The nucleic acid is either of the deoxyribose type (DNA) (animal and bacterial viruses), or of the ribose type (RNA) (animal, bacterial and plant viruses). The DNA containing viruses vary in shape and size, as well as in DNA and protein content over a wide range. There are some viruses resembling organisms (e.g., vaccinia virus), and others showing the sharp edges and symmetries usually associated only with crystalline arrays of molecules (e.g., the tipula iridescens virus). The RNA viruses also show great variability, but have one striking feature in common: their absolute content in RNA is approximately constant, corresponding to a molecular weight of 1.5 to 2.2×10^6. Since their protein content varies greatly, the RNA/protein ratio of these viruses ranges from about 0.01 to 0.40.

A seemingly common feature of all viruses is the architectural relation of protein and nucleic acid. It appears that the protein always covers the nucleic acid. Thus, all surface properties of viruses are determined by the protein coat, and

the nucleic acids are rendered accessible only by disruption of the particle.

The virus which has been the subject matter of the greatest number of studies, and which has marked most of the milestones of virology, is the tobacco mosaic virus, TMV. It was the tobacco mosaic disease which led to the discovery of a new type of infective agent, a « filtrable virus » over 60 years ago. It was this particular virus which about 26 years ago led W.M. STANLEY to the conviction that viruses are physicochemically definable and uniform particles which can be crystallized. It was again TMV which led to the advances of the past seven years in our understanding of the role of the two viral components, which represents the subject matter of this lecture.

TMV is a rod-shaped virus, composed of about 2,200 protein molecules aggregated in a macro-helical array around one RNA molecule. The protein molecules (or subunits of the virus particle) are seemingly identical peptide chains, held together in helical array and thus forming the virus particle only through their high affinity to bind one another in specific manner. This binding is due only to secondary forces such as hydrogen bonds, hydrophobic interactions and ionic bonds. The native protein retains this associative affinity even when separated, and will reversibly aggregate and disaggregate depending on conditions of pH, ionic strength, and temperature. If the RNA is also present, a coaggregation of the two components gives a considerably more stable structure than that formed by protein alone, and complete virus is thus preferentially formed under suitable conditions. It was the discovery of this « reconstitution reaction » in our laboratory in 1955, which has triggered the present era of virus research.

While it is now recognized that the RNA of viruses carries their genetic activity, it is obvious from the above discussion, and other data to be discussed later, that the proteins play an important and specific role in the structure and function of

viruses. This has led chemists to elucidate the structure of virus proteins, and in the case of the protein of TMV, the complete amino acid sequence analysis has been achieved at the Virus Laboratory of the University of California, Berkeley, as well as at the Max Planck Institute für Virologie at Tübingen. Much of the structure was reported by the Berkeley group in 1958, while the first positive contributions from TÜBINGEN appeared in 1960. An almost complete structure was proposed by each of these groups during that year [1, 2]. Most of the minor differences have largely been corrected, but slight uncertainty remains in the part of the molecule which is near the single -SH group of this protein. One feature of particular interest in the virus protein is the absence of a terminal amino group, this being acetylated. All viruses which have been investigated show this same feature, but acetylation was recently observed to block the terminal amino groups also in several non-viral proteins (MSH, ovalbumin, fibrinogen). Another seemingly important feature of the TMV-protein is the proline residue near the C-terminus, which limits enzymatic attack from that end to the release of only the terminal threonine. While the native TMV protein is water soluble, and forms rods by aggregation, as described above, treatment by various denaturants leads to an intractable material, insoluble over the range of pH 3 to 10. Recently it was shown, however, that this denaturation could under suitable conditions be reversed, and seemingly native protein could be regenerated. This important observation of ANDERER's [4] in conjunction with similar ones made with other proteins such as pancreatic ribonuclease strongly suggest that the chain conformation of proteins, which is responsible for their 3-dimensional structure, is a consequence of their amino acid sequence. Thus, sequential information becomes of greater importance then anticipated. Since the biological role and function of TMV-protein is largely if not entirely structural, elucidation of the detailed mechanism by which sequence leads to chain conformation becomes of

primary importance. Such studies have only recently been initiated, and nothing definite can as yet be concluded. However, an inspection of the amino acid sequence reveals features which may be of significance in this connection. Thus, over half of the leucines and isoleucines are bunched in pairs or greater numbers, and the same is true of the amide groups, and of the serine residues. Also 12 of the 16 serine residues occur within 15 residues from either end of the 158-residue chains, and 11 of the 15 COO^- groups are centrally located. Future studies may well reveal the functional significance of these observations.

We will now turn our attention to the RNA of the virus. As mentioned earlier, the RNA has a molecular weight of about 2×10^6. All evidence favors the concept that this is a single unbranched polynucleotide chain, consisting of about 6500 nucleotides. The length of such a polynucleotide can be calculated to be about 30,000 A°, and measurements from electronmicrographs are in accord with this calculation. It can also be calculated that a helix of RNA of the diameter and pitch derived from the architecture of TMV would be 3000 A° long, and this is the actual length of the TMV rod. Thus, it is now clear that the length of the stable rod is the result of the lengths of the RNA molecule. In the virus rod there is no interaction of bases along the polynucleotide chain, but in salt solution the purine and pyrimidine bases tend to interact in RNA as they do in DNA, with the resultant decrease in U.V. absorption (hypochromicity), and increase in optical rotation. Since this interaction is a secondary and freely reversible phenomenon, it is not regarded as critical to the biological function of the RNA.

The biological function of viral RNA is the same as that of the DNA of other viruses, and all organisms, namely that of transmittal of genetic information. It was shown in 1956 in Berkeley and Tübingen that TMV-RNA alone can initiate the infection, and it was shown in further work from our

laboratory that the nature of the disease produced by the RNA was indistinguishable from that produced by the complete virus. TMV, like most viruses, represents a generic term, covering a great number of variants or strains, and it was shown by Knight that these strains usually show detectable differences in their amino acid composition. This definite phenotypic expression was utilized in establishing the genetic activity of the RNA. For it was possible to show that RNA infection leads to progeny virus identical even in its protein structure to that of the original virus from which the RNA was prepared. And this was true, even if the RNA was reconstituted with a heterologous strain protein: the progeny had protein homologous to the RNA, and different from that used for reconstitution and inoculation. Thus the RNA was unequivocally shown to carry all the genetic information necessary to describe a virus.

The fact that the RNA is very much less infective in the free state than in combination with the protein (about 0.2%) is attributed to its great sensitivity to enzymes. Nucleases and phosphatases are ubiquitous, and it is believed that the enzymes of the test leaf inactivate much of the RNA before it can initiate an infection. This belief is supported by many observations showing that the relative RNA infectivity varies markedly under different test conditions. The recent observation that the application of bentonite, an RNase inhibiting clay, to the leaf surface greatly potentiates the infectivity of RNA supports this belief.

From these and other experimental observations the concept has arisen and become widely accepted, that infective RNA consists of a population of polynucleotide chain molecules, unform in nucleotide sequence and number. Each preparation of RNA contains also some (25-50%) shorter polynucleotides, presumably largely fragments of the former, and inactive. This concept calls for elucidation of the specific nucleotide sequence of a given RNA, a task of enormous dif-

ficulty which will probably require several decades for its successful completion. The difficulty is primarily due to the length of the intact polymer (6500 residues), and to its lack of characteristic features (only four components, adenylic, guanylic, cytidylic, and uridylic acids). Methods are now available to determine the end groups of the chain, and the difficulties of applying these methods to so large a polymer were found surmountable, largely through the use of virus carrying sufficient radioactive label to detect one in 6500 residues. We have employed predominantly P^{32} and C^{14} labelled virus, containing several million cpm/mg virus. These studies, performed by Mr. T. SUGIYAMA and Mrs. B. SINGER in our laboratory have led to the conclusion that TMV-RNA carries unphosphorylated adenosine on both ends of the chain [4]. Current studies by Dr. P. WHITFELD give as yet preliminary information about neighboring residues.

An alternate approach to the chemical nature of the infective molecule is through the preparation of derivatives. The nucleotides in the RNA can be subjected to a variety of chemical or physical agents. Thus methyl groups and higher alkyl groups can be introduced, as well as bromine, hydroxyls, etc. The most interesting of these reactions is that with nitrous acid which replaces amino by hydroxyl groups [5].

By means of extrapolation, or of direct analysis (when radioactive reagents were employed), it was possible to show that the infectivity was lost when one or two of the 6500 nucleotides were altered by any of these reagents.

Of particular interest was the observation that some of these reagents led to the appearance of mutants among the surviving particles. Such mutants, like natural strains or variants, were detected by the altered disease symptoms they produced. Mutants were observed not infrequently after methylation and bromination. Very frequent mutation, however, characterized only the deamination reaction [6]. This dramatic effect, first observed by MUNDRY and GIERER, has since been

[6] *Fraenkel-Conrat* - pag. 6

amply confirmed in other laboratories, and extended to other systems. Almost all survivors appear to be mutants if RNA is treated with nitrous acid. This extraordinary mutagenic efficiency may be attributed to the fact that this, and only this, reagent when acting on RNA (not DNA) transforms one nucleotide (cytosine) into another typical RNA compound, uracil. Thus, this reaction may alter the genetic message by changing one letter in the script, without introducing disfigured letters, as do other mutagenic reactions, such as methylation, bromination and others.

All types of mutants can be propagated and progeny virus can be isolated from them. Since many natural strains show differences in the amino acid composition from common TMV, it appeared of obvious interest to ascertain whether chemically evoked strains also show differences. This problem was attacked both by WITTMANN in Tübingen, and by TSUGITA in our laboratory. WITTMANN has concentrated his efforts on strains which are biologically not very different from the parent strain, while our laboratory has investigated more of the strains differing greatly from common TMV. Both laboratories have found strains not distinguishable from the parental in amino acid make-up. In both laboratories many strains were also detected, which differed by one replacement of an amino acid for another [7, 8]. In our laboratory several strains were also detected which showed 2 or 3 replacements, and another group (of 3 strains) showing 16 and 17 replacements and resembling a natural variant of TMV much more closely than they resemble the parental strain. Much further work is required before this last mentioned phenomenon can be rationally interpreted. At present it appears certain that the same type of protein alteration can result from different mutagenic reactions, and that actually there is a marked tendency for the same amino acid residues to become interchanged. Changes appear to be correlated with distinct biological effects, and to result with similar probability from any mutagen.

Recent studies on the viability of various strains, as ascertained from the infective centers present in a single local lesions under standard conditions, have led to some interesting conclusions. It appears that all strains, whether natural or chemically evoked, are less productive, or viable, than common TMV. The interpretation suggests itself that common TMV is the strain optimally adapted to the natural environment, and that all mutations lead to more or less inferior particles, of lower survival value under the existing circumstances. This conclusion is beautifully illustrated by the more detailed study of a particular mutant (≠ 171) isolated in our laboratory [9]. This mutant had 3 amino acid interchanges. One of these, the loss of one proline residue, was shown by TSUGITA to involve the proline near the C-terminus of the peptide chain, which as previously mentioned limits the attack of the enzyme, carboxypeptidase A, to the release of one residue, threonine. In contrast, from strain ≠ 171, threonine, alanine and leucine are rapidly released. This clearly shows that this particular proline had been replaced by leucine in strain ≠ 171. Furthermore, this observation illustrates how a chemical change in the RNA (deamination) can lead to a chemical alteration of the protein coded by that RNA (amino acid replacement), which directly can result in biological inferiority, due to an increased susceptibility to a degradative enzyme. Thus, a beginning understanding of virus structure and function has enabled us to consider evolution in chemical terms.

The rapid advances taking place in the field concerned with the mechanisms of transfer of information from DNA to RNA, and from RNA to protein, coupled with the availability of pure genetically interrelated nucleoprotein system in the viruses, presages a complete elucidation of the physico-chemical principles of heredity in the not too distant future.

[6] *Fraenkel-Conrat* - pag. 8

BIBLIOGRAPHY

[1] ANDERER F.A., UHLIG H., WEBER E. and SCHRAMM G., *Nature, 186,* 922 (1960).

[2] TSUGITA A., GISH D.T., YOUNG J., FRAENKEL-CONRAT H., KNIGHT C.A., STANLEY W.M., *Proc. Natl. Acad. Sci. U.S., 46,* 1463 (1960).

[3] ANDERER F.A.,, *Z. Naturforsch., 14-b,* 642 (1959).

[4] SUGIYAMA T. and FRAENKEL-CONRAT, *H. Proc. Natl. Acad. Sci. U.S., 47,* 1393 (1961).

[5] SCHUSTER H. and SCHRAMM G., *Z. Naturforsch., 13-b,* 697 (1958).

[6] GIERER A. and MUNDRY K.W., *Nature, 182,* 1457 (1958).

[7] TSUGITA A. and FRAENKEL-CONRAT H. In press. J. Molecular Biol. (1962).

[8] WITTMANN H.G., *Z.f. Vererbungslehre, 90,* 463 (1959).

[9] TSUGITA A. and FRAENKEL-CONRAT H., *Proc. Natl. Acad. Sci. U.S., 46,* 630 (1960).

DISCUSSION

RICH

I suppose it would be quite reasonable that one might imagine that you could detect a mutant showing an amino-acid exchange near the end of the protein and you might find a difference in the end group of RNA?

FRAENKEL-CONRAT

This would be quite possible, and we are investigating this possibility.

SCHRAMM

We did some experiments to degrade RNA by phosphorylase, and we found that the degradation is very slow. I wonder whether some experiments are done in your laboratory about stepwise degradation by phosphorylase?

FRAENKEL-CONRAT

We have not at our disposal a sufficiently pure polynucleotide phosphorylase, but have done extensive work using snake venom phosphodiesterase in limited and stepwise manner. We have in this manner tentatively established the sequence of 5 nucleotides from the glycol end.

TUPPY

Does TMV-protein, in your opinion, combine with other species of ribonucleic acids to form rods?

FRAENKEL-CONRAT

Only with the RNA of other strains of TMV. We get no definite rod formation when we let TMV protein interact with yeast RNA or with ascites tumor RNA.

JACOB

With respect to the mutants in which no change in the amino-acid sequence of the viral protein can be detected, I would like to ask whether there exists any evidence for or against the existence of an « early » protein, say an enzyme necessary for the multiplication of the viral RNA? Is it known, for instance, whether inhibition of protein synthesis at an early phase of infection prevents the synthesis of viral RNA?

FRAENKEL-CONRAT

Unfortunately the TMV-tobacco leaf system does not lend itself to the type of experiments necessary to answer conclusively the questions raised by Professor JACOB. As far as I know, no definite information is available concerning those questions.

SUR LE MODE D'ACTION DES GENES ET LEUR REGULATION

FRANÇOIS JACOB et JACQUES MONOD

Services de Génétique microbienne et de Biochimie cellulaire, Institut Pasteur, Paris (France)

Les méthodes de la biologie expérimentale, comme celles de la biochimie et de la génétique sont avant tout analytiques. Elles visent à disséquer la cellule en ses éléments. Au cours des cinquante dernières années, ces méthodes ont permis de réunir une masse étonnante de renseignements sur la structure des macromolécules, les voies de biosynthèse et le métabolisme énergétique de la cellule. Dans chaque cellule, un (ou deux) exemplaires d'une séquence donnée de nucléotides détermine une fonction différente de celle déterminée par d'autres séquences de nucléotides et il en est ainsi pour des milliers de fonctions différentes. Aujourd'hui, il ne semble pas déraisonnable d'espérer comprendre, dans un avenir proche, comme sont formées les macromolécules spécifiques constituant la cellule.

Dans le même temps, il est apparu clairement que chacun des éléments qui constitue le système cellulaire fonctionne dans des conditions beaucoup plus subtiles dans la cellule intacte qu'après isolement « in vitro ». Nous commençons à reconnaître non seulement les mécanismes qui permettent à la cellule de produire un constituant essentiel, mais aussi ceux qui lui permettent de ne pas le produire quand il n'est pas né-

cessaire. Ainsi, de même qu'un organisme pluricellulaire représente une société de cellules intégrées pour former l'individu, de même une cellule correspond en quelque sorte à une société de molécules organisées pour coopérer. Une société exige que chacun de ses membres soit soumis aux règles générales de la communauté; la désobéissance à ces règles pouvant se traduire par la destruction de l'harmonie communautaire, c'est-à-dire la mort de l'individu ou de la cellule. Une société implique donc un système complexe et précis de communications et de signaux entre les constituants membres de la communauté. Chez les organismes pluricellulaires, l'existence de systèmes complexes de communications intercellulaires chimiques ou nerveuses est connue depuis longtemps. A l'intérieur de la cellule, au contraire, c'est au cours des dernières années qu'ont été mis en évidence de tels systèmes de communication. Ce sont les systèmes réglant le transfert d'information intervenant dans la synthèse des protéines qui seront discutés ici. Ces résultats sont encore limités aux seules bactéries, organismes chez lesquels les différents modes de transfert de matériel génétique se prêtent particulièrement bien à l'analyse des mécanismes biologiques au niveau de la cellule.

Transfert d'information structurale dans la synthèse des protéines

Il est généralement admis aujourd'hui que la structure des protéines est déterminée par les gènes. Plus précisément la séquence des acides aminés dans le ou les polypeptides constituant une protéine particulière semble représenter une « transcription », selon un chiffre qui reste à découvrir, de la séquence des nucléotides dans un segment d'acide désoxyribonucléique.

Toutefois, il existe de nombreuses raisons de penser que l'agencement des acides aminés en chaîne polypeptidique ne

se fait pas directement sur le matériel génétique même, mais dans le cytoplasme, au niveau de centres formateurs de protéines, vraisemblablement les ribosomes. Le processus de transcription implique donc au moins deux étapes: synthèse par le gène d'un « intermédiaire » cytoplasmique porteur de l'information génétique; synthèse par l'intermediaire de la chaîne polypeptidique.

Ce n'est que récemment qu'on été obtenues, chez les bactéries, des précisions sur le mécanisme de transfert d'information des gènes aux centres formateurs de protéines et sur la nature de l'intermédiaire. Ce dernier paraît être un polyribonucléotide, désigné sous le nom d'ARN messager, de poids moléculaire élevé et formé à l'image de la séquence de désoxyribonucléotides du gène correspondant, par un enzyme spécifique. Cet ARN messager possède, en outre, la propriété de se fixer aux ribosomes et d'être métaboliquement instable, se renouvelant rapidement. Selon toute vraisemblance, les ribosomes se comportent comme des machines non spécialisées capables de former différents types de protéines. C'est l'ARN messager qui, en se fixant sur un ribosome lui dicte les « instructions » pour la synthèse d'une protéine donnée. Le fonctionnement de l'ARN messager entraîne rapidement sa destruction.

Selon ce schéma, qui fait du produit primaire des gènes un intermédiaire métaboliquement instable, le seul élément cellulaire contenant de façon durable l'information structurale pour la synthèse des protéines est le matériel génétique lui-même. L'expression d'un gène exige un flux continu d'information structurale du matériel génétique aux centres formateurs de protéines par l'intermédiaire d'ARN messagers. Les seules erreurs de copie affectant de façon durable la cellule sont donc celles qui atteignent le matériel génétique même, c'est-à-dire les mutations.

Systèmes génétiques réglant le transfert d'information structurale

Si la présence d'un gène donné est une condition nécessaire pour la synthèse de la protéine correspondante, elle n'est souvent pas une condition suffisante. Chez les organismes pluricellulaires, par exemple, il est généralement admis que chaque cellule contient l'ensemble des gènes de l'organisme quoique les cellules appartenant à différents tissus produisent des protéines différentes. En outre, l'expression d'un gène donné dépend souvent des conditions de milieu. Chez les bactéries, par exemple, la synthèse de certaines protéines peut être provoquée (induction) ou inhibée (répression) sous l'influence de facteurs spécifiques d'origine externe.

L'analyse de ces phénomènes a montré que, chez les bactéries, la transcription des gènes de structure, c'est-à-dire des gènes qui déterminent l'organisation moléculaire des protéines, est réglée par des systèmes complexes de signaux spécifiques intracellulaires. Des gènes, dits régulateurs, gouvernent l'expression des gènes de structure par l'intermédiaire d'un produit cytoplasmique spécifique ou répresseur. Un même gène régulateur, et par conséquent un même répresseur peut inhiber l'expression de plusieurs gènes de structure et, par conséquent, la synthèse de plusieurs protéines appartenant à une même séquence biochimique. On ignore encore la nature exacte du répresseur, mais il semble bien que ce soit par interaction avec les répresseurs endogènes que les facteurs externes exercent leur effet, soit en activant le répresseur endogène (répression), soit en l'inactivant (induction).

Le répresseur inhibe la transcription des gènes de structure par combinaison avec un récepteur spécifique désigné sous le nom d'opérateur. L'analyse a montré que, dans certains systèmes au moins, un seul opérateur gouverne l'expression de plusieurs gènes de structure adjacents, situés sur le même chro-

mosome. En d'autres termes, un opérateur sensible à un répresseur donné peut gouverner l'expression de plusieurs gènes de structure adjacents. Le segment chromosomique contenant un opérateur et les gènes de structure adjacents dont il gouverne l'expression constituent une unité de transcription coordonnée à laquelle a été donné le nom d'opéron. Il a pu être montré récemment que la régulation agit directement au niveau du matériel génétique sur la première transcription, c'est-à-dire sur la synthèse des ARN messagers par les gènes de structure.

Les phénomènes de régulation bactériens comportent donc des systèmes émetteurs-récepteurs de signaux intracellulaires spécifiques. A tout instant, en fonction de l'état du cytoplasme et des conditions du milieu, ils déterminent le taux de transcription des gènes de structure en permettant ou en bloquant la synthèse des messagers correspondants.

Régulation et évolution

Cette analyse des bactéries démontre que le matériel génétique ne contient pas seulement les plans d'architecture déterminant la structure de molécules protéiques individuelles, mais un programme de synthèses coordonnées et les moyens dont on controle l'exécution. Les phénomènes d'adaptation par mutation suivie de sélection peuvent donc affecter aussi bien que les structures protéiques elles-mêmes, les systèmes de régulation qui déterminent les taux de synthèse. Il est évident, par exemple, que toute modification du matériel génétique qui chez une bactérie permettra à une macromolécule de n'être synthétisée que lorsque nécessaire, donnera à l'organisme un avantage sélectif immédiat sur un organisme similaire dépourvu de ce système. On admet généralement que les gènes gouvernant une série de réactions séquentielles peuvent s'être formés par reduplications successives suivies de modifications. La même hypothèse peut être appliquée pour rendre compte de l'appari-

tion des systèmes émetteurs-récepteurs que constituent les gènes régulateurs-opérateurs.

Chez les bactéries, les systèmes de régulation ont essentiellement pour fonction de maintenir un état homéostatique et, en fonction de l'état des bactéries et des conditions de milieu, de déterminer à tout instant quelle fraction du génome est transcrite. Ce système est adapté à la croissance rapide d'un organisme unicellulaire chez lequel la régulation est parfaitement réversible, tout élément du génome pouvant à tout instant être exprimé si nécessaire.

La régulation des organismes pluricellulaires présente un ordre de complexité supérieur, car selon le principe des sociétés organisées, la faculté d'effectuer certains groupes de fonction est déléguée à certains groupes de cellules par la différentiation. Dans une cellule d'un tel organisme, aux systèmes de régulation homéostatique et réversible s'ajoutent donc des systèmes entraînant l'inactivation permanent de certaines fonctions. Bien que les mécanismes qui régissent la différenciation restent totalement inconnus, il semble probable qu'ils agissent aussi au niveau du matériel génétique pour déterminer quelles fractions du génome pourront être transcrites et bloquer de façon permanente la transcription du reste du génome. L'intérêt des systèmes régulateurs-opérateurs, tels qu'ils peuvent être définis chez les bactéries, est de fournir les éléments de base avec lesquels il est possible de construire des modèles de circuits répondant aux exigences les plus diverses. L'analyse génétique de cellules somatiques différenciées devrait permettre dans un proche avenir de reconnaître si la différenciation cellulaire est bien le résultat d'activation, ou d'inactivation, de segments génétiques par de tels circuits.

NOTE — Les arguments expérimentaux et les références concernant les problèmes du transfert d'information et de sa régulation dans la cellule bactérienne ont été rapportés en détail dans l'article: JACOB F. et MONOD J., *J. Mol. Biol.*, *3*, 318-356 (1961). Ce travail a bénéficié de subventions de la « National Science Foundation » et de la « Jane Coffin Childs Memorial Fund » des Etats-Unis d'Amérique, ainsi que du Commissariat à l'Energie Atomique.

DISCUSSION

CHANTRENNE

It will be interesting to find out to what extent this type of regulation mechanisms is operative in higher organisms and whether their operation can result in cell differentiation. The fact that protein synthesis continues for a long time after enucleation in Acetabularia and in reticulocytes indicates that the messenger has a longer life in higher organisms than in *E. coli*. There is also evidence from enucleation experiments that the synthesis of certain proteins is more sensitive to enucleation than that of other proteins, as if certain messengers had a much longer average life than others.

JACOB

With respect to your second question, one must clearly say that very little is known as yet about messenger RNA in organisms other than bacteria and yeasts. The only clear report, which has recently come from several laboratories, is that short pulses of radioactive bases to certain cells of mammals result in a preferential labelling of an RNA fraction which is distinct from ribosomal RNA, exhibits a rapid turnover and seems to have a higher molecular weight than transfer RNA. These properties seem to be very similar to those of the messenger found in bacteria. On the other hand, even in bacteria, we do not know precisely the rate of renewal of messenger RNA because of difficulties in estimations of the size of the nucleotide pools. It is not yet known whether one

messenger molecule can act as a template for the synthesis of one or several molecules of a protein. It is certainly an interesting possibility that different messengers may have a differential rate of renewal.

With respect to your first question, it is not possible for the present time to determine to what extent the type of regulation mechanisms found in bacteria are operative in higher organisms and play a role in differentiation. One of the paradox in differentiation is that, whereas the same set of genes seems to be present in all the cells of an organism, the expression of these genes varies from cell to cell. The problem is somewhat different from that raised in bacteria. In the bacterial cell, the regulation mechanisms which operate have been selected because they satisfy the requirements of an homeostatic system. At any time, any gene of the bacterial cell can be expressed or not, depending on the environmental and intracellular conditions. In differentiation, the genes appeared to be turned on or off in a permanent way according to a precise time schedule.

The study of regulation mechanisms in bacteria, however, provides a model which may account for the differential expression of genes in the cells of higher organisms. What the study of bacteria unravels are the elements of regulation circuitery: regulator genes-repressors-operators. In the same way as an engineer, given the elements of electronic circuiteries such as valves, relays, etc., is able to build electronic circuits satisfying the most various requirements, it is possible, using more or less complex systems of regulator genes-repressors-operators, to devise a variety of regulation circuits. For instance, one may produce a model circuit such that a group of genes could be switched from the inactive to the active state by transient contact with a specific inducer, produced for example only by another tissue. Once activated, the system could not be switched back. Such a model could account for some known features of cellular differentiation and it should be amenable to experimental test in the near future.

[7] *Jacob* - pag. 8

ZAMECNIK

Is it possible that a ribosome which is involved in making a particular protein at a rapid rate, such as would be the case where one had an induced enzyme, has a shorter life-span than the average ribosome in the same cell? One might regard the ribosomal template as becoming imperfect and no longer usable as a result of repeated use in the synthesis of a given long peptide chain. If one locked at the problem in this way, would it be possible to consider the labeled messenger RNA simply as a template being used more rapidly than the average ribosomal template, and therefore undergoing renewal at a similarly rapid rate?

JACOB

This seems to me rather unlikely for two reasons. First, as investigations proceed, we learn that the rate of synthesis of more and more proteins is under the control of some « inducible » or « repressible » system. Until a few years ago the synthesis of enzymes belonging to biosynthetic pathways was considered to be «constitutive», that is to be independent of environmental conditions until their repressible character was discovered. Recently, the synthesis of the enzymes of the KREBS cycle in bacteria has also been found, in Dr. KREBS' laboratory, to obey some regulatory mechanism. I really think that the synthesis of most, if not all, proteins will turn out to be regulated in some way. Furthermore, in the systems studied so far, i.e. catabolic systems such as lactose or galactose utilization, and biosynthetic pathways such as tryptophan or arginine synthesis, the breakdown of one regulation system by mutation results in an enormous constitutive production of the corresponding proteins. This production amounts to nearly 10% of the total protein synthesis in the investigated cases. It is clear therefore that a bacterial cell could not survive the breakdown of many regulation systems. I would like to think that all structural genes work in a similar way and in particular that they all eventually may be expressed at a similar maximal rate. The second argument comes from studies of protein synthesis in phage infected

bacteria. When *E. coli* are infected with T even phages, protein synthesis proceeds linearly, but no detectable amount of new ribosomes is made. The messenger RNA of the phage becomes associated with the ribosomes which were formed before infection and which, before infection when associated with bacterial messenger RNA, were manufacturing bacterial proteins. After 10 or 12 minutes, some 70 to 80% of the total proteins produced correspond to one molecular species, the phage head building block. It seems that after infection, the amount of ribosomes present is the factor limiting the rate of protein synthesis and that the bulk of the ribosomes is engaged in the production of phage head protein.

MESELSON

Doctor SYDNEY BRENNER and I have done heavy isotopic labeling experiments with ribosomes to see how long a ribosome once synthesized is conserved as such. We found that most of the 50 and 30S ribosomes of *E. coli* remain intact for many generations after they are synthesized. Only a small fraction of the ribosomes could be unstable in the manner suggested by Doctor ZAMECNIK.

ZAMECNIK

Do you have a picture of how the messenger RNA fits on the ribosome? Does it overlay an RNA molecule which is already wound around it?

JACOB

Unfortunately, we have no picture as yet. We begin to detect the various species of the system: the transfer RNA with the amino-acids which bring the building blocks, the messenger RNA which provides a template and the ribosome which apparently is the synthesizing machine. How these pieces fit together and how they work remain to be shown.

CHARGAFF

There are some slight chemical difficulties in understanding the very rapid turnover of the so-called messenger RNA. Chemically speaking, there is really not much difference or no difference whatever in the type of linkage holding together the nucleotides in s-RNA, in ribosome RNA and the messenger. You would have to postulate a very mystical type of structure to understand why this particular messenger RNA which is said to mimic a DNA should break down so particularly rapidly.

Moreover, there is really no proof that the messenger RNA — if it exists — imitates the sequence of DNA. All one can say is that in the total base composition, it is somewhat similar or more similar than the rest of the regular s-RNA. But this is in itself not sufficient.

JACOB

I agree that chemically speaking we do not know of any difference in the type of linkage in the various RNA species. Yet there are some known differences in their properties. For instance, it seems that ribosomal RNA and, I believe, transfer RNA are phosphorolized very slowly by the polynucleotide-phosphorylase, whereas TMV-RNA is phosphorolized at a much faster rate. Recent reports indicate that messenger RNA is also phosphorolized at a faster rate than ribosomal RNA.

Of course, there is no proof that the nucleotide sequence in the messenger RNA is indeed a copy of the sequence in the DNA since, for the present time, nobody can determine a nucleotide sequence. The total base composition and the DNA-RNA hybridization are the only available tests and it just happens that they fit for messenger but not for other RNAs.

WATER-INSOLUBLE ENZYME DERIVATIVES. THEIR PREPARATION, PROPERTIES AND USE IN THE STUDY OF NATIVE MACROMOLECULES

EPHRAIM KATCHALSKI

Department of Biophysics, The Weizmann Institute of Science,
Rehowoth (Israel)

The conversion of enzymes into water-insoluble products possessing specific catalytic activity is of interest since such « water-insoluble » enzymes may readily be removed from the reaction mixture, and can be used for the preparation of columns with enzymatic activity. If stable, they may be employed repeatedly to induce specific chemical changes in relatively large amounts of substrate. A short incubation of low or high molecular weight native substrates with a water-insoluble enzyme may facilitate the detection and isolation of some important intermediates formed during an enzymatic reaction. Furthermore, the study of the biological properties of the waterinsoluble enzyme derivatives may shed new light on the properties and mode of action of the native water-insoluble enzymes. In order to strengthen the linkage between the enzyme and the water-insolnble carrier, the attachment was carried out via covalent links with functional groups non-essential for enzymatic activity.

In the work hitherto carried out in our laboratory [1] hydrolases were the enzymes chosen for coupling with the water-

insoluble carrier. The synthetic procedures worked out, however, can be extended to other enzymes. To ensure high enzymatic activity different synthetic procedures were employed in the preparation of the various water-insoluble enzymes. A direct coupling of the enzyme with the chemically reactive polymer was carried out in the case of papain [2] and ribonuclease. A water insoluble pepsin was obtained by reacting pepsinogen with a water-insoluble polymer and subsequent acid activation of the insoluble pro-enzyme [3]. The attachment of the enzyme to the carrier by polymeric side chains which ensure free movement of the catalyst molecules in the reaction mixture was used in the case of trypsin and chymotrypsin [4]. In all the enzymatic preparations investigated, a copolymer of p-amino-*DL*-phenylalanine and *L*-leucine was used as carrier.

Since papain is known to have a relatively high content of tyrosine (13.7%) and approximately two-thirds of the molecule are nonessential for its activity [5], the enzyme was directly coupled with the diazotized copolymer of p-amino-*DL*-phenylalanine and *L*-leucine. The water-insoluble product obtained retained most of the activity of the original papain. It showed hydrolytic activity towards low and high molecular weight substrates parallel to that of the starting water-soluble papain. The insoluble papain derivative, which could easily be freed of reducing agents, was found to act on proteins in the presence of negligible amounts of cysteine. It was thus possible to carry out the digestion of proteins without reduction of substrate or products. Immune γ-globulin hydrolyzed briefly under such conditions, although susceptible to fragmentation by reduction, showed no alteration in its rate of sedimentation or in its ability to precipitate antigen [2].

Five minute incubation of rabbit γ-globulin with water-insoluble papain caused, in the absence of cysteine, the hydrolysis of approximately 3 to 5 peptide bonds per molecule of globulin. No appreciable change in the sedimentation (6.2 S) or the diffusion coefficients of the globulin occurred during this

procedure. However, when the papain-free hydrolysate was treated with cysteine or thioglycolic acid, the γ-globulin was completely transformed into three fragments with a sedimentation constant 3.3 S, similar to those described by PORTER [6]. When ovalbumin was treated with a papain treated anti-ovalbumin globulin, and the precipitate formed was reduced with thioglycolic acid, complete dissolution of the precipitate occurred. The solubilized precipitate gave two peaks in the ultracentrifuge with sedimentation coefficients of 9.4 S and 3.6 S. The component with the high sedimentation constant (9.4 S) was shown to consist of molecular aggregates, each composed of several antibody fragments bound to one antigen molecule [7]. The slowly sedimenting material, on the other hand, was found to represent the antibody fragment which is devoid of any combining site. Isolation and evaluation of the molecular weight of the fast sedimenting component enabled the conclusion that the soluble molecular complex consists, on the average, of 4.08 antibody fragment molecules attached to one antigen molecule. The soluble complexes of antigen and antibody fragments could be dissociated into their original components, at neutral pH, in the presence of detergents. Dissociation occurred also on lowering the pH to pH 4.0. By techniques similar to those described above complexes containing several different protein antigens, e.g. bovine serum albumin and lysozyme, have been isolated, their molecular weights have been determined, and they have been used for various dissociation studies [7]. Dissolved precipitates which contained antigen and antibody in molar ratios varying over a wide range from the region of antibody excess up to equivalent proportions, have been investigated by means of velocity sedimentation methods. The findings obtained have been used to further elncidate the nature of the original insoluble antigen-antibody aggregate.

Preliminary experiments have indicated that trypsin is readily inactivated by the diazonium salts of *p*-aminobenzoic

acid or *p*-amino-*DL*-phenylalanine. Water-insoluble products of low enzymatic activity were also obtained on coupling the enzyme with diazotized poly-*p*-aminophenylalanine or diazotized copolymer of *p*-aminophenylalanine and leucine. The direct binding of trypsin to a water-insoluble carrier via azo bonds was therefore abandoned and a different route for the preparation of the water-insoluble enzyme worked out. A polytyrosyl trypsin containing tyrosyl peptide side chains was prepared by initiating the polymerization of N-carboxy-*L*-tyrosine anhydride with trypsin. The trypsin derivative thus obtained was water-soluble and retained full enzymatic activity [8]. On coupling with a diazotized copolymer of *p*-aminophenylalanine and leucine it yielded the required water-insoluble trypsin.

The rate of hydrolysis of benzoyl-*L*-arginine methyl ester by water-insoluble trypsin at pH 7.8 and 25° indicated that the catalytic activity of the bound protein amounts approximately to a fifth of that of the free enzyme. The enzymatic activity of water-insoluble trypsin dropped slowly with time when kept under 0.0025 N hydrochloric acid at 2°. A decrease in activity of about 25% was recorded after a period of two months. The relative stability of the water-insoluble trypsin is most likely due to the blocking of the ε-amino groups of the enzyme on the one hand, and to the fixation of the enzyme molecules to the water-insoluble carrier on the other. The former leads to a decrease in the number of enzyme peptide bonds susceptible to tryptic digestion, while the latter inhibits interaction between trypsin molecules, thus diminishing autodigestion. The proteolytic and esteratic activities of water-insoluble trypsin are inhibited by di-isopropyl fluorophosphonate. The rate of inhibition is, however, considerably slower than that of trypsin. Addition of an excess of soybean or pancreatic trypsin inhibitor caused only partial inhibition of the enzymatic activity of the water-insoluble enzyme.

The water-insoluble trypsin was used to digest rabbit myosin to light and heavy meromyosin.

[8] *Katchalski* - pag. 4

A trypsin column was prepared by mixing water-insoluble trypsin with an inert polyvinyl resin and pouring the mixture into a glass column. On passing through the column a solution of poly-*L*-lysine hydrobromide, the effluent was found to contain dilysine, trilysine, tetralysine and traces of lysine. Protamine and insulin, oxidized by performic acid, could also be digested on the column.

An equation describing the effect of enzyme concentration (E_o), height of column, h, and rate of flow, V (given in cm./sec.) of the substrate, S, on the extent of hydrolysis, may be derived readily from the well known Michaelis-Mention relation:

$$-\frac{d(S)}{dt} = \frac{k(E_o)(S)}{K_M + (S)} \qquad (1)$$

In eq. (1) K_M is the MICHAELIS constant and k is the specific constant determining the rate of decomposition of the enzyme substrate complex. Eq. (1) yields on integration:

$$(S_o) - (S_h) + K_M \ln \frac{(S_o)}{(S_h)} = k(E_o)t \qquad (2)$$

where (S_o) is the initial concentration of substrate and (S_h) that of the effluent. Substitution of the time of exposure of substrate to enzyme in the column by the ratio h/V, gives

$$(S_o) - (S_h) + K_M \ln \frac{(S_o)}{(S_h)} = k(E_o)\frac{h}{V} \qquad (3)$$

Eq. (3) shows that the extent of degradation of substrate is determined by the effective concentration of enzyme, by the height of the column, and the rate of flow of substrate.

When $(S) \geqslant K_M$ eq. (3) reduces to eq. (4)

$$\frac{(S_o) - (S_h)}{(S_o)} = \frac{k(E_o)h}{(S_o).V} \qquad (4)$$

The last equation could be verified experimentally for the case of *L*-arginine methyl ester. In the presence of a relatively large excess of substrate the amount of ester hydrolysed was inversely proportional to the rate of flow and directly proportional to the height of the column.

Water-insoluble ribonuclease and deoxyribonuclease were prepared by coupling of the corresponding enzymes with a diazotized copolymer of *p*-aminophenylalanine and leucine. Both preparations showed low activity towards RNA and DNA respectively. Preliminary data indicate, however, that water-insoluble RNAse has a relatively high activity towards the low molecular substrate, cytidine -2′,3′-cyclic phosphate. It is thus possible that the interaction between the chemically bound enzyme and the high molecular weight substrates is sterically hindered by the water-insoluble carrier. In this connection it is pertinent to note that most of the hydrolases acting on high molecular weight substrates appear in the cell in free form; enzymes acting on small molecules, on the other hand, such as the oxido-reductive enzymes, are often attached in the cell to water-insoluble membranes. In the first case it is the enzyme which approaches the high molecular weight substrate, while in the second case it is the low molecular weight substrates which approaches the enzyme.

The examples given above illustrate some of the possible methods for the synthesis of water-insoluble enzymes. The principles used, particularly that of attachment of the enzyme to the water-insoluble carrier via a linear polymeric chain, might be used for the preparation of other water-insoluble proteins with biological activity. In such cases it will be of primary importance to determine the functional groups which are non-essential for the activity, and to find the suitable chemical reactions for forming the desired biologically active water-insoluble products.

[8] *Katchalski* - pag. 6

The water-insoluble hydrolytic enzymes prepared so far should be of particular use, as demonstrated above, in the degradation studies of peptides, proteins and nucleic acids.

The columns with enzymatic and other biological activities open up new possibilities for basic and applied research. They may lead to new models for the study of specific surface reactions, and to the development of novel laboratory and industrial techniques for the isolation and modification of biologically important compounds.

REFERENCES

[1] KATCHALSKI E., *Monograph on Poly-alpha-amino Acirs*. University of Wisconsin Press, Madison, Wisconsin (1962), in press.

[2] CEBRA J.J., GIVOL D., SILMAN H.I. and KATCHALSKI E., *J. Biol. Chem. 236*, 1720 (1961).

[3] NEUMAN H. and KATCHALSKI E. Unpublished results.

[4] BAR-ELI A. and KATCHALSKI E., *Nature, 188*, 856 (1960).

[5] HILL R.L. and SMITH E.L., *J. Biol. Chem., 235*, 2332 (1960).

[6] PORTER R.R., *Biochem. J., 73*, 119 (1959).

[7] CEBRA J.J., GIVOL D. and KATCHALSKI E., *J. Biol. Chem., 237*, 751 (1962).

[8] GLAZER A.N., BAR-ELI A. and KATCHALSKI E., *J. Biol. Chem., 237*, 1832 (1962).

DISCUSSION

TISELIUS

Thank you Dr. KATCHALSKI for this very interesting presentation. You certainly gave us the impression that one can do a great many things in these new solidified enzymes. May I perhaps ask if there is evidence from your presentation that one could use this method to follow successive steps in the successive enzymatic breakdown of large molecules as proteins, nucleic acids and so on on columns of the materials prepared by you. In this connection it would be extremely interesting if you could isolate the very first stage in the breakdown.

Furthermore, what have you done with your antibodies?

KATCHALSKI

The data obtained so far suggest that water-insoluble derivatives of hydrolytic enzymes might be of considerable use in the study of the successive stages of the enzymatic hydrolysis of high molecular weight substrates, such as proteins and nucleic acids, as well as in the preparation of the very first degradation products. By means of water-insoluble trypsin, for example, it was possible to degrade myosin under controlled conditions, and to prepare light and heavy meromyosin after a very limited extent of hydrolysis (1). Water-insoluble papain was used to hydrolyse various proteins, such

(1) A. BAR-ELI and E. KATCHALSKI, *Nature, 188*, 856, 1960.

as bovine serum albumin, ovalbumin and γ-globin in their native state. When rabbit antibody to ovalbumin was incubated with the water-insoluble enzyme for three to five minutes, three to four peptide bonds were split. The thus modified γ-globulin retained full immunological activity (2). The water-insoluble papain was also useful in the preparation of soluble complexes of antigen and antibody fragments (3). These complexes represent a most suitable system to elucidate the nature of the forces prevailing between antigen and antibody. I hope that they will be of use in the evaluation of immunological parameters such as specificity and valence of antigen.

TISELIUS

About the theoretical treatment of these columns: If you have some components which are retarded in their movements through your enzyme column, will they have more time to become attacked? In this case you will get a rather complicated theoretical treatment!

KATCHALSKI

The theoretical treatment put forward to describe the extent of substrate degradation by an enzyme-column is a very elementary one. I feel, however, that it can readily be extended to more complicated cases. If some of the components formed are retarded in their movement throngh the column, their contact with the enzyme is prolonged and further degradation can be expected.

SPIEGELMAN

Doubts have been recently raised concerning the precision with which one can define PORTER's fractions I and II.

I wonder whether your experience has led you to any conclusions on this question?

---

(2) J.J. CEBRA, D. GIVOL, H.I. SILMAN and E. KATCHALSKI, *J. Biol. Chem.*, *236*, 1720, 1961.

(3) J.J. CEBRA, D. GIVOL and E. KATCHALSKI, *J. Biol. Chem.*, *237*, 751 (1962).

KATCHALSKI

We have not carried out any new experiments to characterize PORTER's fractions I and II. We could, however, repeat with ease the separation of the fractions by the procedure described (4).

JACOB

Do you know whether the antibodies treated with the polymer containing papain, without treatment with cysteine, still exhibit their allotypic specificities?

KATCHALSKI

No attempt was made till now to determine the allotypic specificities of antibodies treated with water-insoluble papain in the absence of cysteine. Such a study might be of considerable interest.

LIPMANN

May I ask more information about insoluble RNAse?

KATCHALSKI

The water-insoluble ribonuclease (RNAse) was obtained by the coupling of RNAse, via azo links, to a copolymer of *p*-aminophenylalanine and leucine. Preliminary data indicate that the water-insoluble enzyme has a relatively high activity towards cytidine-2′, 3′-cyclic phosphate (20% of that of intact RNAse) and a relatively low activity towards RNA (1.5 to 2% of that of RNAse). It seems that the activity of the bound enzyme, particularly towards the high molecular weight substrate, is hindered sterically by the insoluble carrier. It is possible that a water-insoluble enzyme with a considerably higher activity towards RNA will be obtained by the attachment of the enzyme to the insoluble carrier *via* a long, flexible, hydrophylic side chain.

(4) R.R. PORTER, *Biochem. J.*, *73*, 119, 1959.

MYOGLOBIN: STRUCTURE AND FUNCTION (*)

J. C. KENDREW

Medical Research Council - Laboratory for Molecular Biology,
Cambridge (England)

The X-ray analysis of a crystalline protein became feasible in principle when it was shown by PERUTZ *et al.* [1] that the method of isomorphous replacement could be applied to crystals as complex as those of proteins. This method involves the labelling of each molecule in the crystal with various groups containing heavy atoms, the consequent changes in the intensities of the X-ray reflexions being used to determine the phases of the reflexions; solution of the structure then becomes a straightforward problem in computation — straightforward but of such great magnitude that it is necessary to make use of the fastest electronic computers available. The amount of data to be handled is roughly proportional to the molecular weight of the protein; myoglobin was selected by the present author partly because, as proteins go, it is a very small one, and partly because it is a representative member of the class of haem proteins. Though myoglobin itself has been relatively neglected, the physical chemistry and physiological functions of haem proteins as a class have probably been studied more intensively than those of any other kind of protein.

Even the simplest protein is so much more complex than

(*) The discussion of this paper follows Prof. PERUTZ's lecture.

the molecules generally studied by X-ray analysis that a complete solution at one stroke would be too much to hope for. By attacking the structure in stages, of increasing resolution, the method could be tested out as the work progressed and successively sharper pictures of the molecule could be obtained. The stages chosen in myoglobin were resolutions of 6A., 2A., and 1.5-1.3A.: these involved the handling of about 400, 10,000 and 25,000 reflexions respectively (the number of reflexions increases as the cube of the resolution desired). The first stage of the myoglobin analysis was completed in 1957 and the second in 1959; the third stagt is still in progress. One reason for the choice of these particular steps was that they corresponded roughly to the degree of resolution required to « see » the tertiary, the secondary and the primary structures of the molecule, respectively — that is, the spatial arrangment of segments of polypeptide chain producing a compact molecule, the type of folding in individual segments of chain, and the identity and interactions of the side chains. In the event it transpired that more information could be derived at each stage than had been predicted.

Myoglobin at 6A. resolution

The first stage of the analysis [2] resulted in a three-dimensional representation of the molecule at a resolution of 6A., which was insufficient to reveal the side chains or even the configuration of the polypeptide chain itself, but which did give a picture of the general arrangement of the polypeptide chain and of the haem group — in other words of the tertiary structure of the molecule (Figure 1). This tertiary structure proved to be unexpectedly irregular and complex: clearly the simple geometrical arrangements proposed in many earlier theories of protein structure could not be correct.

[9] *Kendrew* - pag. 2

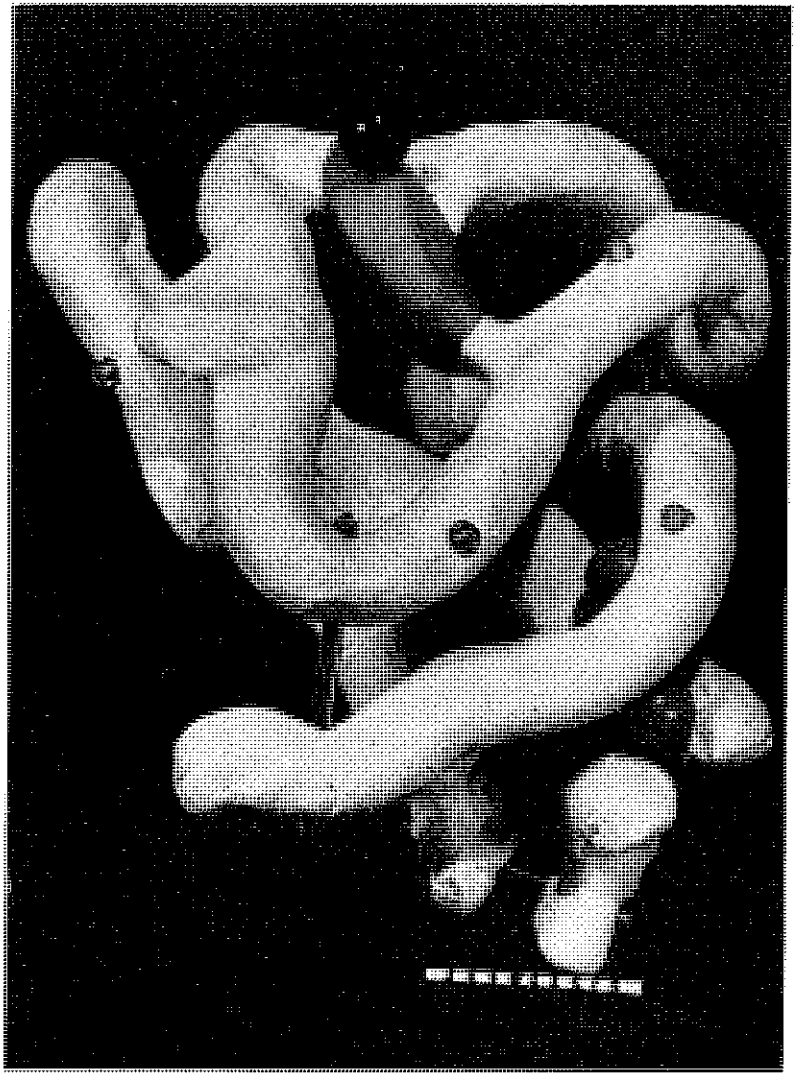


FIG. 1 — Model of the myoglobin molecule derived from the 6A. Fourier synthesis.

MYOGLOBIN AT 2A. RESOLUTION

More recently the resolution of the analysis has been increased to 2A. [3], [4] (see Figure 2). Although neighbouring covalently bonded atoms are still not distinctly resolved, it is now possible to separate atoms which are hydrogen-bonded or in VAN DER WAALS contact, with the result that the atomic arrangement of most of the molecule can be inferred. At this resolution the appearance of the haem group with its central iron atom corresponds closely with theoretical expectation (see Figure 3), thus providing a valuable check on the correctness of the analysis, and it can be seen that the iron atom is attached to a neighbouring polypeptide chain by means of a group which is almost certainly the imidazole ring of a histidine residue. Most of the polypeptide chain consists of a series of segments of right-handed α-helix, with parameters closely corresponding to those proposed by PAULING and COREY in 1951 [5] (see Figure 4). This was the first time that the α-helix had been directly seen in a protein molecule; the results confirmed earlier plausible evidence that this configuration was important in the structure of globular proteins as well as in those of many fibrous proteins. It turns out that the myoglobin molecule consists of eight segments of α-helix joined by irregular regions of varying length: the helical segments comprise 75% of the amino-acid residues, in agreement with estimates made on the basis of optical rotation and deuterium exchange studies.

One reason for the choice of 2A. as the resolution of the second stage of the myoglobin analysis was that the secondary structure (or polypeptide chain configuration) should be clearly revealed at this resolution, as indeed proved to be the case; it was not anticipated that side chains would be identifiable, so that progress could be made towards a direct determination of the amino acid sequence by X-ray methods alone. In the event expectation was much exceeded, for it was possible to

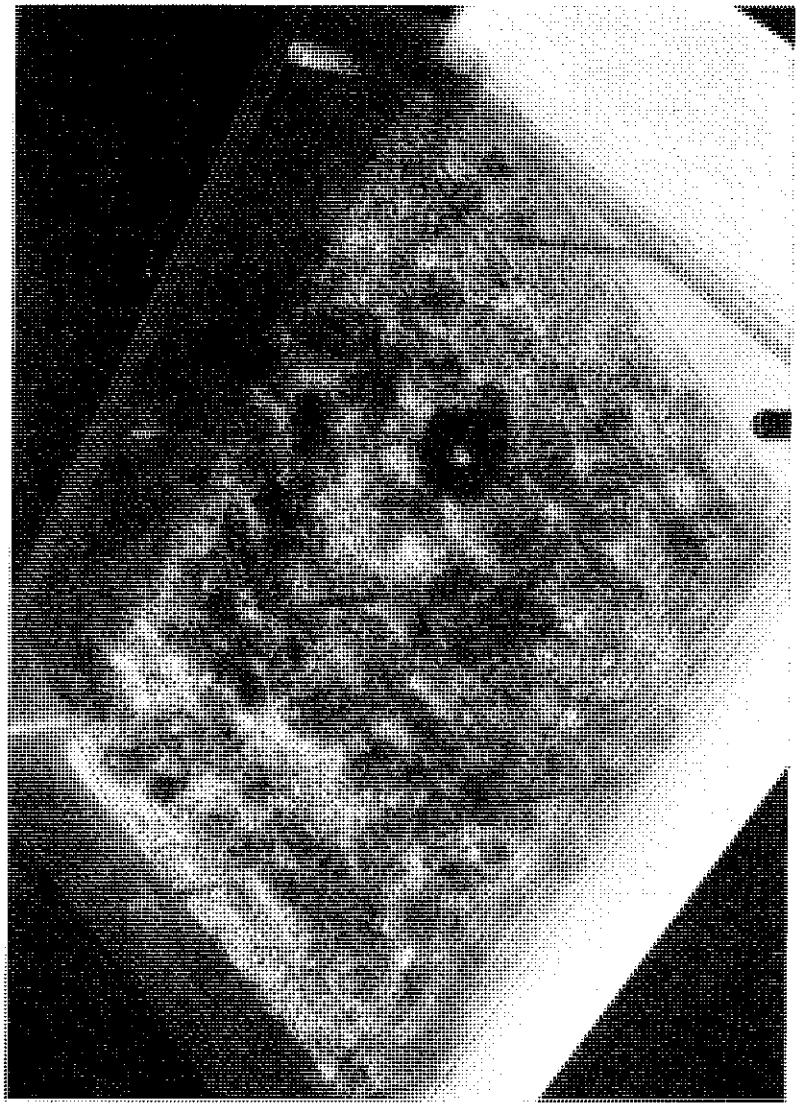


FIG. 2 — 2A. Fourier synthesis of myoglobin with end-on view of a helical segment of polypeptide chain, showing the hole down the centre.

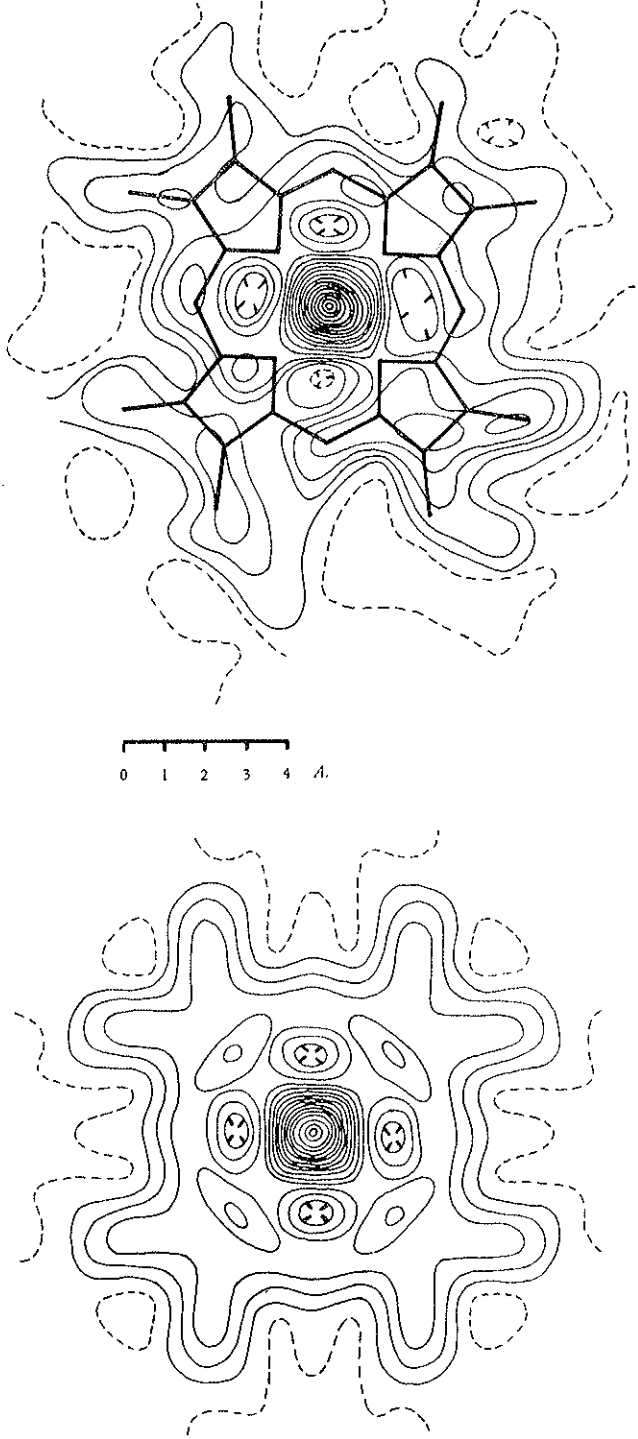


FIG. 3 — The haem group at 2A. resolution as observed in the myoglobin Fourier synthesis (above). Below is a theoretical Fourier synthesis at the same resolution, for comparison.

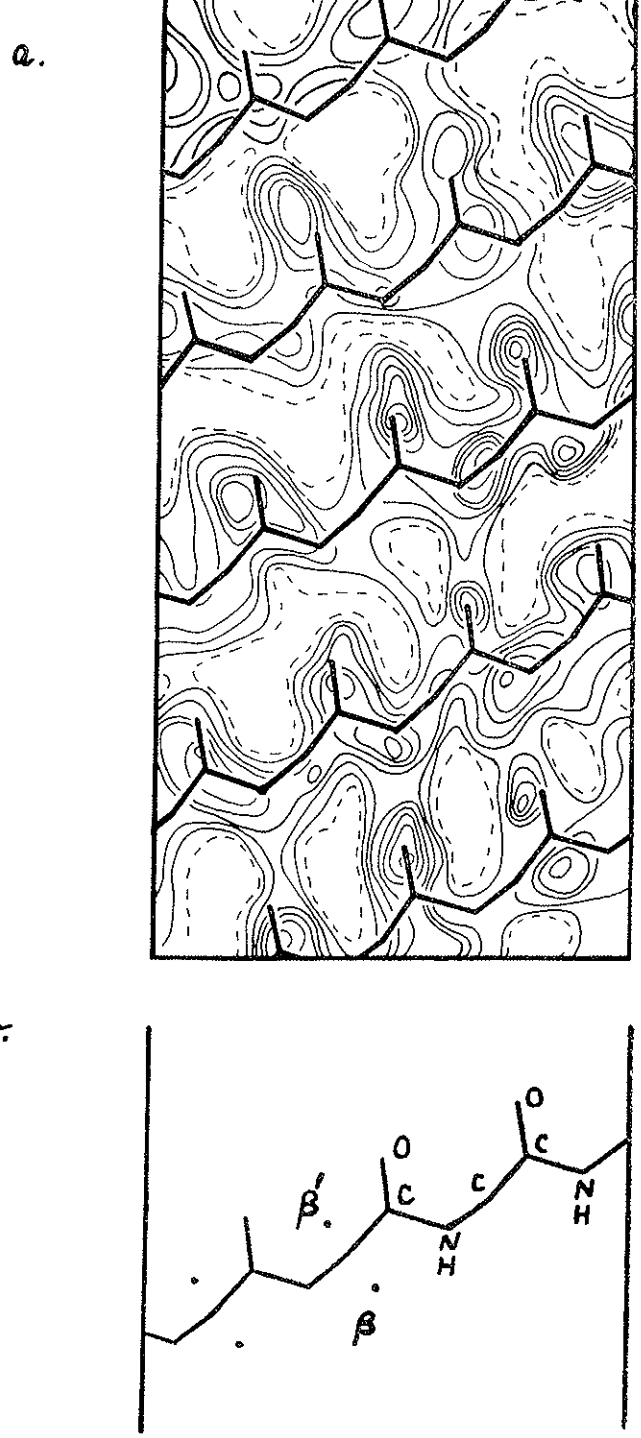


FIG. 4 — (*a*) Cylindrical section of a helical segment of polypeptide chain, in the 2A. Fourier synthesis of myoglobin; a length of PAULING α-helix is superposed for comparison. (*b*) Key to the structure of the *α-helix*.

see the individual side chains as dense regions emerging from the helical main chain at intervals corresponding to the parameters of the α-helix, and a close examination of these regions often made it possible to identify side chains with certainty; in other cases some ambiguity remained but the choice of side chain could be reduced to two or three (see Figure 5). At the present resolution about one-third of the side chains can be identified with certainty, and another third with fairly high probability (see Figure 6). It now becomes possible to cor-

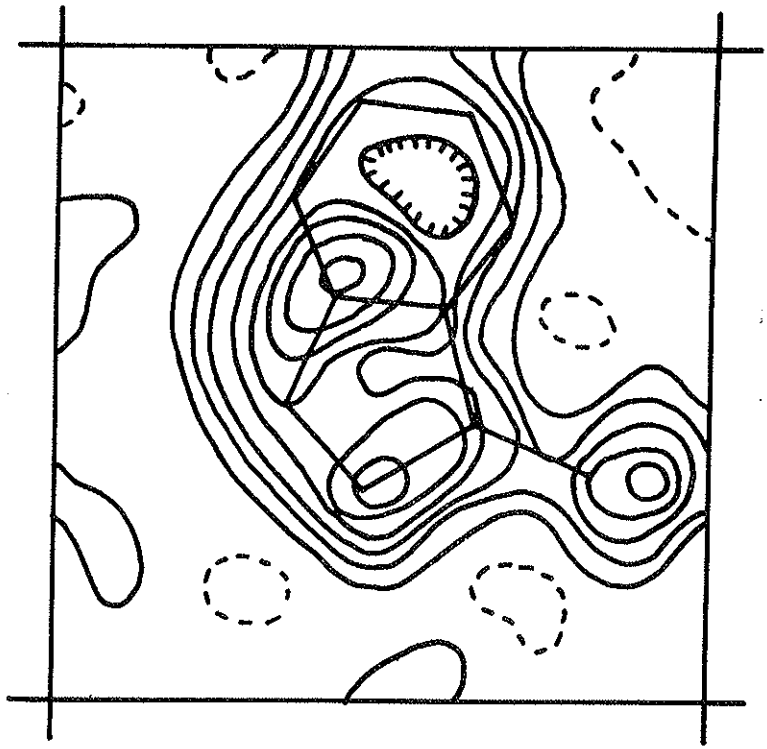


FIG. 5 — An example of a side chain at 2A. resolution, as seen in the 2A. Fourier synthesis of myoglobin.

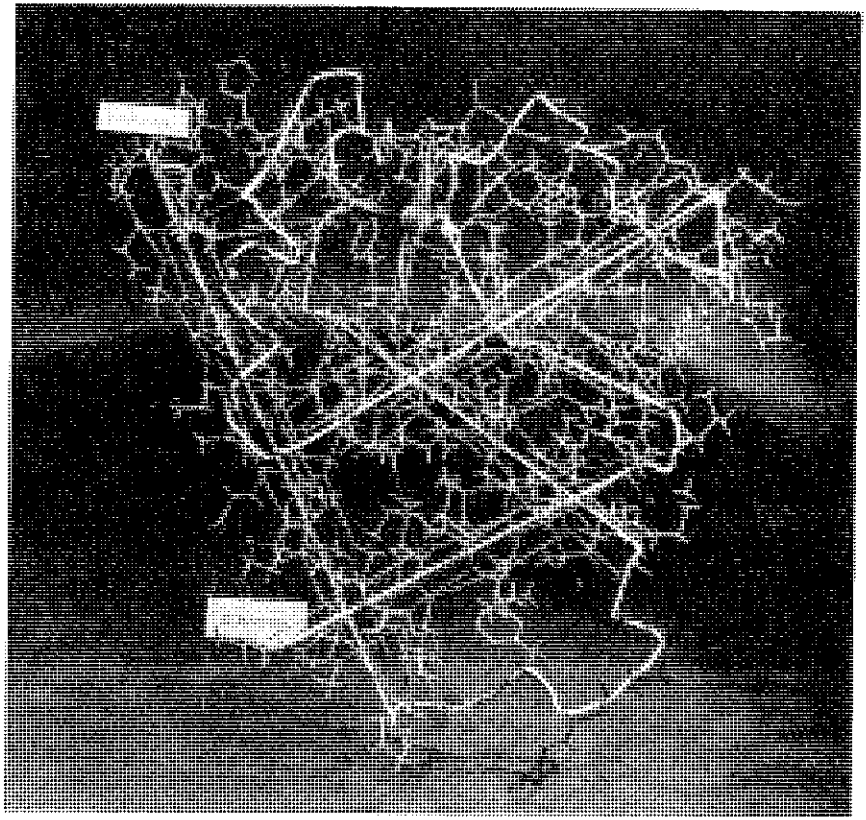


FIG. 6 — Model of the myoglobin molecule derived from the 2A. Fourier synthesis, including most of the side chains.

relate these X-ray results with the preliminary data obtained by Dr. A. EDMUNDSON, who is engaged in working out the amino acid sequence of myoglobin by chemical methods [6]. He has broken down the molecule into short peptides by means of tryptic digestion, and he has determined the composition, and in a few cases the internal sequence, of these peptides. By comparing his results with our own it has proved possible to

place almost all the tryptic peptides along the polypeptide chain, and the order of peptides thus ascribed corresponds with the order suggested in a few cases by EDMUNDSON on the basis of preliminary work using chymotryptic digestion. Although some discrepancies remain and although the amino acid sequence has not yet been completely determined, its main features are now beyond doubt. We are at present engaged in an attempt to increase the resolution of the X-ray results still further, and we hope that the remaining ambiguities will then be removed.

GENERAL NATURE OF THE STRUCTURE

Without going into details, we may summarize some general features of the structure as it has so far been revealed.

1) *The molecule is compact.* There is no water inside the molecule, with the exception of a very small number of single water molecules presumably trapped at the time the molecule is folded up. There are no channels through the molecule.

2) *Almost all the polar groups are on the surface.* With very few exceptions all the lysine, arginine, glutamic, aspartic, histidine, serine, threonine, tyrosine and tryptophan residues are on the outside of the molecule; the rare exceptions are polar side chains performing a special function within the molecule, e.g. the haem-linked histidine.

3) *The interior of the molecule is made up of non-polar residues,* almost everywhere close packed, in VAN DER WAALS contact with their neighbours. Not all the non-polar residues are inside; some, especially the short ones (glycine, alanine) are at the surface.

4) *Bound water molecules* are attached to all polar groups at the surface, including main chain CO and NH groups. Apart

from these bound water molecules there is no obvious sign of order in the liquid regions; these regions have not yet, however, been studied in detail.

In general it is clear that the VAN DER WAALS interactions between non-polar residues are quantitatively much more important in stabilizing the structure than are electrostatic and hydrogen-bond interactions between polar groups. On the other hand it may well be that certain key polar interactions have a special significance in determining the overall configuration of the molecule. In particular, interest is focussed on the non-helical regions and on the forces which determine the point at which each helical segment terminates. Proline, which cannot be accommodated within an α-helix, is clearly important here, but there are 8 non-helical regions and only 4 prolines. Several types of polar interaction seem also to be important in this connexion, but they do not occur more than once each, so generalization is impossible — indeed all the non-helical segments are different.

The interactions of the haem group deserve special consideration, since they are largely responsible for the physiological function of myoglobin. An isolated haem group does not exhibit the phenomenon of reversible oxygenation; its special behaviour in myoglobin and haemoglobin must be a consequence of its interactions with the protein moiety of these molecules. We have seen that the fifth co-ordination position of the iron atom is occupied by a ring nitrogen atom of a histidine residue; on the other side of the haem group, and occupying its sixth co-ordination position, is a water molecule, as would be expected in ferrimyoglobin in the region of pH 7. Beyond the water molecule, in a position suitable for hydrogen-bond formation, is a side chain which is almost certainly a second histidine. These results strikingly confirm earlier suggestions based on chemical evidence (for a discussion of these, see the review by WYMAN) [7]. The polar propionic acid side groups

of the protoporphyrin are at the surface of the molecule; the rest of the group is in the interior and is surrounded almost entirely by non-polar residues with which it is in VAN DER WAALS contact. It is difficult to see how at any rate the larger ligands known to combine with the haem group can do so without causing a substantial rearrangement of side chains, for the sixth co-ordination position of the iron atom is by no means exposed to the ambient solution.

STRUCTURE AND FUNCTION

Though we can now specify with some precision the environment of the haem group in myoglobin and haemoglobin, we do not yet know by what mechanism this special environment modifies the properties of the group so that it combines reversibly with oxygen and displays the other subtelties of behaviour which have been studied in such detail by several generations of biochemists. The solution of this central problem lies in the future — as does that of a much wider one, the problem of explaining protein behaviour generally in terms of protein structure. The structural features observed in myoglobin no doubt exemplify general principles, some of which may emerge from the comparative studies of myoglobin and haemoglobin now in progress. It is now clear, however, that proteins display great variety of structure as well as of function: generalization will not be possible until the structures of some other proteins have been worked out in detail.

REFERENCES

[1] GREEN D. W., INGRAM V. M. and PERUTZ M. F., *Proc. Roy. Soc., A225,* 315 (1954).

[2] BODO G., DINTZIS H. M., KENDREW J. C. and WYCKOFF H. W., *Proc. Roy. Soc., A253,* 70 (1959).

[3] KENDREW J. C., DICKERSON R. E., STRANDBERG B. E., HART R. G. and DAVIES D.R., *Nature, 185,* 422 (1960).

[4] KENDREW J. C., WATSON H. C., STRANDBERG B. E., DICKERSON R. E., PHILLIPS D. C. and SHORE V. C., *Nature, 190,* 666 (1961).

[5] PAULING L. and COREY R. B., *Proc. Nat. Acad. Sci.,* Washington, *37,* 235 (1951).

[6] EDMUNDSON A. B., and HIRS C. H. W., *Nature, 190,* 663 (1961).

[7] WYMAN J., *Advances Protein Chem., 4,* 410 (1948).

STRUCTURAL STUDIES OF METAL ION - ENZYME INTERACTION

I. LINDQVIST

Uppsala Universitet, Kemiska Institutionen, Uppsala (Sweden)

By tradition the inorganic chemistry and the biochemistry have been considered as representing the two branches of pure chemistry most different from each other. Accordingly the mutual knowledge of the other field has often been very small without causing much trouble for any part. When it comes to the role played by metal ions in biochemical reactions, and particularly in enzymatic processe, such ignorance might, however, be unfortunate. The biochemist ignorant about basic inorganic chemistry, might very well suggest mechanisms and models which are in serious disagreement with our basic knowledge about the laws governing the complex formation of metal ions in aqueous solutions. In the same way the inorganic chemist with a rather dim view about the biochemical processes might believe that he is studying model systems for some biochemical reaction although a serious consideration of the kinetics of that reaction would show that this cannot be the case.

During the last ten years many steps have been taken towards a deeper understanding of the function of metal ions in enzymatic catalysis and these efforts have been excellently summed up in some recent review articles [1]. It has thus

been shown that the kinetics of the process must first be studied in detail in order to prove that the metal ions are really activating it. Even then it remains to investigate if the metal ions can interact with both enzyme and substrate strongly enough to justify the suggestion that the metal ions can have the function of joining the two reactants. Studies of this type have been carried out for a number of metal ion activated enzymes during the last years and it therefore seems to be proper time to start accurate crystal structure determinations to obtain a better knowledge about the detailed nature of the interactions involved. This cannot possibly be achieved only by structure determinations of the whole enzymes, because, as has here been pointed out by KENDREW and PERUTZ, in a protein structure determination we cannot hope to come down to a better resolution than 1.5 Å. We must thus study some model compounds and the information we reach in this way is of course not confined to interaction with enzymes but will help the understanding of all metal ion - protein interaction.

As an example of a possible model for metal ion-active site interaction we can choose the zinc ion interaction with carboxy peptidase. It has recently been presented strong evidence [2] that the zinc ion is bonded to one cysteine and to one nitrogen containing residue (possibly histidine) of the carboxypeptidase molecule, and that the chelation with 2-mercaptoethylamine is a model for this interaction. We had indepently come to the same conclusion and therefore started a structure determination of the zinc salt $Zn(SC_2H_4NH_2)_2$ (STRANDBERG and ÅKERBY). The structure has not been finally refined but the present picture, given in Fig. 1 cannot be much changed. The most obvious features are: 1) the distorted tetrahedral coordination around zinc; 2) the assymmetry of the molecule formed. (We will try to separate the two enantiomorphic crystal types and study the optical activity); 3) the Zn-S bond length is 2.37Å, almost the same as in ZnS(2.35 Å); 4) the Zn-N bond length is 2.06 Å; 5) the bond angle S-Zn-N within each chelate ring

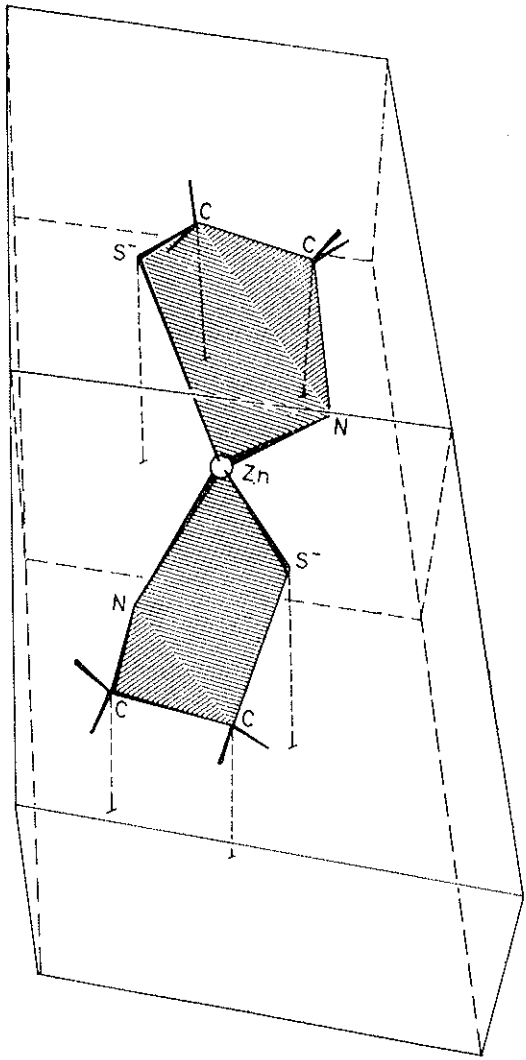


FIG. 1 — The coordination around the zinc ion in zinc(2-mercaptoethylamine)$_2$.

is as small as 89°. Any model considerations of chelation with zinc must thus take the possibility of great variations in the bond angle into account.

The weak point of this model is certainly the amino group which is rather different from the imidazol group in histidine. One possibility would be to try to compare a more similar chelate but first we wanted to see if the Zn-N bond lengths with imidazol would be very different. (The great stability of a chelate is mainly due to the entropy term and not to a difference in bond strengths). Crystals of $ZnCl_2.4C_3N_2H_4$ and $Zn(NO_3)_2.4C_3N_2H_4$ have therefore been prepared and the crystal structures are being studied. No final results have, however, been obtained as yet. Parallel studies of the imidazol salt obtained with zinc in alkaline solutions, $Zn(C_3N_2H_3)_2$, give rather surprising results, shown in Fig. 2 (STRANDBERG). The average Zn-N bond length (in the unrefined structure) is as short as 1.8 Å. Even if we could expect a stronger interaction with the negative ion, the negative charge being equally distributed on the two nitrogen atoms, than with the molecule, the difference from the value found in the preceding compound is remarkable. In fact the imidazol molecule could very well give as strong bonds as the negative ion, if the resonance form placing a negative net charge on the ligand nitrogen atom and a positive net charge on the other nitrogen atom is of importance. All these questions will find an answer when the structures of the neutral imidazole complexes have been studied.

It is not only the interaction with the protein that is of interest but also with the substrate. The structural studies have also so far mainly dealt with peptides as substrates. Extensive studies have been made by FREEMAN in Sydney of the structures of copper complexes with different polypeptides. The detailed results will not be given here, only some general observations of interest. The terminal amino group is always involved in the complex formation and also the terminal car-

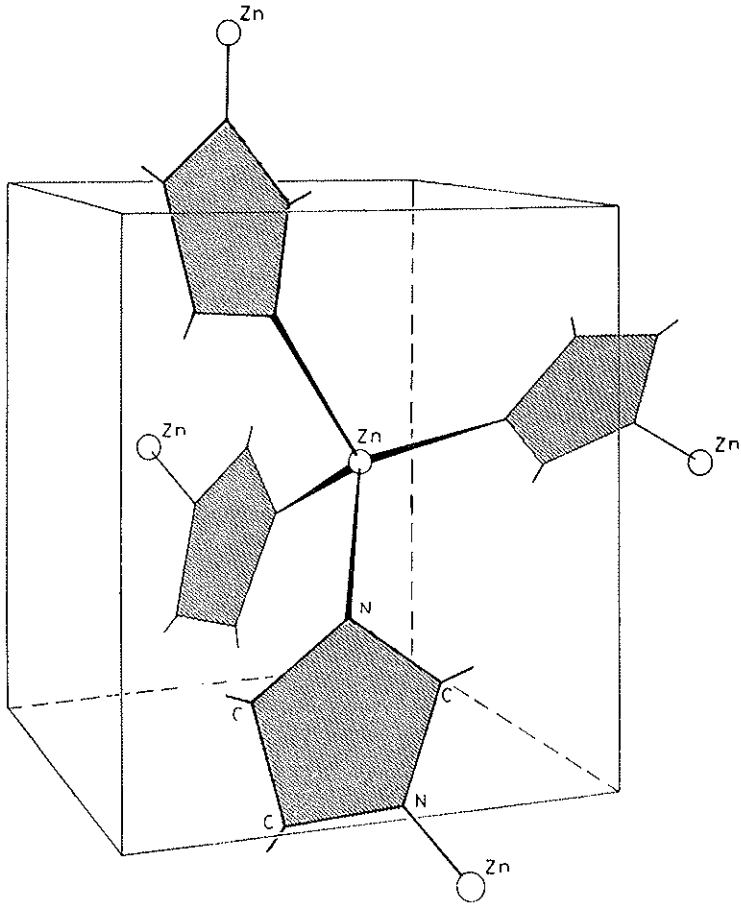


FIG. 2 — The coordination around the zinc ions in zinc(imidazol)$_2$.

boxylic group. In some cases chelation can be obtained with the carbonyl group next to the terminal amino group. The peptide nitrogen is never used for complex formation, unless the proton is split off, making a lone electron pair available for complex formation. This is the case in the copper compound studied by us $Cu(NH_2CH_2CONCH_2COO).3H_2O$. The structure [3] of the chelate molecule (two independent molecules) is shown in Fig. 3. Interesting features are the short bond length to the peptide nitrogen, 0.15Å shorter than to the amino group and the fact that this strong interaction does not change the general shape of the dipeptide appreciably. The addition of a fifth ligand (a water molecule) on a much longer distance is also noticeable. The dichelate molecules are joined with hydrogen bond to each other and to interstitial water molecules.

It is obvious that the structural studies made so far does not solve any biochemical problems but the information obtained about bond lengths and bond angles will certainly be necessary for the final clarification. We have tried to use the information we have obtained so far by considering the possibility of chelate formation between zinc ions and a straight helical chain. We have then used the same assumptions about the structure of the chain as KENDREW in the detailed discussion of the myoglobin structure and added our knowledge about the complex formation with zinc ions. We then find as a general rule that chelation with two side groups is only possible if there are three amino acids between the two chelating groups. Even then there are rather restricted possibilities, but it was shown that cysteine and histidine can and should chelate in these positions. A systematical study was made of all possibilities and will be published when the models have been checked carefully. The study of such chelation might be of some diagnostic value in the study of the tertiary structure, if it is remembered that chelation of straight chains might be masked because the chains are imbedded in the tertiary structure and that on the other hand chelation might be obtained

[10] *Lindqvist* - pag. 6

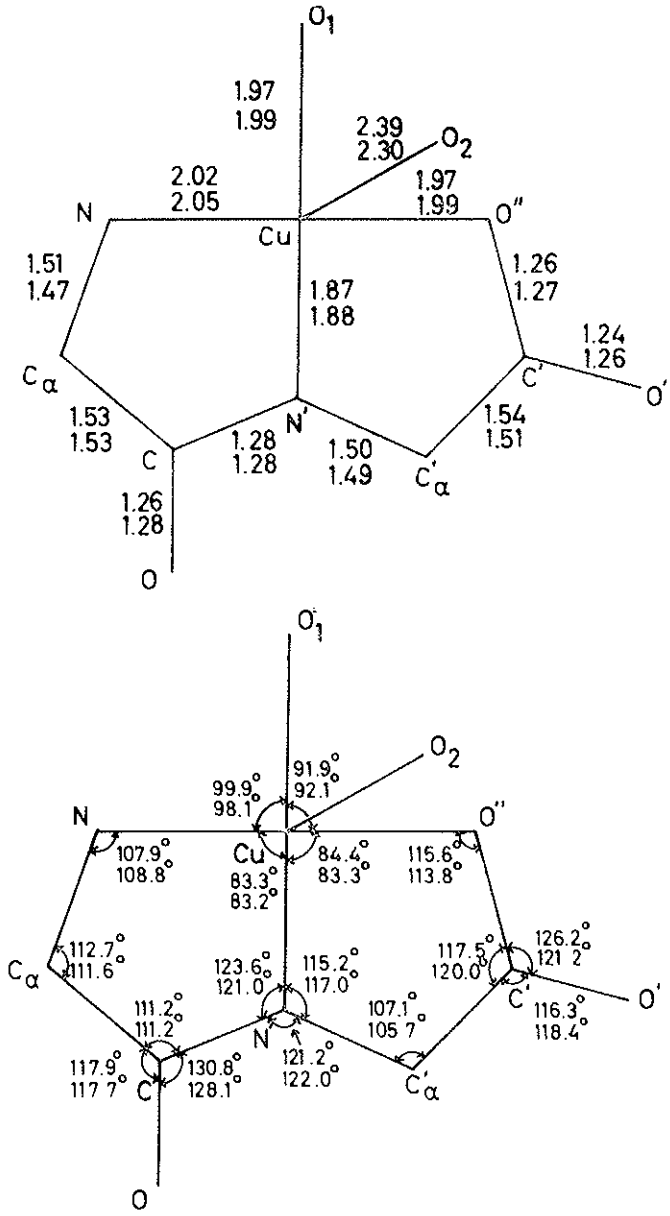


FIG. 3 *a* and *b* — The coordination around the copper (II) ions in copper (II) monoglycylglycine trihydrate. In *a* the bond distances and in *b* the bond angles are given for both the independent molecules.

from quite different parts of the protein. If the number of zinc ions which can be attached is greater than what can be predicted from the amino acid sequence such crosslinking chelation must be of importance. If on the other hand the number of zinc ions attached is much smaller than predicted there cannot be many straight helical chains available for chelation for one reason or another.

REFERENCES

[1] *Federation Proceeding Suppl.* no. *10*, page 60 (Malmström, B. 1961), page 71 (VALLEE B.L., 1961).

[2] COLEMAN J.E. and VALLEE B.L., *J. Biol. Chem.*, *236*, 2244 (1961). VALLEE B.L., WILLIAMS R.J.P. and COLEMAN J.E., *Nature*, *190*, 633 (1961).

[3] STRANDBERG B., LINDQVIST I. and ROSENSTEIN R., *Zeitschrift für Kristallographie*, *116*, 265-289 (1961).

DISCUSSION

PERUTZ

I am interested in Prof. LINDQVIST's copper coordination compound because copper is one of the few ions which inhibit crystallization of hemoglobin in very small amounts: 2 equivalents of copper per molecule of hemoglobin inhibits crystallization, and according to his results, copper could chelate at many different points and thereby produce random distortions of the hemoglobin molecule.

KATCHALSKI

May I ask whether the interesting investigations of Prof. LINDQVIST on the structure of zinc-imidazole and zinc-thioethanol-amine have led to any hypothesis or suggestions as to the role of zinc in determining the enzymatic activity of carboxypeptidase?

Poly-*L*-histidine has been prepared in our laboratory some time ago (1). It readily combines with zinc and other cations. In collaboration with Dr. BLOUT, it has been shown recently that poly-*L*-histidine attains three different macromolecular conformations in the suitable solvent systems. Would it, therefore, be desirable to study the structure of the zinc derivatives of poly-*L*-histidine, which have been obtained under conditions favoring a helical or a random coil conformation?

---

(1) A. PATCHORNIK, A. BERGER and E. KATCHALSKI, *J. Am. Chem. Soc.*, *79*, 5227, 1957.

LINDQVIST

I think it is too early to propose even hypotheses about the role of zinc. It would be very interesting to study metal complexes of poly-*L*-histidine.

SCHRAMM

I don't know whether this copper reaction has to do with the biuret reaction?

LINDQVIST

The biuret reaction requires that more strong Cu-N bonds are formed.

STUDIES ON THE SYNTHETIC FUNCTION OF THE RIBOSOME-AMINOACYL-sRNA SYSTEM

FRITZ LIPMANN, DANIEL NATHANS,
and GÜNTER VON EHRENSTEIN

The Rockfeller Institute, New York, N.Y. (U.S.A.)

The problem of replication of the genetic material in deoxyribonucleic acid (DNA) has been resolved into a coupling between polymerization [1] and base pairing [2] through hydrogen bonding to the mother DNA. An entirely analogous mechanism has now been found for the replication of DNA sequences in ribonucleic acid (RNA) [3, 4], and a continuity of information transfer through sequence replication by base pairing extends from DNA to RNA, which means to the template for protein synthesis. Therewith, the problem of protein synthesis emerges even more clearly as representing a translation of a message contained in the 4-symbol language of the nucleic acids into the 20-symbol language of polypeptide sequence. The molecules involved in this transformation process are presented in Fig. 1. On the left side are the four bases of RNA, leaving out the unusual ones. On top, the two purines; below, two pyrimidines, adenine pairing with uracil and guanine with cytidine. On the right side we see the array of amino acids characterized by the monotonous presence of the carboxyl α-amino configuration with the enormous variety of chemistry in the side chains.

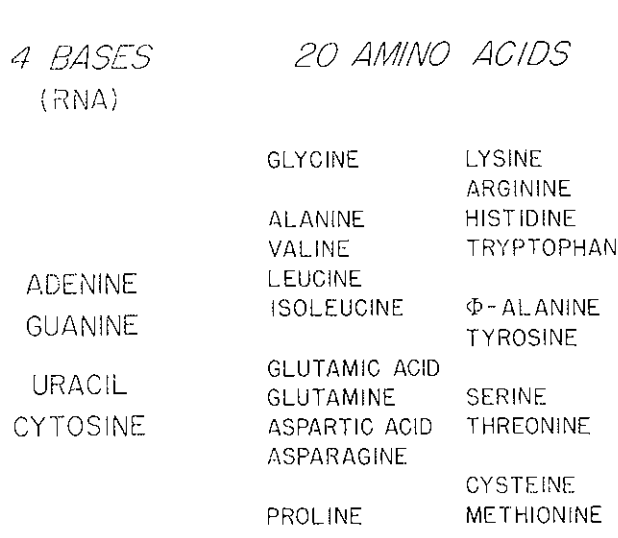
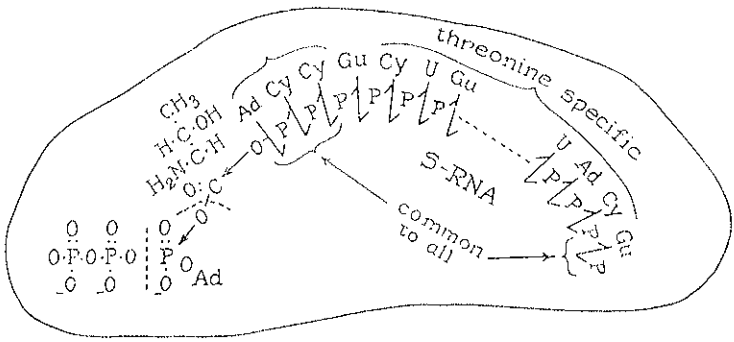


FIG. 1

An important advance came when it appeared that amino acid-specific activation enzymes catalyze not only the amino acid-ATP reaction, but also transfer the activated carboxyl to the corresponding soluble ribonucleic acid (sRNA) [5, 6]. The process is pictured in Fig. 2, which describes, as an example, the activation of threonine which we studied more extensively with HARTMANN [7]. The activation by displacement of the pyrophosphate in ATP to form the adenyl derivative of the amino acid is followed by transfer of the activated carboxyl to the adenosine terminal of the sRNA. Fig. 3 shows this interesting bonding in greater detail. The link forms between the carboxyl group and the 2′ or 3′ hydroxyl group of the ribose of terminal adenosine [8]. The energy of this ester bond is rather high, and is equivalent to the phosphoryl bond in ATP. The over-all equilibrium constant is near 1. The binding of amino acid to its specific sRNA represents the translation of amino acid into polynucleotide coding; the positioning of

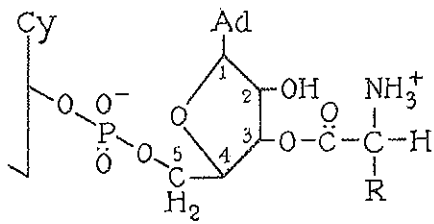
Threonine Activation Enzyme



ATP + Threonine + RNA ⇌ Threonine~RNA + AMP + PP

FIG. 2

sRNA Amino Acid Ester



or

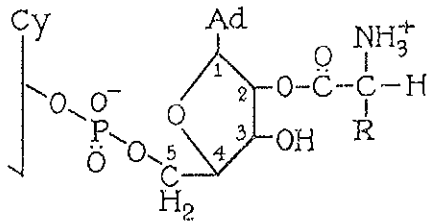


FIG. 3

the amino acid in the prescribed sequence can be imagined to occur by a pairing of an amino acid-specific sequence of 3 to 4 bases through hydrogen bonding between sRNA and template RNA, as proposed by CRICK [9]. A rather primitive picture of this process is attempted in Fig. 4.

The product formed in the activation-transfer step, the aminoacyl-sRNA, seems to be freely diffusable in the cytoplasm. The mixture of 20 amino acids, linked to their specific RNA's, can be obtained by reacting a mixture of amino acids containing a radioactive one as marker, with a mixture of activating enzymes, ATP, and sRNA, and then isolating by phenol extraction [10]. The activating enzymes are present in the 100,000 × g supernatant fraction of cell homogenates.

We have concentrated on the study of protein synthesis using aminoacyl-sRNA's as donors of active amino acid for peptide bond formation on the ribosome. That we are really studying here a process of protein synthesis was proven when this system was shown to produce hemoglobin [11]. The

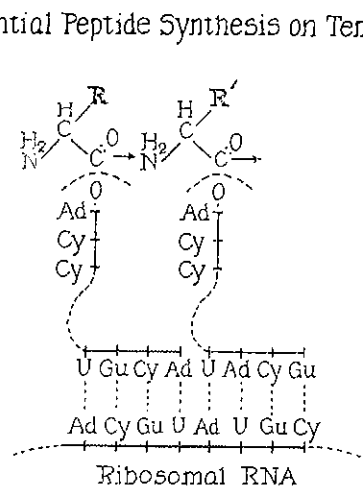


FIG. 4

[11] *Lipmann* - pag. 4

mixture of the aminoacyl-sRNA's had appeared to be unspecific with regard to species or to the protein formed [12]. To convince ourselves more thoroughly of this remarkable non-specificity, meaning a universality of code, we used in our hemoglobin synthesis the aminoacyl-sRNA mixture obtained from the activating enzymes and the sRNA's of *E. coli*. The ribosomes were prepared from reticulocytes according to SCHWEET [13]. In addition, a supernatant fraction, GTP, and free sulfhydryl groups are essential. This minimal system is outlined in Table 1.

TABLE I

Components of Protein Synthesis System

Aminoacyl-sRNA
Ribosome, Mg^{++}
Peptide polymerase, in supernatant factor
Other enzymes in supernatant factor?
GTP
Sulfhydryl groups

Using this system, the newly formed hemoglobin was collected and, after addition of some carrier hemoglobin, the mixture was fractionated on an carboxymethyl cellulose column. As shown in Fig. 5, radioactivity and color analysis give nicely matching peaks, indicating that the radioactive material is hemoglobin. A further confirmation was obtained by the trypsin digest « fingerprinting » method of INGRAM which, in Fig. 6, shows an overlap between the radioactivity and the ninhydrin reaction, indicating clearly that the labelled

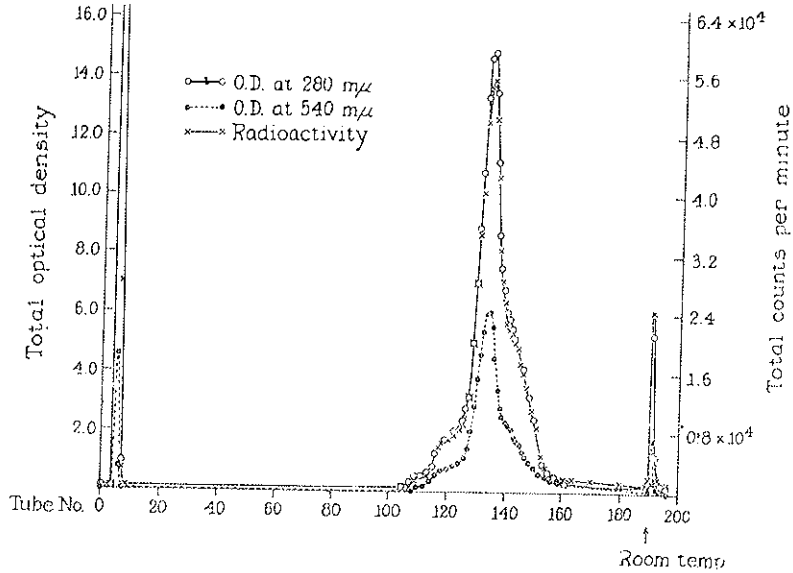

Total optical density
16.0
14.0
12.0
10.0
8.0
6.0
4.0
2.0
O.D. at 280 mμ
O.D. at 540 mμ
Radioactivity
6.4 ×10⁴
5.6
4.8
4.0
3.2
2.4
1.6
0·8 ×10⁴
Total counts per minute
Tube No. 0
20
40
60
80
100
120
140
160
180
200
t
Room temp

FIG. 5

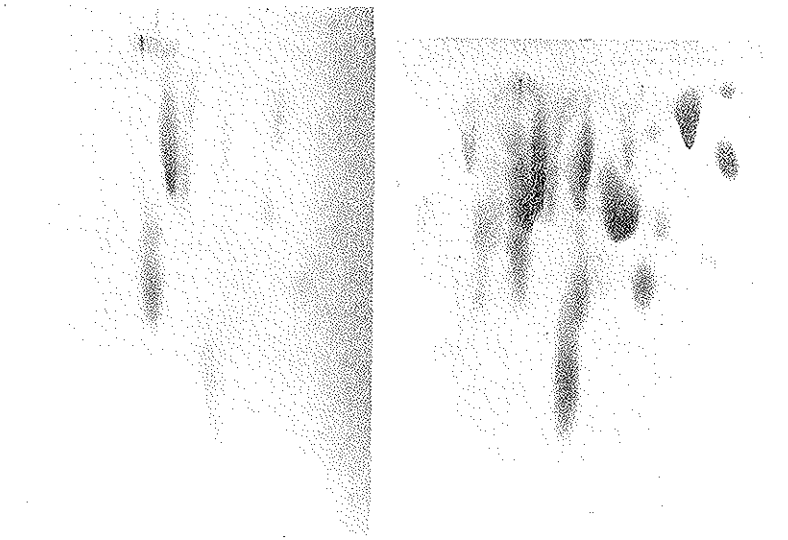


FIG. 6

amino acid, leucine, was incorporated into the peptides of hemoglobin.

Some progress toward a better understanding of the specific roles of the components of this system was obtained by a purification of the supernatant factor containing the peptide polymerase. For assay of the factor, the ribosomes had to be washed with a surface active reagent using 0.5% deoxycholate, which removes the strongly adsorbed transfer enzyme. Preparations of transfer factor have been obtained from liver [12], from reticulocytes [11], and from *E. coli* [10]. In view of our greater experience with the *E. coli* system, further studies were done with it. Fig. 7 shows an assay for the supernatant factor using washed ribosomes, and Fig. 8 indicates a peak obtained by chromatography on a DEAE-cellulose column. The tested amino acids all fall on to the same peak, and similar tests have been done with three more amino acids (Fig. 9). This indicates that we are dealing with a single factor that is identical for all amino acids.

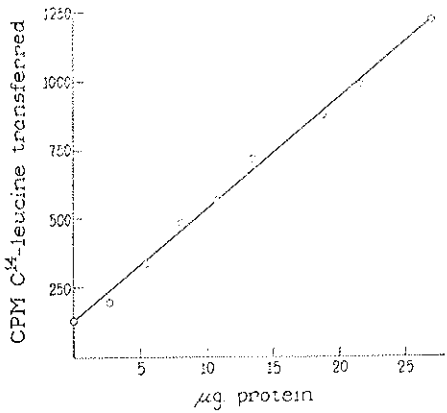


FIG. 7

[11] *Lipmann* - pag. 7

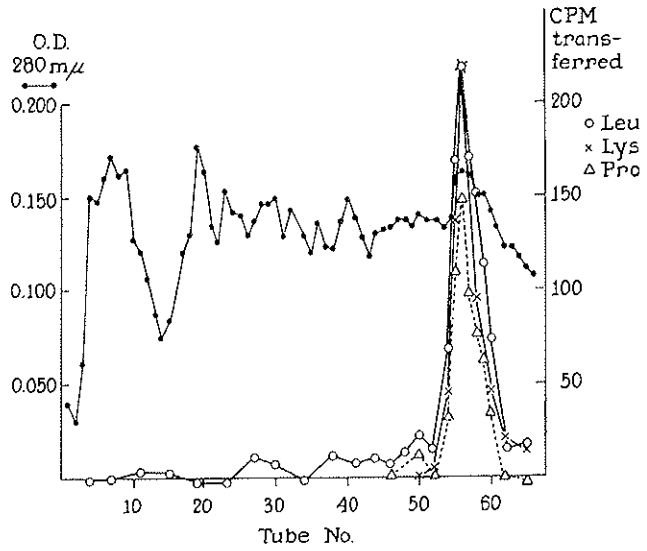


FIG. 8

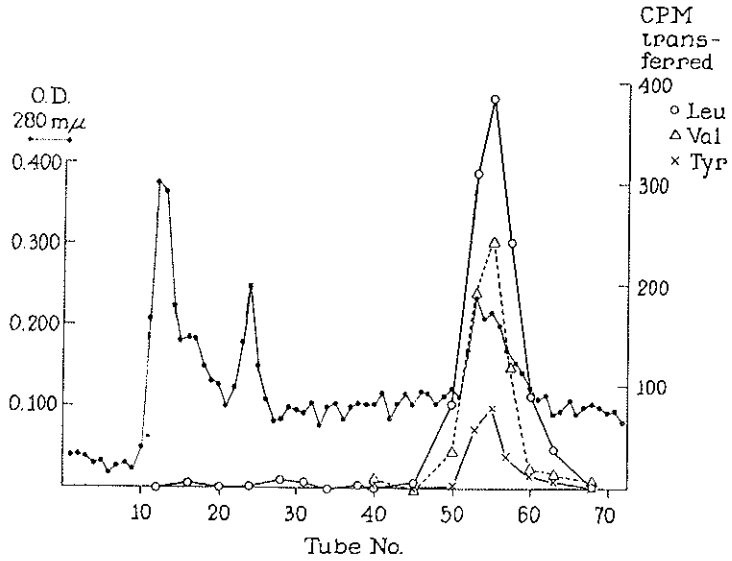


FIG. 9

[11] *Lipmann* - pag. 8

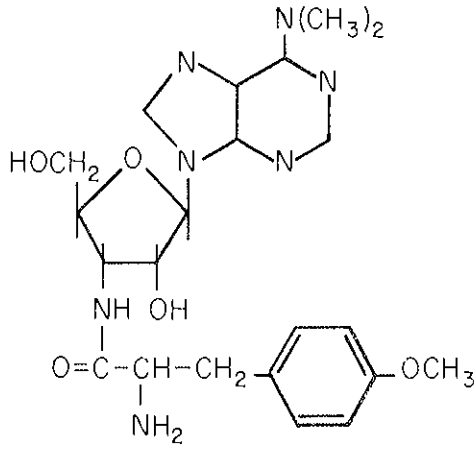
We have pointed out that the aminoacyl-sRNA appears to be non-selective with regard to the ribosome. Therefore, we were somewhat surprised to find that the supernatant factor is species specific. As shown in Table 2, the *E. coli* supernatant will not supplement liver particles, nor will liver supernatant supplement *coli* particles. In all cases, *E. coli* aminoacyl-sRNA was used as amino acid donor. This supernatant specificity seems interesting.

TABLE II

Ribosome Specificity of Transfer Factor

Ribosomes	Supernatant fraction	C^{14}-leu trans. c.p.m.
E. coli	None	90
E. coli	Liver	80
E. coli	E. coli	665
Liver	None	21
Liver	E. coli	19
Liver	Liver	244

Chloramphenicol and, from the work of YARMOLINSKY et al. [14], puromycin, have been shown to inhibit protein synthesis rather specifically. Puromycin, which in its structure (Fig. 10) has a great similarity to the amino acid-carrying adenosine terminal of sRNA (cf. Fig. 3), was proven to be a general inhibitor of the transfer from the aminoacyl-sRNA to the ribosome. In our hands, however, chloramphenicol has shown activity only with microbial systems. Table 3 shows that chloramphenicol at a high concentration does not inhibit the reticulocyte system, while puromycin inhibits both the reticulocyte and the microbial systems. The same is true for liver (unpublished experiments).



Puromycin

FIG. 10

TABLE III

Inhibition of Aminoacyl-sRNA—Ribosome System (*)

Ribosome	Inhibitor	C^{14} leucine inc. c.p.m.
Reticulocyte	None	6,110
Reticulocyte	10^{-3} M chloramphenicol	5,950
Reticulocyte	5.10^{-4} M puromycin	232
E. coli	None	1,725
E. coli	2.10^{-4} M chloramphenicol	577
E. coli	4.10^{-4} M puromycin	46

(*) E. coli aminoacyl-sRNA used in all.

In the course of studying the puromycin effect, it appeared that in the presence of puromycin, amino acid incorporation is replaced by an equivalent amount of amino acid fixation in a trichloroacetic acid and 60% alcohol soluble fraction. The puromycin-promoted fixation shown in Table 4 is dependent upon all the components that are necessary for protein synthesis and is essentially equivalent to the amino acid incorporation. Chloramphenicol does not have a comparable effect and, as seen in line 4, it inhibits the puromycin effect of fixation in a low molecular fraction. A better identification of this low-molecular fraction is now being attempted.

TABLE IV

TCA-Soluble Compounds- vs. Protein-Synthesis

Inhibitor	Per cent TCA-soluble	Per cent in protein
1. None	0, 0	35, 36
2. Puromycin 5×10^{-4} M . .	29, 31	1.2, 0.8
3. Chloramphenicol 5×10^{-4} M	0, 0	15, 15
4. 2+3	9, 12	1.8, 1.4

The recent discovery by NIRENBERG et al. [16] that poly-uridylic acid will act as template in an exclusive manner for the polymerization of phenylalanaine and poly-cytidylic acid for the polymerization of proline, has opened new avenues for studying the template problem. I will present some data with the NIRENBERG system, using phenylalanyl-sRNA with our transfer factor and washed *E. coli* ribosomes. Since the basic experiments of NIERENBERG et al. were done in a crude system using the supernatant, ATP, and phenylalanine, we were quite eager to check their observation with our minimal system using

phenylalanyl-sRNA, transfer factor, and ribosomes. The characteristics of this reaction using « active phenylalanine » as the precursor are shown in the next tables and figures. They bear out our presumption that this system can be considered a standard one for the study of the mechanism of protein synthesis.

In table 5 it can be seen that all factors required for the incorporation of leucine into protein are also required for the polymerization of phenylalanine on poly-uridylic acid. Figure 12 presents a time curve and shows a brief induction period which seems to be real and worth attention. We next figure, 12, brings out the fairly low ribosome concentration at which saturation is reached, about 1/6th to 1/5th of that generally used in the previous experiments. If phenylalanine is tested with increasing poly-uridylic concentration, an upper limit is reached at a ratio of uridylic acid to phenylalanine incorporation of 3.2, after which the curve plateaus (Fig. 13). A similar ratio obtains if poly-uridylic acid is kept constant and the amount of phenylalanyl-sRNA made available is increased (Fig. 14). Here again, a plateau develops when the ratio of uridylic acid versus phenylalanine incorporation is 3.8.

TABLE V

Cofactor Requirements for Polyuridylic acid-dependent Incorporation from Phenylalanyl sRNA

	Components	cpm	mμMoles
1)	Complete system	562	0.097
2)	- Poly U	28	0.0048
3)	- Supernatant factor	12	0.0021
4)	- Ribosomes	0	0
5)	- GTP, PEP and binase	14	0.0024
6)	+ C^{12} - Phenylalanine (1 μMole) . .	590	0.102

[11] *Lipmann* - pag. 12

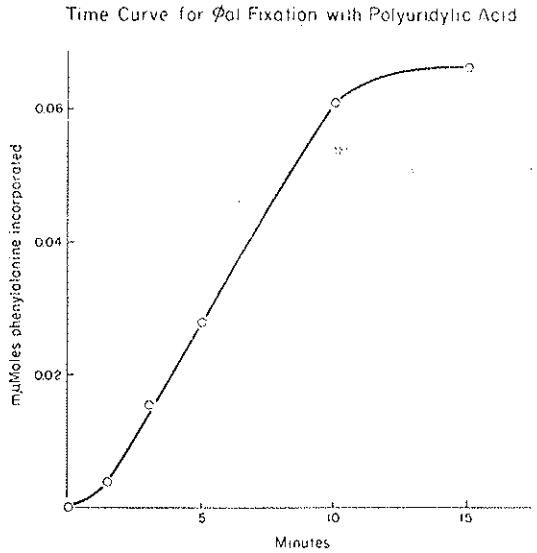

Time Curve for Φal Fixation with Polyuridylic Acid
mμMoles phenylalanine incorporated
0.06
0.04
0.02
5
10
15
Minutes

FIG. 11

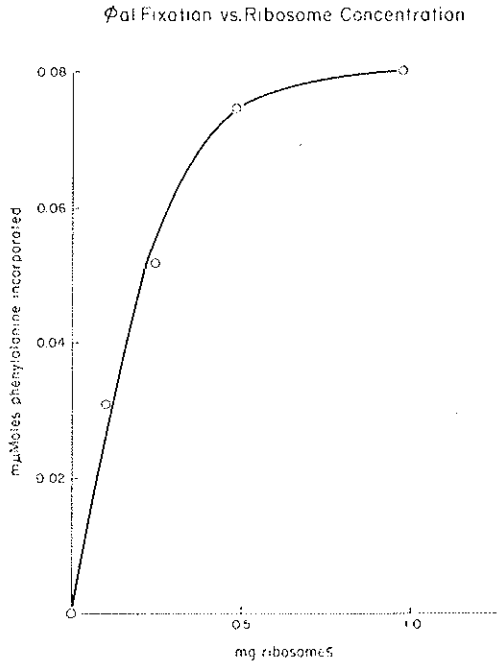

Φal Fixation vs. Ribosome Concentration
mμMoles phenylalanine incorporated
0.08
0.06
0.04
0.02
0.5
1.0
mg ribosomes

FIG. 12

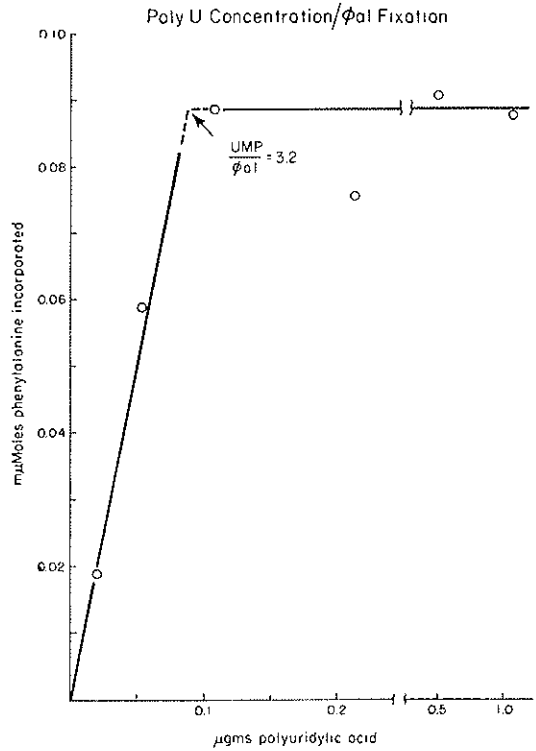


FIG. 13

These observations were to be in accord with the assumption that this template cannot be re-used, and an upper limit value of 3 could correspond with a 3-nucleotide symbol for the amino acid. Dr. NIRENBERG informs us, however, that in his experiments with the cruder system, after first finding a ratio of about 3, their values recently have tended to go down to almost 1, which does not fit in with a non-catalytic function of the template since a code of less than 3 is difficult to imagine. This would be in better accord with our experience with the reticulocyte system which does indicate a stable template. In

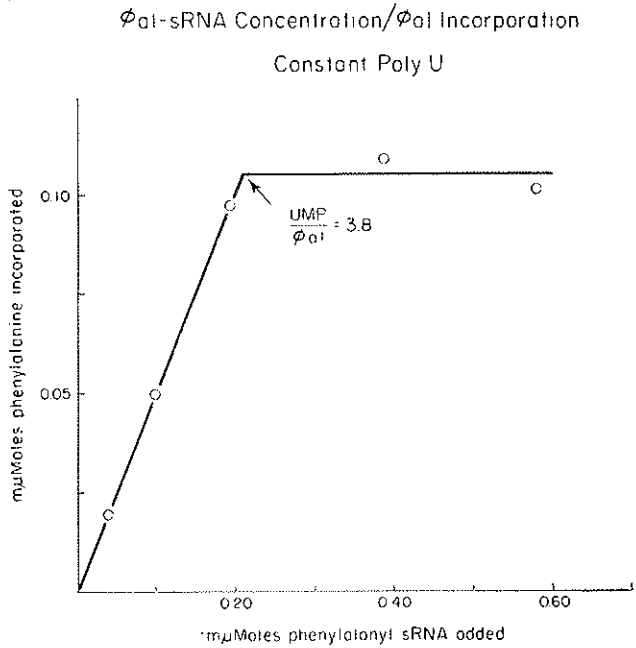


FIG. 14

preliminary results, it was found that the living reticulocyte, using radioactive phosphate as a tracer, shows only a very slow incorporation as compared with the incorporation of amino acids (Table 6). The incorporation of phosphate, however, cannot be equated without reservation with the turnover of the RNA without also assessing phosphate incorporation into the nucleotide pool which may be rather slow in the reticulocyte. Further experiments on this problem are now in progress.

Nevertheless, the present results, together with SCHWEET's recent work [16] encourage the assumption that in the reticulocyte, and possibly more generally in the mammalian cell, the template RNA may be rather stable in contrast to the high turnover of the template or messenger RNA in the microbial

TABLE VI

Incorporation of Phosphate into RNA Fractions and of Amino Acids into Hemoglobin by Rabbit Reticulocytes

Incorporation	μμM per 10 ml of packed cells per minute
Amino acids	990,000
Soluble RNA	52
Ribosomal RNA	
1 M NaCl precipitate	4.7
1 M NaCl supernatant . . .	4.5
Debris RNA	
1 M NaCl precipitate	27
1 M NaCl supernatant	30

cell. It is best, therefore, to postpone a definite decision on the rather crucial question as to whether a breakdown of the template RNA is coupled with the synthesis of the polypeptide chain of a protein, or if the linking of the amino acids takes place on a stable polynucleotide chain which may serve many times over.

No discussion of our problem can be complete without touching upon the undoubted role of guanosine triphosphate (GTP). Studies are complicated by contamination of the system by « GTPases » of various kinds. We are attempting, with some success, to clean up the system, but so far not enough to yield any reasonable chemical correlation. Present results, obtained by R. MONRO indicate only that it is the terminal phosphoryl of GTP that reacts.

[11] *Lipmann* - pag. 16

REFERENCES

[1] KORNBERG A., *Science, 131,* 1503 (1960).

[2] WATSON J.D. and CRYCK F. H. C., *Nature, 171,* 737 (1953).

[3] WEISS S.B. and NAKAMOTO T., *Proc. Nat. Acad. Sci., U.S., 47,* 694 (1961).

[4] HURWITZ J.A., BRESLER A. and DIRINGER R., *Biochem. and Biophys. Research Communications, 3,* 15 (1960).

[5] HOAGLAND M.B., ZAMECNIK P.C. and STEPHENSON M.L., *Biochim. et Biophys. Acta, 24,* 215 (1957).

[6] BERG P., *J. Biol. Chem., 233,* 601 (1958).

[7] LIPMANN F., HÜLSMANN W.C., HARTMANN G., BOMAN H.G. and ACS G., *J. Cell. Comp. Physiol., 54,* Sup. 1, 75 (1959).

[8] ZACHAU H.G., ACS G. and LIPMANN F., *Proc. Nat. Acad. Sci. U.S., 44,* 885 (1958)

[9] CRICK F.H.C., *Symp. Soc. Exp. Biol., 12,* 138 (1958).

[10] NATHANS D. and LIPMANN F., *Proc. Nat. Acad. Sci. U.S., 47,* 497 (1961).

[11] VON EHRENSTEIN G. and LIPMANN F., *Proc. Nat. Acad. Sci. U.S., 47,* 941 (1961).

[12] NATHANS D. and LIPMANN F., *Biochim. et Biophys. Acta, 43,* 126 (1960).

[13] SCHWEET R., LAMFROM H., and ALLEN E., *Proc. Nat. Acad. Sci. U.S., 44,* 1029 (1958).

[14] YARMOLINSKY M.B. and DE LA HABA G.L., *Proc. Nat. Acad. Sci. U.S., 45,* 1721 (1959).

[15] NIRENBERG M.W. and MATTHAEI J.H., *Proc. Nat. Acad. Sci. U.S., 47,* 1580, 1588 (1961).

[16] BISHOP J. and SCHWEET R., *Nature, 191,* 1365 (1961).

DISCUSSION

ANFINSEN

As far as understood about two weeks ago, polyphenylalanine was easily detected because of its insolubility. This was the macromolecular substance which, upon hydrolysis, gave only phenylalanine. There was no clear idea of the size of the polymer. There is one important point that might be kept in mind, i.e. that polyaminoacids composed of all *L*, or all *D* nonpolar amino acids have a tremendous lyophobicity. Such lyophobicity would constitute a strong driving force due to precipitation and might help drive a nonspecific reaction.

LIPMANN

Against such unspecific precipitation of an insoluble product speaks the need for all the components of the system known to be essential as shown in one of my tables. In its requirements for activation of amino acid, supernatant factors, and GTP, polyphenylalanine synthesis corresponds exactly to the systems proven to synthesize specific protein.

FRAENKEL-CONRAT

In cooperation with Dr. NIRENBERG, TSUGITA and I have been attempting to establish the nature of the protein synthesized by the *E. coli* system under the direction of TMV-RNA. It appears that a good part of the labeled material is related to TMV coat protein,

but not quite identical with it, as judged both from reconstitution experiments and from protein structural work, i.e., isolation of labeled peptides after trypsin digestion of the protein.

LIPMANN

I understand that in the *in vitro* synthesized TMV protein the terminal acetyl group is missing, as one might expect. For this reaction, a particular enzyme system should be required. This alone might explain why this protein cannot combine with RNA to form a good virus. For obvious reasons, I have been rather curious about the nature and production of this acetylation enzyme.

KATCHALSKI

Has any study been made on the chain length or physical chemical properties of the polyphenylalanine obtained in NIERENBERG's system?

LIPMANN

I think NIRENBERG has made attempts towards determination which seem to be difficult because of the difficulty to find a solvent for the phenylalanine polymer.

[11] *Lipmann* - pag. 20

INTERACTION BETWEEN DNA AND POLYCYCLIC AROMATIC HYDROCARBONS

A.M. LIQUORI

Centro Nazionale di chimica delle Macromolecole
Istituto Chimico dell'Università, Napoli (Italia)

F. ASCOLI, C. BOTRÉ

Istituto di Chimica Farmaceutica dell'Università, Roma (Italia)

The ability of several purines to solubilize polynuclear aromatic hydrocarbons which are virtually insoluble in water has been reported long ago (WEIL-MALHERBE, 1946; BROCK, DRUCKREY and HAMPERL, 1938).

Crystalline molecular complexes between purines and aromatic hydrocarbons have also been reported to form from organic solvents (BOOTH and BOYLAND, 1953; BOOTH, BOYLAND and ORR, 1954).

Mixed crystals containing a very small number of aromatic hydrocarbons in the lattice of the purine have also been studied (BOOTH and al., 1954).

GEOMETRY OF PURINE-HYDROCARBON COMPLEXES

An X-ray investigation of the crystal structure of 1:1 and 1:2 molecular complexes between a purine, namely 1-3-7-9 tetramethyluric acid (TMU) and aromatic hydrocarbons such as pyrene, benzpyrene and coronene was undertaken a few

years ago in our laboratory with the aim to obtain information on their geometry which could in turn give some indication on the kind of forces promoting both complex formation and solubilization of aromatic hydrocarbons in purine water solutions.

The crystal structure of TMU was studied first. Centrosymmetrical pairs of TMU molecules lying in parallel planes with a perpendicular separation of 3.4 A were found in one crystalline modification by DE SANTIS, GIGLIO and LIQUORI (1960) ($a = 15.59$ A; $b = 8.87$ A; $c = 8.46$ A; $\beta = 118°$; $P2_1/c$; $z = 4$) (see fig. 1). The same type of structure, although with a different mode of packing was independently established by SUTOR (1960) for a second crystalline modification.

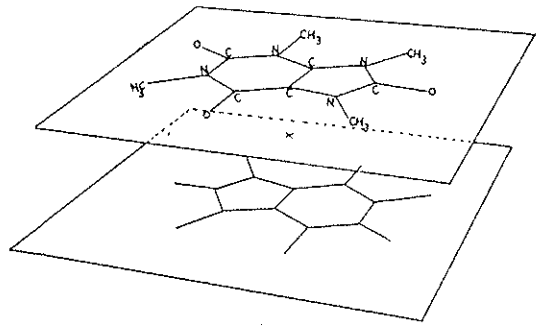


FIG. 1* — Arrangement of centrosymmetrical pair of molecules in the lattice of TMU.

Next the crystal structure of the 1:1 complex between TMU and pyrene was investigated ($a = 9.71$ A; $b = 8.00$ A; $c = 15.04$; $\beta = 117°$; $z = 2 + 2$; $P/c - C^2_s$) (DE SANTIS, GIGLIO, LIQUORI and RIPAMONTI (1961).

(*) Reprinted from *Nature, 188,* 46 (1960).

FOURIER projections on two planes of the unit cell revealed the presence of alternate stacks of pyrene and TMU molecules lying in parallel planes with a perpendicular separation of 3.4 A (see fig. 2).

A graphite-like displacement from coincidence of the perpendicular projections of the atoms in one molecule with respect to the atoms in the next molecule can be observed. This mode of packing is typical of molecular complexes the stability of which is commonly ascribed to « polarization bonding » between the components (WALLWORK, 1961), the complexes between aromatic hydrocarbons and s-trinitrobenzene being a well known example. Weak charge transfer and electrostatic iuteractions between the polar purine and the polarizable aromatic molecule must therefore make the dominant contribution to the stability of these complexes.

The crystal structure of the 2:1 complexes between bezpyrene and TMU and between coronene and TMU is still under study. The results thus far obtained have already shown that a plane to plane stacking is present also in these complexes.

In these cases however a sandwich-like arrangement of the aromatic molecule between the two purine molecules rather than an infinite stacking characterize the geometry of the molecular complexes.

It is interesting to notice that the TMU molecule retain in the 2:1 complexes the same arrangement they have in their own crystal lattice.

These studies on the crystal structure of molecular complexes between purines and aromatic molecules are still in progress but the results thus far obtained have provided a very convincing evidence that perpendicular interactions can be established between aromatic molecules and purines, which may be ascribed to polarization bonding. Furthermore it may be inferred that the same forces promote the solubilization ability of purine water solutions toward aromatic hydrocarbons observ-

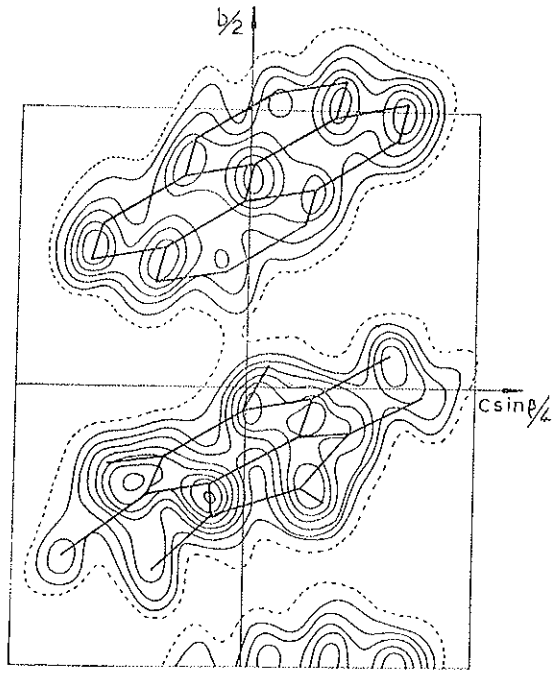


FIG. 2-*a* (*) — Fourier projection of the electron density on (100).

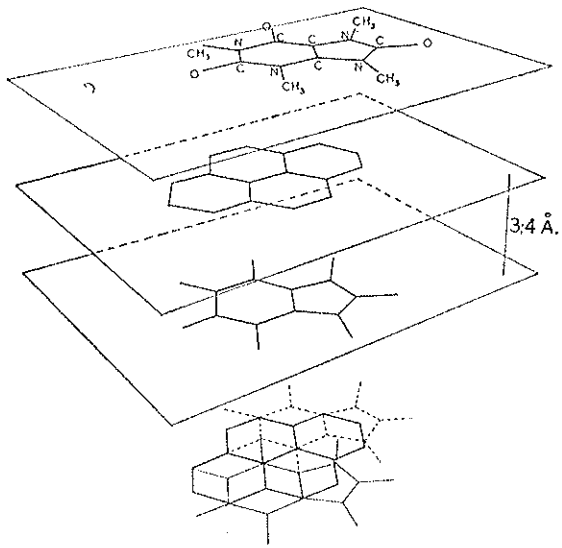


FIG. 2-*b* (*) — Schematic drawing of the perpendicular separation of the molecules in one stack through the unit cell.

(*) Reprinted from *Nature*, *191*, 900 (1961).

ed by WEIL-MALHERBE, by formation of aggregates essentially similar to those present in the crystalline complexes.

Considering the arrangement of the purine and pyrimidine bases within the core of DNA double stranded helix we were induced to think of the possibility, that vertical interactions of the above described type might confer to this molecule the ability to solubilize large aromatic hydrocarbons by occluding them in its core, provided that some degree of disorder could be introduced into CRICK and WATSON structure which can be regarded under many respects as an anphiphillic system.

An investigation was therefore carried out with the collaboration of Dr. ASCOLI, Dr. BOTRÉ and Dr. TRASCIATTI with the aim to exploit this possibility using two large aromatic hydrocarbons, benzpyrene (BP) and dibenzanthracene (DBA).

The number of solubilization sites available for the hydrocarbon molecules was determined in various conditions and U.V. absorption and fluorescence spectra of the solubilized hydrocarbon were recorded. Some preliminary microspectrographic observations have also been made by Prof. DE LERMA in order to investigate the possibility of interactions between benzpyrene and the nuclear components of living cells, which will be reported elsewhere.

SOLUBILIZATION OF AROMATIC HYDROCARBONS IN DNA WATER SOLUTIONS

The solubilization of BP and DBA in solutions of calf thymus DNA prepared according to KAY and al. (1952) was determined at 20° C., *a*) at different DNA concentrations in absence of salt and *b*) at a single DNA concentration in presence of 10^{-2} M NaCl following exposure of the solutions to different temperatures, between 40° C. and. 95° C.

In order to determine the maximum amount of the hydrocarbon which could be solubilized, the solutions were equilibrated

with very fine crystals of the hydrocarbon for one day or more, in black glass tubes.

After careful filtration across a sintered glass, known volumes were extracted with known volumes of cyclohexane or benzene and the hydrocarbon content in the organic phase was analyzed spectrophotometrically.

The number of hydrocarbon molecules solubilized per monomer unit is plotted in fig. 3 against DNA concentration for BP (curve *a*) and DBA (curve *b*). Both curves show a remarkable trend.

The number of hydrocarbon molecules solubilized per DNA monomer unit remains practically invariant with DNA concentration until a range of dilution is reached which marks the onset of a very steep rise.

In a previous investigation (ASCOLI, BOTRÉ, CRESCENZI, LIQUORI and MELE, 1959a) on the influence of counterion binding on the conformational stability of DNA, the existence of a critical narrow range of concentration has been pointed out corresponding to the « dilution-denaturation » of DNA first observed by CAVALIERI, ROSOFF and ROSENBERG (1956). Comparison with the curves of fig. 3 shows that the steep increase of the solubilization power of DNA is observed at the onset of dilution denaturation.

These results therefore indicate that a small number of solubilization sites are available in the DNA molecule which are independent from the number of DNA molecules present in solution. At the critical dilution, where denaturation of the structure takes place, many more sites are rapidly created.

This interpretation is strongly supported by the results of the solubilization experiments carried out on the beat treated DNA solutions in presence of 10^{-2} M NaCl, which show that the number of BP molecules solubilized per DNA monomer unit depends critically on the temperature of the heat treatment. Also in this case, it may in fact be inferred that a large number of new solubilization sites are created in the DNA molecule in

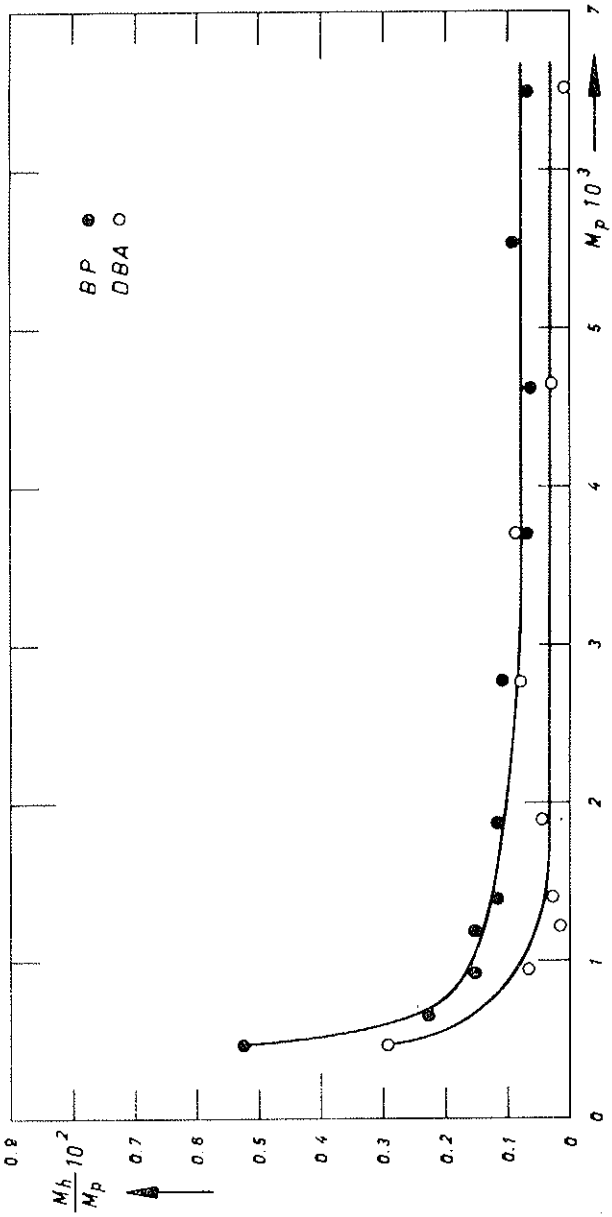


FIG. 3 — Number of moles of benzpyrene (BP) and dibenzanthracene (DBA) solubilized per DNA monomer unit as a function of DNA concentration.

a very narrow temperature range, where the well known order-disorder transition or thermal denaturation of DNA occurs (DOTY et al., 1959).

In fact, fig. 4 shows that the critical changes of optical rotation, optical density at 260 $m\mu$ (DOTY et al., 1959) and activity of Na^+ ions (ASCOLI et al., 1961) which DNA undergoes on thermal denaturation, occur in the same narrow temperature range where the increase of solubilization sites available for the aromatic hydrocarbon is observed.

U.V. ABSORPTION AND FLUORESCENCE SPECTRA

U.V. absorption spectra of the optically clear solutions of DNA containing solubilized BP were recorded by means of a DU BECKMAN spectrophotometer in a region where the absorption of DNA does not interfere with the characteristic spectra of the hydrocarbon. For comparative purpose, BP was solubilized in several purine and pyrimidine water solutions and in a sodium laurylsulphate solution.

The position of the characteristic absorption peaks in the near ultraviolet spectra of BP dissolved in a DNA solution (7.9×10^{-3} Mp) in cyclohexane, in water solutions of caffeine and sodium laurylsulphate are reported in Table I.

TABLE I

benzpyrene in	λ max mμ	λ max mμ	λ max mμ
Cyclohexane	347	363	383
Sodium laurylsulphate . . .	348	365	386
Caffeine water solution . . .	353	370	389
DNA water solution . . .	354	373	394

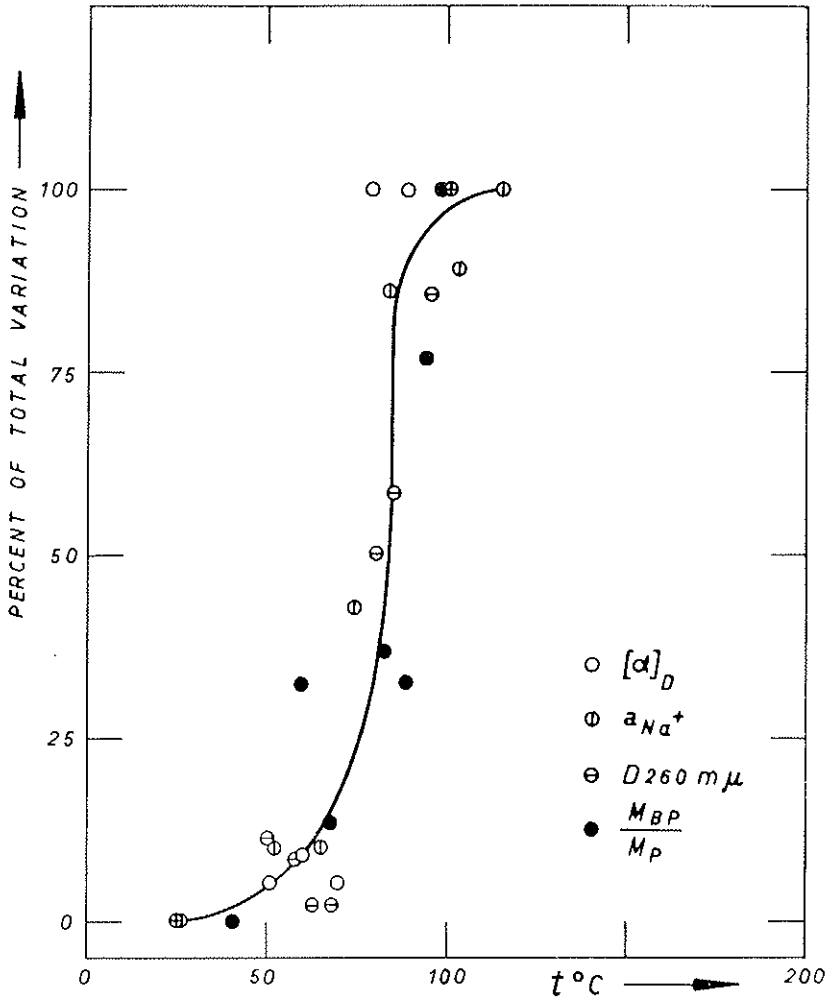


FIG. 4 — Percent of total variation in specific rotation, activity of Na^+ ions, optical density at 260 $m\mu$ and number of moles of BP dissolved per DNA monomer unit, as a function of previous thermal treatment of DNA solutions.

As it may be seen, the positions of the maxima in the near ultraviolet spectra of BP dissolved in caffeine already reported by BOOTH et al. (1953, 1954) and in DNA solution are both slightly red shifted with respect to the spectra in cyclohexane or in water containing sodium laurylsulphate.

Though the effect is not very large, it may be taken as an evidence that caffeine and DNA exert a similar perturbation on the electronic transitions of the hydrocarbon molecule.

The aromatic hydrocarbon occluded within the core of the micelles formed by the sodium laurylsulphate may be considered instead as dissolved in an inert solvent like cyclohexane (MCBAIN and HUTCHINSON, 1959; KLEVENS, 1948 and 1950).

Even more evident are the environmental influences on the fluorescence spectrum of BP. Fig. 5 shows the emission spectra

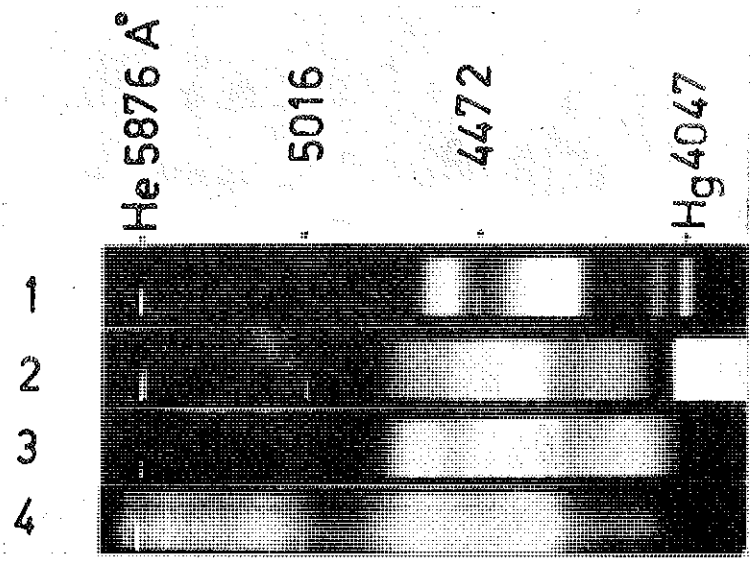


FIG. 5 — Fluorescence spectra of BP in different environments:
1) BP in cyclohexane solution
2) BP in DNA solution
3) BP in caffeine solution
4) BP in caffeine mixed crystals.

of BP in a cyclohexane solution (1), in DNA water solution (2), in caffeine water solution (3), in mixed caffeine crystals (4) prepared according to WEIL-MALHERBE (1946).

The spectra were recorded both with a ZEISS and a HILGER constant deviation emission spectrograph with a glass prism. The emission was excited with light of wave length 3650 A (not strictly monocromatic) from a mercury lamp. A fluorescence microscope was connected to the spectrograph when the mixed caffeine crystals were examined, according to the technique described by DE LERMA (1958).

Like the U.V. absorption spectra in the spectral region of 3889 A and 5016 A, the fluorescence bands of BP in DNA and in caffeine water solutions, as well as those in caffeine mixed crystals are red shifted and considerably broadened with respect to the spectrum in cyclohexane. It is of interest that similar broadening effects in the emission spectra of typical charge transfer complexes between aromatic hydrocarbons and trinitrobenzene have been observed (WALLWORK, 1961; REID, 1952).

The results very strongly suggest that the solubilization of the aromatic hydrocarbons in DNA solution is due to interactions other than VAN DER WAALS with the purinic and pyrimidinic bases. It seems very plausible to ascribe such interactions to weak charge transfer complexes formation and to dipole induced dipole interactions. However, a more quantitative analysis of the spectra will be necessary in order to obtain a more detailed information.

The above described experiments show that large aromatic hydrocarbons such as BP and DBA are solubilized in DNA water solutions to an extent depending in a very critical way on the degree of order of the DNA structure. In fact, under

conditions which do not alter the stability of CRICK and WATSON structure, a limited number of solubilization sites are available for the aromatic hydrocarbon. When denaturation of the structure of DNA is induced either by thermal treatment or by simple dilution at room temperature in salt free solution, a rather large number of new solubilization sites are created. In fact M_{BP}/M_P increase from 0.4×10^{-2} at 41° C. to 1.4×10^{-2} at 95^0 C.

As to the specific nature of these sites, the previously reported solubilization ability of several purines corroborated by our own experiments on the solubilization ability of some pyrimidines (not reported), strongly suggest that the bases contained in DNA confer to this molecules its solubilization power toward aromatic hydrocarbons.

Spectroscopic evidence indicates that the interactions between the bases of DNA and the hydrocarbon molecules, although rather weak, must imply the same kind of polarization bonding which would account for complex formation between purines and aromatic hydrocarbons.

Our previous studies on the changes of the charge density of DNA brought about by dilution (ASCOLI et al., 1959a) or thermal denaturation (ASCOLI, BOTRÉ and LIQUORI, 1961) of DNA have pointed out the possible coexistence in the disordered structnre of denatured DNA of sections with CRICK and WATSON helical conformation and sections with less ordered and much more extended conformations, whose population increases upon heating or on dilution below a critical range.

The average repeat of the extended conformations would correspond from a gross estimate based on polyelectrolyte theories (OOSAWA, 1957; ASCOLI et al., 1959b) to about twice the characteristic repeat of CRICK and WATSON helix.

All the above evidences suggest that the solubilization sites in DNA molecules might be identified with empty spaces between neighbouring paired bases in the stretched conformations.

It should be mentioned that recently a similar mechanism has been proposed to explain the interaction between DNA and acridine orange or other related cationic dyes (LERMAN, 1961). It seems plansible to reconcile such a proposal with the commonly accepted view that electrostatic interactions play a very important role in these cases (PEACOCKE and SKERRET, 1956; BRADLEY and FELSENFELD, 1959) by assuming that the two mechanisms, intercalation and electrostatic binding, are competing. Such competition is ruled out for the aromatic polycyclic hydrocarbons chosen for our studies because of their virtual insolubility in water and the lack of polar groups. However, it might be reasonable to attempt to correlate the fraction of acridine molecules intercalated in DNA from the results obtained with the aromatic hydrocarbons.

As already mentioned the work described has been carried out also in collaboration with Dr. TRASCIATTI and Prof. DE LERMA and it is a pleasant duty to acknowledge them.

REFERENCES

Ascoli F., Botré C., Crescenzi V., Liquori A.M. and Mele *A.*, *Nature, 184,* 1482 (1959a).

— *J. Polymer Sci., 40,* 169 (1959b).

Ascoli F., Botré C., Liquori A.M., *J. Mol. Biol., 3,* 202 (1961).

Booth J. and Boyland E., *Biochimica and Biophisica Acta, 12,* 75 (1953).

Booth J., Boyland E. and Orr S.F.D., *J. Chem. Soc.,* 598 (1954).

Bradley F. and Felsenfeld G., *Nature, 184,* 1920 (1959).

Brock N., Druckrey H. and Hamperl H., *Arch. Exp. Path. Pharmac., 189,* 709 (1938).

Cavalieri L.F., Rosoff M. and Rosenberg B.H., *J. Am. Chem. Soc., 78,* 5239 (1956).

De Lerma B., *Handbuch der Histochemie,* Vol. I (1958). Gustav Fischer Verlag, Stuttgart.

De Santis P., Giglio E. and Liquori A.M., *Nature, 188,* 46 (1960).

De Santis P., Giglio E., Liquori A.M. and Ripamonti A., *Nature, 191,* 900 (1961).

— unpublished.

Doty P., Boedtker H., Fresco J.R., Haselkorn R., and Litt M., *Proc. Nat. Acad. Sci., 45,* 482 (1959).

Kay E.R.M., Simons *N.S.* and Dounce A.L., *J. Am. Chem. Soc., 74,* 1724 (1952).

Klevens H.B., *J. Phys. Chem., 52,* 130 (1948); ibidem, *54,* 1012 (1950).

Lerman L.S., *J. Mol. Biol., 3,* 18 (1961).

McBain and Hutchinson E., *Solubilization and Related Phenomena,* Academic Press, New York (1959).

Oosawa F., *J. Polymer Sci., 23,* 421 (1957).

Peacocke A.R. and Skerrett J.N.H., *Trans. Far. Soc., 52,* 261 (1956).

Reid C., *J. Chem. Phys., 20,* 1212 (1952).

— *Excited States in Chemistry and Biology,* Academic Press Inc., New York (1957).

Sutor J., *Nature, 188,* 48 (1960).

Wallwork S.C., *J. Chem. Soc.,* 494 (1961).

Weil-Malherbe H., *Biochemical Journal, 40,* 351 (1946).

DISCUSSION

RICH

Is the DNA employed in your experiments denatured or native, and what happens if you do the experiments with salts?

LIQUORI

We have employed in our experiments with consistent results several samples of calf thymus DNA which were native according to the values of the optical density and optical rotation as measured in standard conditions.

As I have mentioned, the solubilization experiments were carried out both with DNA salt free solutions in a range of concentrations ($0.5—8.10^{-3}$ M_p) and with DNA solutions at a given concentration (8.10^{-3} M_p) in presence of 10^{-3} M NaCl following heat treatments at different temperatures (30°-100°).

Before use, the salt free DNA solutions were kept at concentrations well above the critical range of concentrations where dilution denaturation takes place ($1—2.10^{-3}$ M_p) (1).

SPIEGELMAN

It might be interesting to use such agents as dimethylsulfonate to disrupt the hydrogen bonds rather than heat.

---

(1) A.M. LIQUORI and cow., *Nature*, *184*, 1482 (1959).

Have you examined binding the aromatic hydrocarbons by RNA? And what kind of RNA have you used?

LIQUORI

We will certainly try to follow your suggestion in planning future experiments. I feel however that the formation of solubilization sites in a DNA molecule avaiable to the aromatic hydrocarbon might not necessarily require hydrogen bond fissions between the bases.

We have made some exploratory experiments with ribosomal RNA and have not been able to demonstrate a significant solubilization ability. We plan to repeat these experiments using a more sensitive method.

CHARGAFF

Then it cannot be a purine effect actually, because you would have the purine even more accessible in RNA than in DNA.

LIQUORI

DNA probably acts as an anphiphillic system with respect to its solubilization ability toward aromatic molecules. The occlusion of the hydrocarbon within the core of DNA would be promoted by perpendicular interactions with the bases. These anphiphilic sites might be present in a much smaller number in ribosomal RNA which is believed to be single stranded.

CHARGAFF

Pyrimidines have no effect in solubilizing these aromatic hydrocarbons?

LIQUORI

Pyrimidines also solubilize aromatic hydrocarbons but to a smaller extent than purines.

CHARGAFF

Is the absorption of the DNA bases changed by uptake of hydrocarbons?

LIQUORI

It would be extremely difficult to try to detect changes in the absorption spectrum of the bases of DNA in presence of the aromatic hydrocarbons because of the small fraction of bases which actually interact with the hydrocarbon molecules. On the contrary significant shifts both in tbe absorption and emission spectrum of the hydrocarbon can easily be detected.

CHARGAFF

Did you measure directly the amount of hydrocarbon taken up or extrapolate from the actual absorption spectrum of the DNA? Do you know the absorption of the amount of hydrocarbon which has been taken up, so that you could reconstruct this spectrum? I have wondered whether you see a hypochromic or a hyperchromic effect?

LIQUORI

We have measured the amount of benzpyrene solubilized by a given amount of DNA by extracting the DNA water solutions containing benzpyrene with known volumes of cyclohexane and measuring the optical density of the hydrocarbon in this solvent.

A hypochromic effect besides a bathochromic shift can be noticed in the spectrum of benzpyrene solubilized in DNA water solutions with respect to that of the hydrocarbon dissolved in cyclohexane.

SPIEGELMAN

It may be that the chain of soluble RNA was not a happy one since evidence suggests that it is much more extensively hydrogen bonded than ribosomal RNA.

LIQUORI

I am sorry about some misunderstanding. We have used ribosomal RNA.

RANZI

All these hydrocarbons are carcinogenetic substances. Is it possible that the point of attack in the carcinogenetic process is an interaction between hydrocarbons and DNA or RNA?

LIQUORI

This is obviously a possibility but I would not like to make any speculative extrapolation of our results at the present stage.

RUBIO HUERTOS

About the experiments with cells. Has the nucleous a selective absorption of the product? I mean, is it just the nucleous which takes up the hydrocarbon and not the cytoplasm?

LIQUORI

Professor DE LERMA has been able to show that when dissolved in the cytoplasm of rat liver cells, benzpyrene gives an emission spectrum essentially similar to that of the hydrocarbon in a lipophilic solvent. The emission spectrum of benzpyrene localized in the nuclei is very close to that of benzpyrene in DNA water solution or in caffeine mixed crystals.

PERUTZ

I have not quite understood the nature of this intercalation. You obtained much larger absorption of benzpyrene when you denatured the DNA. Have you ever tried to anneal the DNA afterwards according to DOTY and MARMUR and see whether the benzpyrene still stays in the DNA or whether gets out again?

[12] *Liquori* - pag. 18

LIQUORI

The hydrocarbons we have selected for our studies are virtually insoluble in water and it might prove experimentally very difficult to demonstrate that on annealing, the benzpyrene is extruded from DNA.

However I should again point out that the denaturing conditions we have used in all the experiments I have reported do not allow to assume strand separation.

GENETIC RECOMBINATION AT THE MOLECULAR LEVEL

M. MESELSON
The Biological Laboratories Harvard University
Cambridge, Mass. (U.S.A.)

Genetic methods of analysis have been of crucial importance in the study of biological processes at the molecular level but the mechanism of genetic recombination itself is largely unknown. However, many of the problems can be formulated precisely in molecular terms and some central features of mechanism have been established. Recombination by breakage and joining of DNA molecules is thought to occur with bacterial viruses and very probably with bacteria as well. The processes of molecular breakage and joining are not understood, nor is it known whether breakage and joining is the only mechanism of genetic recombination. The ensuing general discussion considers the present state of some of these problems.

Genetic information is carried by molecules of DNA and in the case of certain micro-organisms there is no doubt that the sites of genetic recombination are predominantly, if not entirely *intra*molecular. The number of mutant sites between which recombination has been observed in various bacteria and bacteriophages is far greater than the nnmber of separate DNA molecules they contain. In fact, there appears to be only one DNA molecule in bacterial viruses and in the nucleus of

E. coli [1] [2] [3] [4]. This situation may extend even to the chromosomes of organisms with recognizable meiosis, since the DNA content of an *Aspergillus, Neurospora,* or *Drosophila* chromosome is close to that of a nucleus of *E. coli.*

Our knowledge of recombination involving RNA is derived from a few studies of its occurrence between RNA viruses where, however, it might result from segregation of separate RNA molecules [5] [6]. In this discussion we wish to consider only that recombination which seems to occur at *intra*molecular sites in DNA.

The hereditary information carried by DNA is thought to reside in the linear sequence of its nucleotides. Chemical, electron microscopic, and autoradiographic examinations of DNA indicate the complete absence or extreme rareness of branches in its structure. In striking parallel, genetic linkage maps, which summarize the frequencies of recombination between mutations, show no evidence of branching. The simplest conclusion is that the nucleotide sequences of DNA molecules and the recombination maps of the genetic information they carry are co-linear. The assumption of co-linearity is supported by studies of physical and genetic linkage in transformation [7], bacterial conjugation [8], and bacteriophage recombination [2] [3].

The co-linearity of physical carriers of heredity and genetic maps has been established for many years for organisms with chromosomes susceptible to detailed cytologic invesigation [9] [10] [11]. However, it is not yet clear whether those recombinations which can be correlated with chromosome morphology occur within or between DNA molecules.

How do two or possibly more homologous DNA molecules interact to produce molecules whose nucleotide sequence is partly that of one of the interacting molecules and partly that of another? Whatever the process, it is exceedingly exact in that rearrangements of map order or mutations of other sorts seem never to result from recombination except in certain rare

cases which can be attributed to inversions or other gross inhomologies between the interacting genetic structures.

Possible models for genetic recombination may be classified according to whether or not recombinant molecules are synthesized *de novo* or, instead, contain lengths of the polynucleotide chains of the interacting parental molecules. In the first case, called copy-choice, parental molecules would act alternately only as templates for the direct or possibly indirect specification of the recombinant sequence. In the second, breakage of at least one of the interacting parent molecules is required and is followed by utilization of a fragment or fragments thereof for the construction of recombinant molecules. According to this scheme, a recombinant might be constructed entirely of parental fragments in which case the process may be described as break-and-join, or the recombinant could consist of some parental fragments and some DNA synthesized *de novo* on a parental template. This variant of breakage recombination might be called break-and-copy.

Both models, breakage and copy choice, have been entertained as explanations for recombination in bacteriophage and bacteria, and for *inter*-cistronic recombination in organisms with recognized meiosis [12] [17]. It is not known whether recombination at these two levels occurs by similar molecular mechanisms. However, it is natural that older considerations, drawn primarily from studies of *inter*-cistronic recombination in *Drosophila,* maize, and fungi, have been applied to the explanation of recombination in bacteriophage and bacteria. In spite of the usefulness and hopefulness of this comparison, it has occasionally resulted in confusion. This confusion has concerned the question of whether breakage must be associated with reciprocity in the production of recombinants. It is found in higher organisms that *inter*-cistronic (but not *intra*-cistronic) recombinants are produced almost invariably in reciprocal pairs. Such recombination is thought to occur by reciprocal breakage and joining of chromosomes. Although there is no

direct proof for this mechanism, it provides a simple explanation of three and four strand exchanges and of certain other genetic and cytological observations. Because of the apparent association of reciprocal recombination with breakage and joining in higher organisms, it is sometimes said that the two are necessarily connected. Whether or not such a connection exists, neither it nor any other connection between breakage or copy choice and reciprocity is logically necessary. Nevertheless, the apparent lack of reciprocity, especially in bacteriophage recombination where extensive efforts [18] [19] [20] have failed to detect it, has occasionally led to the view that copy choice rather than breakage is at work in such cases. Certain other arguments have been raised against breakage recombination in bacteriophage, but they are not much more compelling than the argument based on analogy to reciprocal inter-allelic exchanges [21] [22].

It has recently been found with the aid of heavy and of radioactive isotopic labels that recombinant bacteriophages contain specific lengths of parental DNA [2] [23]. The details of the distribution of parental isotope among bacteriophage of various recombinant genotypes emerging from crosses between labeled parental phage performed in unlabeled cells lead to the following conclusions:

1. Recombination occurs by breakage of parental DNA molecules although other mechanisms cannot be excluded.
2. Recombination can proceed without separation of the two strands of DNA, and therefore may not require DNA replication.
3. Distances along the genetic map are proportional to lengths along the viral DNA molecule at least for the intervals examined.

Evidence for the presence of fragments of parental DNA in recombinants comes also from studies of bacterial transforma-

tion [24] [25], so that there is now little room for doubt that breakage of parental DNA molecules and the incorporation of their fragments into recombinants is a general aspect of recombination at least in bacterial systems. It is not yet clear, however, whether breakage is followed by joining or instead by copying. A strong indication that joining of fragments does indeed follow their production by breakage comes from the quite recent finding of recombinant phages containing DNA entirely of parental origin among the progeny of labeled parents crossed in unlabeled cells [26]. It remains to be seen, however, whether the different parental contribuions to the DNA complement of such phage particles are actually joined together or whether, instead, such phages merely contain completely separate pieces of DNA each derived from only one parent. Only in the former case may it be concluded that joining has followed breakage. Nevertheless, it may be contended that the evidence for breakage and joining in bacteriophage recombination is at least as strong as that for breakage and joining in the production of *inter*-cistronic reciprocal recombinants in higher organisms.

If recombinants are produced by breakage and joining of fragments of parental DNA molecules, what is the structure of the joint? Recalling the extreme precision of genetic recombination, a plausible structure might involve an overlap between a chain of DNA from one parent and a complementary chain from the other held together through the specific hydrogen bonds of the WATSON-CRICK structure. The joint would then be a region with overlapping homologous nucleotide sequences from both of two parents and might correspond to the partial heterozygotes found to issue from phage and other crosses [27] [22] [28]. Phage heterozygotes selected from the progeny of multi-factor crosses segregate as though they were recombinant (and homozygous) for markers on either side of the heterozygous region [21] [29] [30]. The discovery of this genetic structure led some time ago to the view that phage heterozy-

gotes are intermediates in the formation of recombinants quite spart from any consideration of their physical structure [27] [21]. Subsequently, extensive studies of the segregation of phage T4 particles simultaneously heterozygous for a cluster of close markers led to the suggestion that heterozygous regions correspond to discontinuities in the viral DNA molecule such as might result from breakage and joining [31]. However, even if phage heterozygotes are intermediates in the production of recombinants by breakage and joining, it does not follow that the heterozygous region itself corresponds to a joint.

Before attempting further discussion of the structure of the joint, it may be useful to inquire into its location. We shall consider two extreme models as they apply to the bacteriophage *lambda* for which there is at present the most direct evidence of breakage and joining [2] [23] [26].

The first extreme model follows a classical view of the mechanism of breakage and joining between rod chromosomes. We consider the genetic map of *lambda* to be a faithful image of the viral DNA molecule. The mutations which recombine with highest frequency are placed at opposite ends of the linear genetic map and we might suppose that the nucleotide sequences to which they correspond lie at opposite ends of the DNA molecule. On this picture, break-and-join recombination between any two markers would correspond to breakage and joining in the nucleotide sequence somewhere between the two mutations for which recombinants are being selected. This picture, based on the most naive interpretation of the relation between map and molecule is, however, not the only possible one.

For the second extreme model, we make the assumption that *lambda* DNA molecules are not all exactly alike but that instead each is a cyclic permutation of the nucleotide sequence corresponding to the linear genetic map. Then the terminal markers of the map would usually be located well in the interior of the molecule. Let us further assume that between these markers

there lies a region of DNA or some other material with unusual properties: it is highly susceptible to breakage and joining. Let us assume that it is the *only* site at which *joining* can take place. When two phage molecules undergo breakage and joining exchange at this special site, there can in general result two recombinant products, one longer than the other. The short recombinant will not carry a full complement of genetic information and presumably cannot, in this deficient condition, mature to become a viable phage. The long recombinant contains more than a full genetic complement but in order to become a mature phage particle it must be cut down to the more or less uniform length known to occur in mature *lambda* particles [32]. This cutting is responsible for recombination about a region somewhere in the interior of the linear genetic map. If the DNA molecules are cut slightly oversize before maturation, this region could appear heterozygous. Such heterozygosity would reflect terminal redundancy of the nucleotide sequences and any attempt to correlate its properties with the structure of the joint would be misleading [33]. The boundaries of the heterozygous region would indeed correspond to sites where breakage had occurred but it would not have been followed by joining. Instead, all joining would take place following breakage at the special site which lies between the sequences corresponding to markers at the opposite ends of the simple linear map. So long as joining is restricted to this one spot, repetition of the process would cyclicly permute the nucleotide sequence but not the genetic map order. A circular genetic map [34] need not result and successful predictions of the amount of parental DNA contributed to various recombinant types could still be based on the simple linear map, as is found to be the case for bacteriophage *lambda*. The purpose of discussing this rather bizarre model is to emphasize one of the hazards of too literal an interpretation of the genetic map in the study of the physical basis of genetic recombination. In view of this and other difficulties of interpretation, we shall say no more about

the presumptive joint associated with break-and-join recombination except that nothing positive is known about its structure.

If recombination occurs by breakage, it may be asked what causes the breakage. It is possible in bacteriophage, for example, that breaks are made at random locations; let us say as a result of enzymatic activity. The various fragments produced might then undergo random encounters which would result in joining only when two fragments with homologous terminal sequences succeed in forming a stable overlap region by complementary base-pairing. Nothing known about recombination in bacteriophage rules out the possibility of such a chaotic picture. However, it is difficult to reconcile with knowledge of recombination in bacteria in the course of transformation or conjugation. This is because of the remarkable rapidity and efficiency of genetic integration in these cases. In both forms of bacterial recombination there is a probability approaching 50% that if DNA carrying a given marker enters a recipient cell, this DNA will later be found linked genetically to resident markers [7] [8]. Furthermore, studies of segregation following bacterial conjugation and of the appearance of linkage in re-isolated transforming principal suggest that linkage of the incoming to the resident markers can be accomplished in considerably less than a generation time [8] [35] [36] [37]. If there is random breakage of the host chromosome, then the efficiency and rapidity of genetic integration of even the small DNA fragments used for transformation suggests an extraordinary degree of constant fragmentation. It is easier to imagine that the sites of breakage and joining are somehow dictated by the incoming fragment itself.

The ends of the incoming fragment might penetrate the double helix of the resident chromosome and lead to the intertwining of the fragment with the chromosome. Perhaps some other form of synapsis occurs. In any event, if the breakage and joining which seems to occur in bacteriophage is a mechanism of recombination in transformation and conjugation as

[13] *Meselson* - pag. 8

well, there must follow the breakage of the resident chromosome and the insertion of the fragment, perhaps to give a completely intact recombinant structure as by transphosphorylation or perhaps to give a structure still bearing a discontinuity at the site of exchange. It may be that the detailed elucidation of these structures will require the development of systems in which recombination can proceed *in vitro,* and from which recombinant products can be isolated and characterized before the intervention of further possibly obscuring events.

REFERENCES

[1] RUBENSTEIN I., THOMAS C. A. and HERCHEY A. D., *Proc. Mat. Acad. Sci. Wash., 47,* 1113 (1961).

[2] MESELSON M. and WEIGLE J. J., *Proc. Nat. Acad. Sci. Wash., 47,* 857 (1961).

[3] KAISER A. D., *J. Mol. Biol., 4,* 275 (1962).

[4] CAIRNS J., *J. Mol. Biol., 4,* 407 (1962).

[5] BURNET F. M., « Principles of Animal Virology » (2nd ed.). Acad. Press, New York (1960).

[6] SIMPSON R. W. and HIRST G. K., *Virology, 15,* 436 (1961).

[7] RAVIN A. W., *Advances in Genetics, 10,* 62 (1961).

[8] JACOB F. and WOLLMAN E. L., *Sexuality and the Genetics of Bacteria,* Academic Press, New York (1961).

[9] STURTEVANT A. H., *J. Exptl. Zool., 14,* 43 (1913).

[10] PAINTER T. S., *J. Hered., 25,* 464 (1934).

[11] SWANSON C. P., *Cytology and Cytogenetics,* Prentice-Hall, Englewood Cliffs, N. J. (1957).

[12] BELLING J., *Univ. Calif. (Berkeley) Publs. Botany, 16,* 153 (1931).

[13] MORGAN T. H., *Science, N. S., 34,* 384 (1911).

[14] DARLINGTON C. D., *Recent Advances in Cytology.* (2nd ed.) Churchill, London (1937).

[15] STURTEVANT A. H. in HERSHEY A. D. and ROTMAN R., *Genetics, 34,* 44 (1949).

[16] LEDERBERG J., *J. Cellular Comp. Physiol., 45,* Suppl. 2, 75 (1955).

[17] DELBRUCK M. and STENT G. S., *The Chemical Basis of Heredity,* Johns Hopkins, Baltimore (1957).

[18] HERSHEY A. D. and ROTMAN R., *Genetics, 34,* 44 (1949).

[19] JACOB F. and WOLLMAN E. L., *Ann. Inst. Pasteur, 87,* 674 (1954).

[20] BRESCH C., *Z. Naturforsch., 10b,* 545 (1955).

[21] LEVINTHAL C., *Genetics, 39,* 169 (1954).

[22] JACOB F. and WOLLMAN E. L., *Ann. Inst. Pasteur, 88,* 724 (1955).

[23] KELLENBERGER G., ZICHICHI M. L. and WEIGLE J., *Proc. Nat. Acad. Sci. Wash., 47,* 869 (1961).

[24] FOX M. S., *Proc. Nat. Acad. Sci., 48,* 1043 (1962).

[25] LACKS S., *J. Mol. Biol., 5,* 119 (1962).

[13] *Meselson* - pag. 10

[26] MESELSON M., unpublished results.
[27] HERSHEY A. D. and CHASE M., *Cold Spring Harbor Symposia Quant. Biol., 16,* 471 (1951).
[28] KITANI Y., OLIVE L. S., and EL-ANIM A. S., *Science, 134,* 668 (1961).
[29] TRAUTNER T., *Ph. D. Thesis,* Göttingen (1957).
[30] KELLENBERGER G., ZICHICHI M. L. and EPSTEIN H. T., *Virology, 17,* 44 (1962).
[31] DOERMANN A. H. and BOEHNER L., unpublished manuscript.
[32] WEIGLE J., MESELSON M. and PAIGEN K., *J. Mol. Biol., 1,* 375 (1959).
[33] STREISINGER G. and MESELSON M., personal communications.
[34] STREISINGER G., EDGAR R. and HARRAR G., private correspondence.
[35] FOX M. S. and HOTCHKISS R. D., *Nature, 187,* 1002 (1960).
[36] FOX M. S., *Nature, 187,* 1004 (1960).
[37] VOLL M. J. and GOODGAL S. H., *Proc. Nat. Acad. Sci. Wash., 47,* 505 (1961).

DISCUSSION

SCHRAMM

Are there any methods known which influence the frequency of recombination in phages?

MESELSON

Bacteriophage recombination frequencies are increased by ultraviolet irradiation of parental phages or of the host cells before or shortly after infection. Recombination is also stimulated by thymine deprivation when crosses are performed in a thymineless host. Also, increased recombination frequencies are found from crosses performed at high concentrations of infected cells. So far as I know, the mechanisms of these effects are unknown.

POSSIBLE POLYPEPTIDE CONFIGURATIONS OF PROTEINS

SAN-ICHIRO-MIZUSHIMA
Tokyo Research Institute of Yawata, Kawasaki-City (Japan)

I. Introduction

The structure of molecules was originally studied by chemical methods of investigation among which the most important is one concerned with the study of the nature of chemical reactions in which substances take part. This means that at least some of the chemical reaction are characterized by the spacial configurations of molecules involved in the reactions.

Similarly it has been recognized that some of the biological reactions are closely related to the molecular configurations of proteins, e.g., the characteristic manner of the folding of polypeptide chains seems to be very important for an understanding of serological reactions. Therefore, structural studies of proteins, especially of polypeptide chains, are very interesting not only from the view point of structural chemistry but also from that of biology.

The structural studies of proteins may be divided into two categories: 1) investigations made directly on proteins and 2) determinations of molecular and crystal structures of amino acids, amides, and other simple substances closely related to

proteins. Recent advances in both the experimental and theoretical sides of structural chemistry encourage the hope that in the not distant future considerable light will be thrown on the detailed structure of many protein molecules by the application of these two complementary lines of investigation.

As to the first category experimental results have been published and quite often discussed in relation to PAULING-COREY's α-helix [1]. Internal hydrogen bonds stabilize the structure of this helix.

The simplest polypeptide chain is polyglycine $(-NHCH_2CO-)_n$, consisting of glycine residues, which usually forms a zigzag chain rather than an α-helix.

In the case of silk fibroin, infrared absorption data have revealed that its main part is a polypeptide consisting of alternately arranged glycine and alanine [2], but MARSH, COREY, and PAULING proposed a pleated sheet on the basis of X-ray diffraction studies [3].

The only example of a more complex globular protein whose structure has been determined to a certain extent is myoglobin, studied by KENDREW and others. According to the result of three-dimensional X-ray FOURIER analysis with resolving power of ca. 6A [4, 5] this molecule, comprises a chain of molecules about 10 A in diameter, similar to α-helix, turning and bending a complecated shape. Recent analysis [6] with resolving power of 2 A shows that 65-72% of the whole molecule consists of a regular α-helix. However, the structure becomes different at the corner at which the polypeptide chain turns. An interesting fact in this connection is that this result was derived by X-ray analysis and that all myoglobin molecules constituting myoglobin crystal have the same structure. This is considered to be closely related to the fact that all myoglobin molecules have the same kinds of amino acid linked in the same order, and this important aspect is the one in which the structure of

a native protein differs fundamentally from the structure of a synthetic high polymer.

Naturally, such a structure of a protein molecule may not be the most stable state of a polypeptide chain because application of heat results in denaturation and the molecule clearly undergoes change. However, the fact that these native proteins retain the original state against heat motion at ordinary temperature indicates that this form is at least in a metastable state.

The majority of synthetic fibrous peptides is in a state of equilibrium between α-helix and the so-called random coil in solution. However, this is not in a completely random state, and the stable positions taken by the internal rotation angle are very few in number.

All the polypeptide chain molecules are in some kind of stable or metastable state. Consideration will now be given to the factors determining the stable or metastable state.

II. Bond Lenght, Bond Angle and Internal Rotation Angle

The factors which determine the structure of a molecule are bond length, bond angle, and angle of internal rotation. Once these factors are determined, the spatial configuration of a molecule will be completely established. The molecule will take the most stable position of the energy curves as determined by these variables. However, we must also take into account other factors such as the hydrogen bond, the electrostatic forces between polar bonds, and the VAN DER WAALS forces.

Let us first consider the bond length. The bonding force of chemical bonds [7, 8], is far greater than the energy of thermal motion, RT (ca. 600 cal./mole at ordinary tempera-

ture). The bond length can be varied by application of a large external force. The magnitude of the change can be estimated from the values of force constants calculated from the normal frequencies of a molecule. The energy required to change the bond length by 5% is ca. 570, 3100, 2300 and 2300 cal./mole for hydrogen, oxygen, chlorine, and iodine molecules, respectively. The bond length is affected to some extent by the heat motion of a molecule and undergoes thermal vibration. The mean square amplitude of vibration is calculated from the force constant and also determined experimentally by precision measurement of electron diffraction of a gas. It was found that the variation due to thermal motion is small.

The changes in bond angle by thermal motion can be treated quite similarly as those of bond length; only the force constants are smaller than those of bond length. The energies required to change the angles, $<$HCH, $<$FCF, and $<$ClCCl by 5% are ca. 280, 340 and 200 cal./mole, respectively. Another force that affects the bond length and the bond angle is the repulsive force between atoms which are in close proximity, but not bonded directly [8]. In some cases, this force stretches the bond slightly from its equilibrium length and makes the bond angle larger than that in equilibrium state. This force can be obtained from the analysis of the frequencies of a polyatomic molecules. For example, the energies required to change the distance between Cl and Cl in Cl-C-Cl and that between Br and Br in Br-C-Br by 5% are 960 and 850 cal./mole, respectively.

Thus the force keeping the bond length and bond angle constant is great and is hardly affected by thermal motion. However, as explained above, these two molecular constants cannot determine completely the configuration of a molecule and in addition to these we have to take into account the angle of internal rotation. In order to determine the value of this angle in the polypeptide chain, we have

TABLE I

Stable Configurations of Rotational Isomers. (Figures in parentheses attached to the less stable isomer denote the energy difference in kilocalories per mole).

Molecule	Solid	Liquid	Gas	Ref.
$ClH_2C\text{-}CH_2Cl$	T	T,G(ca.0)	T,G(1.14)	(9)
$BrH_2C\text{-}CH_2Br$	T	T,G(0.73)	T,G(1.70)	(9)
$ClH_2C\text{-}CH_2Br$	T	T,G(0.42)	T,G(1.46)	(9)
$ClH_2C\text{-}CH_2I$	T	T,G		(10)
$CH_3CH_2\text{-}CH_2CH_3$	T	T,G(0.8)	T,G	(10)
$CH_3CH_2\text{-}CH_2Cl$	T	T,G(ca.0)	T,G	(11)
$CH_3CH_2\text{-}CH_2Br$	T	T,G(ca.0)	T,G(ca.0)	(11)
$HOH_2C\text{-}CH_2Cl$	G	G,T	G,T(0.95)	(12)
$Cl\text{-}CH_2\text{-}CH(CH_3)\text{-}CH_3$	—	—	$T,G_I(0.2)$	(13)
$Cl\text{-}CH(CH_3)\text{-}CH_2\text{-}CH_3$	—	T,G_I	$T,G_I(0.7)$	(13)
$Cl\text{-}CH_2\text{-}CH(Cl)\text{-}CH_3$	—	$T,G_I(0.2),G_{II}(0.9)$	$T,G_I(1.2),G_{II}(1.9)$	(13)
$CH_3\text{-}CH(Br)\text{-}CH_2\text{-}Br$	—	$T,G_I(1.0)$	$T,G_I(1.7)$	(13)
$CH_3\text{-}CH(Br)\text{-}CH_2\text{-}CH_3$	—	—	$T,G_I(0.7),G_{II}(1.3)$	(13)
$Cl\text{-}C(CH_3)_2\text{-}CH_2\text{-}Cl$	T	T,G(ca.0)	T,G (less stable)	(14)

to make the assumption that the angle is almost the same as that in a simpler molecule with a structure similar to that of a part of the polypeptide chain. Experimental results on the potential of internal rotation obtained for various molecules are listed in Table. I. In this Table T and G designate the trans form and the gauche form, respectively. In the case of $ClH_2C\text{-}CH_2Cl$ the trans form is a staggered form with two chlorine atoms as far as possible and the gauche forms are obtained from the trans form by internal rotation about the C-C axis by ca. $\pm 120°$ (see Figure 1). Many studies have been made by the present author and their collaborators regarding this potential of internal rotation. The experimental results obtained can be summarized as follows [10, 15].

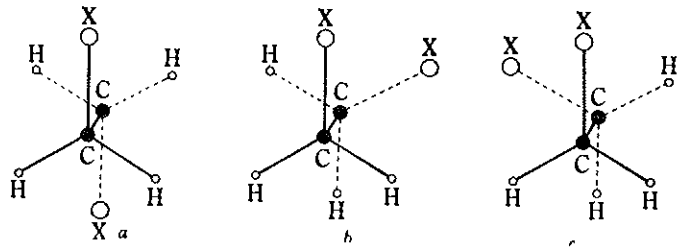


FIG. 1 — Stable configurations of 1,2-dihalogenoethanes.

1) the stable angle of internal rotation is nearly the same in almost all the cases irrespective of the molecules being in the solid, liquid, or vapor state or in solutions. This means that the molecular form of a rotational isomer is independent of the environment;

2) on the other hand, energy of internal rotation changes greatly with the state of aggregation. This means that the ratio of rotational isomers changes greatly with surroundings;

3) the factors determining these stable angles and energy of internal rotation are considered to be the steric repulsion of atoms or atomic groups adjacent to the rotational axis, the dipole moment of the bond adjacent to the axis, the internal hydrogen bonding, and the partial double bond character of the internal rotation axis.

The height of the energy barrier between stable positions of the internal rotation has been obtained only for simple molecules, but the energy required to change the angles of internal rotation slightly from the stable position can be calculated for many molecules from the frequency of torsional vibration about the axis of internal rotation. Some typical values are listed in Table II. In the case of CH_3CH_3 the energy required to change the angle by 10° is found to be 184 cal./mole.

TABLE II

Energy Barrier (Vb) and Frequency (Vr) of Internal Rotation

Molecule	V_b cal./mole	V_r cm.$^{-1}$	
CH_3-CH_3 (gas)	2750	275	(16, 17)
CH_3-OH (gas)	1070	—	(18)
$ClCH_2$-CH_2Cl (liq.) . .			(19)
trans	—	125	
gauche	—	134	

The change of energy with change in bond length, bond angle, and angle of internal rotation has been discussed above. In addition to this, the change of molecular energy with hydrogen bonding plays an important part in determining the stable molecular configuration. For example the energy required to

change the distance of OH...O by the formation of hydrogen bond of OH group in formic acid by 5% is 1250 cal./mole. This change of energy by hydrogen bonding is fairly great, but when a molecule is dissolved in water, a water molecule will be involved in the intermolecular hydrogen bond of the polymers and the effective value of hydrogen bond energy that determines the configuration of a molecule becomes much smaller.

Other forces which may affect the configuration of a molecule, particularly in its angle of internal rotation, are interionic forces, dipolar forces, and VAN DER WAALS forces, but they are smaller than the force of hydrogen bonding.

These forces do not affect the molecular configurations of stable forms (i.e., rotational isomers), but considerably affect their ratio in equilibrium. For example, for rotational isomers with different dipole moments, those with a smaller moment will be stabilized in a nonpolar solvent, whereas those with a larger moment will be more stable in a polar solvent. The stabilization arises from the interaction between molecular dipoles [10, 23].

III. ROTATION ABOUT THE CO-NH BOND AS AXIS

The polypeptide chain $(\text{-NH-CHR-CO-})_n$ has three kinds of axes of internal rotation (cf. Fig. 2). The first is the CO-NH bond, the second the CHR-CO bond and the third the NH-CHR bond. The potential energy of internal rotation about these axes can be estimated from those of simple molecules with structure similar to that of a polypeptide chain. Of these, the peptide bond, the most important part of the polypeptide chain, will be discussed first.

The internal rotation about this axis has been studied from both X-ray diffraction and infrared absorption investigations. From the result of X-ray diffraction studies on acetani-

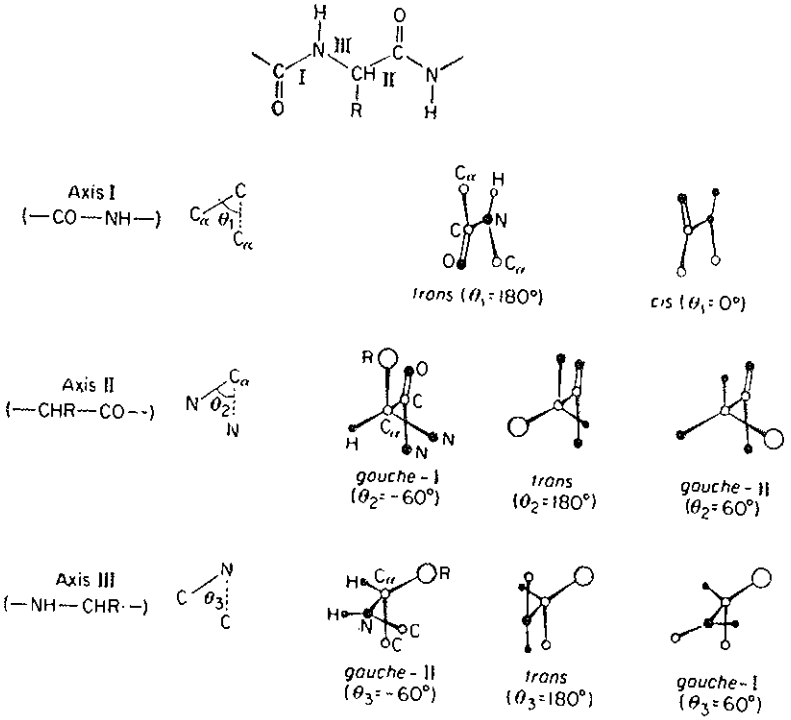
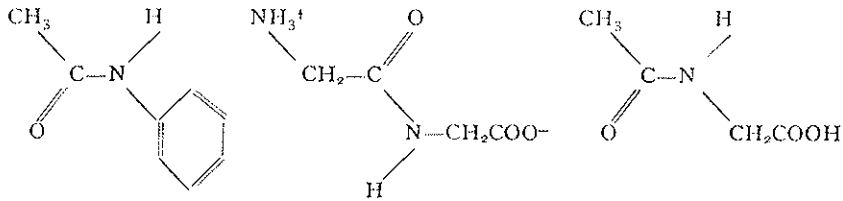


FIG. 2 — Stable angles of internal rotation in the polypeptide chain.

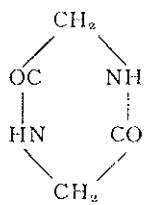
lide [24] (I), glycylglycine [25] (II), and acetylglycine [26] (III), R_1 and R_2 in R_1-CO-NH-R_2 were found to be in the trans position to each other.



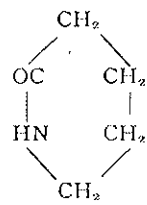
(I) (II) (III)

In connection with infrared absorptions, detailed examinations have been made on N-methylacetamide [27]. Measurements of ultraviolet absorption, RAMAN effect, and dipole moment have also been carried out at the same time with this molecule. The calculation of normal vibrations has also been made, and the molecule was proven to take the trans form in the vapor, liquid, and solid states [28, 29]. With regard to the infrared evidence that the peptide bond takes the trans form following three experimental results are important: 1) appearance of free NH absorption at 3440-3470 cm.$^{-1}$ in carbon tetrachloride solution (in the case of cis form, the free NH absorption appears at a frequency lower by 20-40 cm.$^{-1}$); 2) with increasing concentration of the carbon tetrachloride solution, an associated band appears at ca. 3370 cm.$^{-1}$, which gradually shifts to a lower wavenumber region, and is finally located at ca. 3300 cm.$^{-1}$; 3) in the associated state, another NH stretching band and the so-called amide-II band appear at ca. 3080 and 1550 cm.$^{-1}$, respectively.

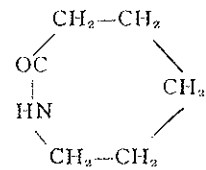
However, the molecules with ring structures such as diketopiperazine (IV), δ-valerolactam (V), and ε-caprolactam (VI) cannot take the trans form.



(IV)



(V)



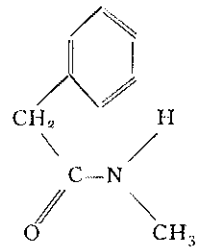
(VI)

By the measurement of the free NH stretching vibration it has been confirmed that a considerable amount of the cis form is present in formanilide in carbon tetrachloride solution and a

[14] *Mizushima* - pag. 10

very minute quantity of the cis form is present in acetanilide under the same condition [30]. This is due to the steric interference between the oxygen atom in the peptide group and benzene ring. In an ordinary peptide not containing the benzene ring, the cis peak arising from the free NH stretching vibration is hardly observed.

The cis peak of formanilide mentioned above was first observed by RUSSELL and THOMPSON [31]. N-methylphenylacetamide (VII) shows a peak apparently in the same position as the cis peak. In this molecule, the NH group is bonded to the benzene ring in a manner similar to that in hydrogen



(VII)

bonding, but the peptide bond itself is considered to take the trans form [32].

It may be concluded from the foregoing that owing to the considerable double bond character of the central CN bond, the CO-NH peptide group takes either the trans or the cis form in the internal rotation. But in general, this bond is more stable in the trans form and rarely takes the cis form, except when a large group like the benzene ring is present next to the NH group.

IV. Rotation about the CHR-CO and NH-CHR Bonds as Axes

Considerations will now be made regarding internal rotation about the CHR-CO and NH-CHR axes, the second and third axes of the polypeptide chain. For the CHR-CO axis the following molecules have been studied:

1) $Cl\text{-}CH_2\text{-}CO\text{-}Cl$ [33]. This molecule is present in the trans and gauche forms with respect to the two Cl atoms. In the solid state, only the trans form is stable. Angles of internal rotation in both the trans and gauche forms are somewhat different from the ideal values, 0 and 120°;

2) $Br\text{-}CH_2\text{-}CO\text{-}Cl$ [33]. Same as in (1). In the vapor state, the trans form is more stable than the gauche form by 1000 cal./mole;

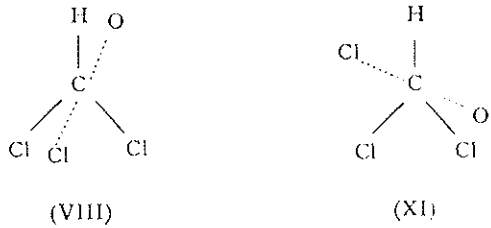
3) $Br\text{-}CH_2\text{-}CO\text{-}Br$ [33]. Same as in (1). In the vapor state, the trans form is more stable than the gauche form by 1900 cal./mole;

4) $Cl\text{-}CH_2\text{-}CO\text{-}CH_3$ [34]. Both the trans and gauche forms with respect to the Cl and CH_3 groups are present. In the solid phase, only the trans form is stable; in the liquid state, the two forms coexist in approximately the same amount; in the vapor state, practically all the molecules take the gauche (more polar) form;

5) $Cl\text{-}CH_2\text{-}CO\text{-}NH\text{-}CH_3$ [34]. Both the trans and gauche forms with respect to the Cl and NH groups are present. In the solid state two crystalline forms are found, one with the trans form alone and the other with the gauche form alone. In the liquid state the two forms coexist, and in the vapor state almost all the molecules are in the gauche form;

[14] *Mizushima* - pag. 12

6) Cl-CH-CO-Cl [36]. Two configurations (VIII) and (IX) shown in Figure 3, are present. In the vapor phase, (VIII) is more stable than (IX) by 200 cal./mole. The energy of the two forms is almost the same in the liquid state, whereas only (IX) persists in the solid state;



(VIII) (XI)

FIG. 3 — Rotational isomers of dichloroacetyl chloride Cl_2CH-COCl as viewed along the C-C axis

7) Cl-CH-CO-Cl [37]. Three configurations, the trans, gauche-I, and gauche-II forms exist. In the solid state, only the trans form is stable, whereas in the liquid state, the trans and the gauche-I forms are present with a small amount of the gauche-II form, the energy difference between the trans and gauche-I forms being almost zero. In the vapor state, all the three configurations are found, the amount of the gauche-II form being greater than that in the liquid state;

8) Br-CH-CO-Cl [37]. Only the gauche-I form is present in the solid state, and both the trans and gauche-I forms are found in the liquid and vapor states;

9) Br-CH-CO-Br [37]. Both the trans and gauche-I forms are present in all the vitreous, vapor, and liquid states;

10) $Cl\text{-}CH(CH_3)\text{-}CO\text{-}NH\text{-}CH_3$ [37]. Only the trans form is present in the solid state and both the trans and gauche-I forms are found in the liquid state. Practically all the molecules are in the gauche-I form in the vapor state. The trans form is more stable than the gauche-I form by ca. 1400 cal./mole in the liquid state.

Of the substances mentioned above, all the molecules of (5) and (10) take the trans form with respect to the CO-NH axis.

As shown above the internal rotation about the CHR-CO bond as axis has three potential minima corresponding to the trans, gauche-I and gauche-II forms. The former two forms are more stable than the gauche-II form.

As to the internal rotation about the NH-CHR bond as axis almost no experimental data are available. However, the similarity in the RAMAN specta between $CH_3\text{-}CH_2\text{-}NH\text{-}CH_2\text{-}CH_3$ and $CH_3\text{-}CH_2\text{-}CH_2\text{-}CH_2\text{-}CH_3$ suggests that the rotation about the NH-CHR axis bears a close resemblance to that about CH_2-CHR axis. Consequently, the trans and gauche-I forms seem to be more stable, and the gauche-II form is less stable.

V. POSSIBLE CONFIGURATIONS OF A POLYPEPTIDE CHAIN

From the studies on the internal rotation explained above, possible configurations of a polypeptide chain can be determined. Actually not all the stable configurations will appear owing to the effect of the structure of the side chain and other factors. The angles of internal rotation are not exactly equal to 60° and 180°, and some deviations from these ideal values will take place owing to the energy relation explained in Section II. In addition to the energy of internal rotation the energy of hydrogen bond must also be taken into consideration.

Let us first assume similar repetition of the equilibrium values of the azimuthal angles θ_1, θ_2, and θ_3 about axes I, II, and III, shown in Figure 2. Then we see that a polypeptide chain has a helical structure which can be described by three coordinates, d, Φ, and ρ, where d is the translation along the helical axis, Φ is the rotation angle, and ρ is the distance of an atom from the axis (See Fig. 4).

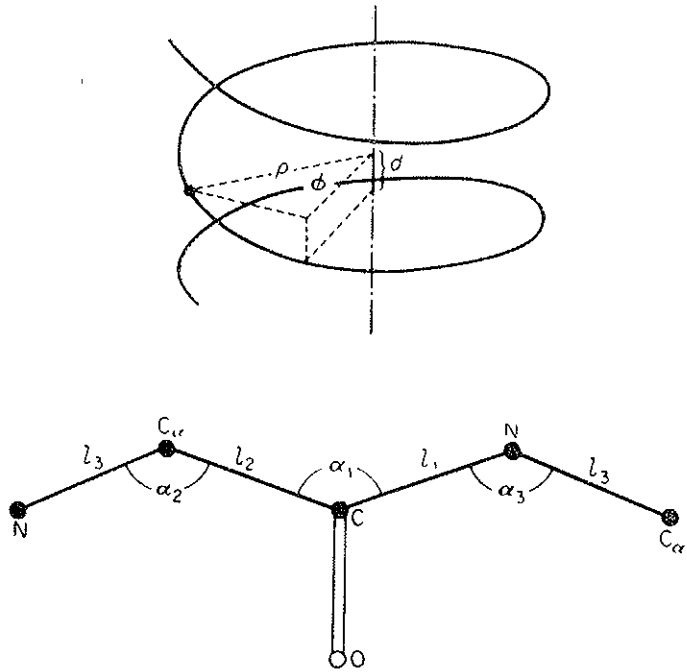


FIG. 4 — Parameters in the polypeptide chain.

The relation between θ_1, θ_2, θ_3 and d, Φ, ρ was discussed partly by PAULING and COREY by the use of a model [38]. However, we obtained a general mathematical expression for this relation using the matrix derived by EYRING [39, 40]. Let

the interatomic distances and bond angles of the polypeptide chain be denoted as shown in Figure 4. Then the values of d, Φ, and ρ of C_α atom in the polypeptide chain can be calculated as [40].

$$d = \frac{b_1 (a_{13} + a_{31}) + b_2 (a_{23} + a_{32}) + b_3 (a_{33} - a_{11} - a_{22} + 1)}{(3 - a_{11} - a_{22} - a_{33}) (a_{33} - a_{11} - a_{22} + 1)^{\frac{1}{2}}} ,$$

$$\cos \Phi = (a_{11} + a_{22} + a_{33} - 1)/2 ,$$

$$\rho^2 = (b_1^2 + b_2^2 + b_3^2 - d^2)/(3 - a_{11} - a_{22} - a_{33}) ,$$

where

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} -\cos \alpha_3 & & -\sin \alpha_3 & & 0 \\ \sin \alpha_3 \cos \theta_3 & & -\cos \alpha_3 \cos \theta_3 & & -\sin \theta_3 \\ \sin \alpha_3 \sin \theta_3 & & -\cos \alpha_3 \cos \theta_3 & & \cos \theta_3 \end{pmatrix}$$

$$\times \begin{pmatrix} -\cos \alpha_1 & & -\sin \alpha_1 & & 0 \\ \sin \alpha_1 \cos \theta_1 & & -\cos \alpha_1 \cos \theta_1 & & -\sin \theta_1 \\ \sin \alpha_1 \sin \theta_1 & & -\cos \alpha_1 \sin \theta_1 & & \cos \theta_1 \end{pmatrix}$$

$$\times \begin{pmatrix} -\cos \alpha_2 & & -\sin \alpha_2 & & 0 \\ \sin \alpha_2 \cos \theta_2 & & -\cos \alpha_2 \cos \theta_2 & & -\sin \theta_2 \\ \sin \alpha_2 \sin \theta_2 & & -\cos \alpha_2 \sin \theta_2 & & \cos \theta_2 \end{pmatrix}$$

$$b_1 = l_2 \cos \alpha_3 \cos \alpha_1 - l_2 \sin \alpha_3 \sin \alpha_1 \cos \theta_1 - l_1 \cos \alpha_3 + l_3$$

$$b_2 = -l_2 \sin \alpha_3 \cos \theta_3 \cos \alpha_1 - l_2 \cos \alpha_3 \cos \theta_3 \sin \alpha_1 \cos \theta_1 - l_2 \sin \theta_3 \sin \alpha_1 \sin \theta_1 + l_1 \sin \alpha_3 \cos \theta_3$$

$$b_3 = -l_2 \sin \alpha_3 \sin \theta_3 \cos \alpha_1 - l_2 \cos \alpha_3 \sin \theta_3 \sin \alpha_1 \cos \theta_1 + l_2 \cos \theta_3 \sin \alpha_1 \sin \theta_1 + l_1 \sin \alpha_3 \sin \theta_3$$

As explained above, the experimental result obtained for simple molecules shows that the most probable values of stable

[14] *Mizushima* - pag. 16

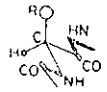
azimuthal angles are: $\theta_1 = 180°$, $\theta_2 = 180°$ and $\pm 60°$, and $\theta_3 = 180°$ and $\pm 60°$. From these values of azimuthal angles we can calculate d, Φ, and φ which determine the configuration of the polypeptide chain.

TABLE III

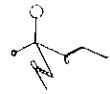
Configuration	-CHR-CO	θ_2	-NH-CHR-	θ_3	d (A)	Φ	φ (A)
a	trans	180°	trans	180°	3.64	180°	0.57
b	gauche-I	− 60°	trans	180°	− 2.53	92°	1.98
c	gauche-II	60°	trans	180°	2.53	92°	1.98
d	trans	180°	gauche-I	60°	2.69	85°	2.00
e	gauche-I	− 60°	gauche-I	60°	2.66	180°	1.38
f	gauche-II	60°	gauche-I	60°	− 1.17	89°	2.58
g	trans	180°	gauche-II	− 60°	− 2.69	85°	2.00
h	gauche-I	− 60°	gauche-II	− 60°	1.17	89°	2.58
i	gauche-II	60°	gauche-II	− 60°	2.66	180°	1.38

According to the results of the calculations, shown in Table III, there are nine different configurations of the polypeptide chain. (See Fig. 5). As can be shown from the study of simple molecules discussed above, the configurations with $\theta_2 = 60°$ or $\theta_3 = -60°$ are less stable than the other configurations. In other words, configurations *c*, *f*, *g*, *h*, and *i* are less stable than configurations *a*, *b*, *d*, and *e*. Of these four, *b* and *d* correspond to helices in which all the NH groups are directed inside and all the CO groups are directed outside, or all the NH groups outside and all the CO groups inside. Accordingly *b* and *d* cannot form stable hydrogen bonds and need not be considered.

Of the two remaining stable forms, *a* is the extended form of the polypeptide chain which was proposed by MEYER and

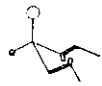
(a) 

β-Keratin proposed by MEYER and MARK

(b) 

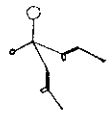
(c) 

(d) 

(e) 
α-Keratin proposed by AMBROSE and HANBY and SHIMANOUCHI and MIZUSHIMA

(f) 
Left-handed α-elix proposed by PAULING and COREY

(g) 

(h) 
Right-handed α-helix proposed by PAULING and COREY

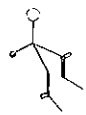
(i) 
α-Keratin proposed by HUGGINS

FIG. 5 — Stable configurations of a polypeptide chain unit.

[14] *Mizushima* - pag. 18

MARK [41]. It has been considered to correspond to the structure of β-keratin and silk fibroin and *e* is the bent or folded form proposed by us [42] and by AMBROSE et al. [43].

Of the four less stable forms, *c* and *g* need not be considered for the same reason as that in the case of *b* and *d*. Form *i* is the folded one proposed by HUGGINS [44] which is similar to *e*, but it quite different from *e* with respect to the position of the side chain. In the remaining two forms *f* and *h*, one C=O group forms the hydrogen bond with NH group in every second residue. If we change the azimuthal angles slightly ($\theta_2 = 55°20'$, $\theta_3 = -48°40'$ or $\theta_2 = -55°20'$ $\theta_3 = 48°40'$) these two forms become identical with helices proposed by PAULING and COREY [38].

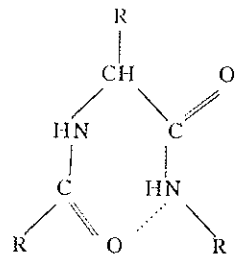
In conclusion we have five stable configurations, *a*, *e*, *i*, *f*, and *h*. Of these *i*, *f*, and *h* are less stable than *a* (extended configuration) and *e* (folded configuration) as far as the energy of the internal rotation is concerned.

Let us discuss each of these five stable configurations in more detail. The extended configuration or *a* corresponds to the structure of β-keratin and silk fibroin. However, if we use the experimental values of bond lengths and bond angles obtained by COREY and DONOHUE [45] the fiber period is calculated as 7.27 A. which is considerably larger than the experimental values 6.64 A. of silk fibroin [45]. In order to explain this experimental values, PAULING and COREY [38] proposed a new pleated sheet which corresponds to the values of azimuthal angle of $\theta_1 = 180°$, $\theta_2 = -120°$, and $\theta_3 = 120°$. The last two values ($\theta_2 = -120°$ and $\theta_3 = 120°$) correspond to potential maxima of internal rotation.

In the ideal extended configuration the value of θ_2 and θ_3 is 180°. However, the actual value may be a little different from that. Moreover, the bond lengths and bond angles of polypeptide chain may be different from those found by COREY and DONOHUE [45] for simple molecules, because the difference in the strength of hydrogen bond formed in the system

of polypeptides will affect considerably the amount of resonance of $-C(=O^-)=N^+H$-structure which will, in turn, affect the bond length and bond angles of the polypeptide chain. If we take into account these changes in azimuthal angle and in bond length and bond angle, we may be able to calculate the fiber period of the nearly extended configuration in good agreement with that observed.

The folded configuration or *e* was discussed by us in 1947 [42] and independently by AMBROSE et al. in 1949 [43]. We could prove definitely the existence of the folded form (X) with internal hydrogen bond for the type of molecules of RCONHCHR'CONHR, which can be considered to correspond to a part of a polypeptide chain [47-53].



(X)

With $\theta_1 = 180°$, $\theta_2 = -60°$, and $\theta_3 = 60°$ the N-H···O distance in the folded form becomes 2.2 A. which would be too small as the non-bonded N···O distance. However, with $\theta_1 = 180°$, $\theta_2 = -40°$ and $\theta_3 = 70°$ we obtain a reasonable distance for N-H···O.

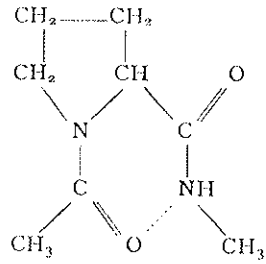
Configuration *i* also forms a seven-membered ring similar to *e*. From the internal rotation potential it will be seen that

this configuration is less stable than *e* in which the side chain is in the general plane of the ring, since *i* contains more gauche structure than *e*. This is compatible with the experimental results obtained for halogenocyclohexanes [54, 57] and the pyranose ring [58] according to which the configuration with the side chain in the general plane of the ring is much more stable than the other configuration with side chain perpendicular to the plane of the ring.

VI Conclusion

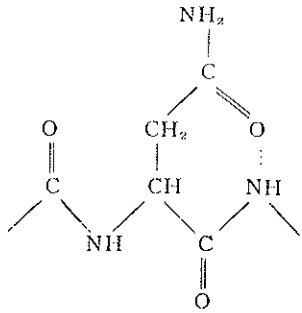
Other factors that affect the structure of a polypeptide chain are the configuration and ionic property of a side chain in an amino acid residue. Some conclusions have been derived for the structure of side chains from the studies made on thiobutane [59] and glutamate [60] and on the crystal structure of various amino acids [61]. These side chains are considered to affect the configuration of the main chain by the electrostatic force or steric hindrance.

The very intricate form of myoglobin may be due to the fact that the order of amino acid residues in its polypeptide chain is complicated and each one of them affects the configuration of the adjacent residues. This explains the fact the myoglobin molecule does not take the random form, but has a regular configuration. In this connection we would like to point out that in our experimental results the equilibrium ratio of the extended and folded forms of $CH_3CONHCHR'CONHCH_3$ depends considerably on the nature of R. An extreme case is that of acetylproline N-methylamide, in which we found only the folded form (XI) [48, 62]. Formation of a hydrogen bond would also affect the configuration of the main chain.

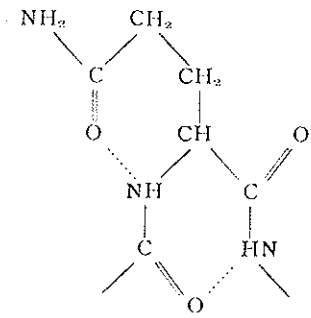


(XI)

For instance, asparagine and glutamine residues form hydrogen bonds which result in different main chain configurations (XII, XIII).



(XII)



(XIII)

From these facts we can conclude that in natural polypeptides some amino acid residues tend to take extended forms whereas other folded forms or helical forms and a specific combination of these forms correspond to the specificity of the protein.

The order of the amino acid residues has been determined in insulin and other hormonal proteins and in some of the enzymes. If an amino acid residue controls the configuration of the main chain in its vicinity, each of the residues will take

the complicated configuration characteristic of the chain. It is considered that the various interesting actions of proteins in vivo are explained primarily by the fact that each of the protein molecules has such a characteristic structure.

Infrared absorption studies indicate that a polypeptide chain with glycine and alanine linked alternately forms the main portion of silk fibroin. In such a case the regularity of the configuration can be understood from the same point of view. It seems natural that synthetic polypeptides form regular structures such as the α-helix, because they are all built from the same amino acid residues.

ACKNOWLEDGEMENT

This is a summary of a past of the work done in our laboratory. I wish to thank those who have collaborated with me in this field, particularly Prof. SHIMANOUCHI who has worked with me for the past twenty years.

REFERENCES

[1] PAULING L. and COREY R.B., *Proc. Nat. Acad. Sci., 37,* 205, 235, 251, 256, 261, 272 and 282 (1951).

[2] ASAI M., TSUBOI M., SHIMANOUCHI T., and MIZUSHIMA S., *J. Phys. Chem., 59,* 322 (1955).

[3] MARSH R.E., COREY *R.B.,* and PAULING L., *Biochim. et Biophys. Acta, 16,* 1 (1955); *Acta Cryst., 8,* 710 (1955).

[4] KENDREW J.C., BODO G., DINTZIS H.M., PARRISH *R.G.,* WYCKOFF H.W., and PHILLIPS D.C., *Nature, 181,* 662 (1958).

[5] BODO G., DINTZIS H.M., KENDREW J.C., and WYCKOFF H.W., *Proc. Roy. Soc., A253,* 70 (1959).

[6] KENDREW J.C., DICKERSON R.E., STRANDBERG B.E., HART R.G., DAVIES D.R., PHILLIPS D.C., and SHORE V.C., *Nature, 185,* 422 (1960).

[7] HERZBERG G., *Spectra of Diatomic Molecules,* Van Nostrand, New York, 1950.

[8] MIZUSHIMA S., and SHIMANOUCHI T., *Infrared Absorption and Raman Effect,* Kyoritsu Shuppan, Tokyo, 1958.

[9] KURATANI K., MIYZAWA T., and MIZUSHIMA S., *J. Chem. Phys., 21,* 1411 (1953).

[10] MIZUSHIMA S., *Structure of Molecules and Internal Rotation,* Academic Press, New York, 1954.

[11] KOMAKI C., ICHISHIMA I., KURATANI K., MIYZAWA T., SHIMANOUCHI T., and MIZUSHIMA S., *Bull. Chem. Soc. Japan, 28,* 330 (1955).

[12] MIZUSHIMA S., SHIMANOUCHI T., MIYAZAWA T., ABE K., and YASUMI M., *J. Chem. Phys., 19,* 1477 (1951).

[13] MIZUSHIMA S., SHIMANOUCHI T., and TSUCHIYA S. (to be published).

[14] HAYASHI M., ICHISHIMA I., SHIMANOUCHI T., and MIZUSHIMA S., *Spectrochim. Acta, 10,* 1-7 (1957).

[15] MIZUSHIMA S., MORINO Y., and SHIMANOUCHI T., *J. Phys. Chem., 56,* 324 (1952).

[16] KEMP J.D., and PITZER K.S., *J. Am. Chem. Soc., 59,* 276 (1937).

[17] KISTIAKOWSKY G.B., LACHER J.R., and STITT FRED, *J. Chem. Phys., 7,* 289 (1939).

[18] IVASH E.V., and DENNISON D.M., *J. Chem. Phys., 21,* 1804 (1953).

[19] ICHISHIMA I., KAMIYAMA H., SHIMANOUCHI T., and MIZUSHIMA S., *J. Chem. Phys., 29,* 1190-91 (1958).

[20] KARLE J., and BROCKWAY L.O., *J. Am. Chem. Soc., 66,* 574 (1944).

[14] *Mizushima* - pag. 24

[21] MIYAZAWA T., and PITZER K.S., *J. Am. Chem. Soc., 81,* 74-79 (1959).
[22] COOLIDGE A.S., *J. Am. Chem. Soc., 50,* 2166-78 (1928).
[23] WATANABE I., MIZUSHIMA S., and MASIKO Y., *Sci. Pap. Inst. Phys. Chem. Research,* Tokyo, *40,* 425 (1943).
[24] BROWN C.J., and CORBRIDGE D.E.C., *Nature, 162,* 72 (1948).
[25] HUGHES D.W., and MOORE W.Y., *J. Am. Chem. Soc., 71,* 2618 (1949).
[26] CARPENTER G.B., and DONOHUE J., *J. Am. Chem. Soc., 72,* 2315 (1950).
[27] MIZUSHIMA S., SHIMANOUCHI T., NAGAKURA S., KURATANI K., TSUBOI M., BABA H., and FUJIOKA O., *J. Am. Chem. Soc., 72,* 3490 (1950).
[28] MIYAZAWA T., SHIMANOUCHI T., and MIZUSHIMA S., *J. Chem. Phys., 24,* 408 (1956).
[29] MIYAZAWA T., SHIMANOUCHI T., and MIZUSHIMA S., *J. Chem. Phys., 29,* 611 (1958).
[30] SUZUKI I., TSUBOI M., SHIMANOUCHI T., and MIZUSHIMA S., *Spectrochim. Acta, 16,* 471-78 (1960).
[31] RUSSELL R.A., and THOMPSON H.W., *Spectrochim. Acta, 8,* 138 (1956).
[32] SUZUKI I., TSUBOI M., and SHIMANOUCHI T., *Spectrochim. Acta, 16,* 467 (1960).
[33] NAKAGAWA I., ICHISHIMA I., KURATANI K., MIYAZAWA T., SHIMANOUCHI T., and MIZUSHIMA S., *J. Chem. Phys., 20,* 1720 (1952).
[34] MIZUSHIMA S., SHIMANOUCHI T., MIYAZAWA T., ICHISHIMA I., KURATANI K., NAKAGAWA I., and SHIDO N., *J. Chem. Phys., 21,* 815 (1953).
[35] MIZUSHIMA S., SHIMANOUCHI T., ICHISHIMA I., MIYAZAWA T., NAKAGAWA I., and ARAKI T., *J. Am. Chem. Soc., 78,* 2038 (1956).
[36] MIYAKE A., NAKAGAWA I., MIYAZAWA T., ICHISHIMA I., SHIMANOUCHI T., and MIZUSHIMA S., *Spectrochim. Acta, 13,* 161 (1958).
[37] MIZUSHIMA S., SHIMANOUCHI T., and FUKUSHIMA K. (to be published).
[38] PAULING L., and COREY R.B., *Proc. Natl. Acad. Sci. U.S., 37,* 241 (1951).
[39] EYRING H., *Phys. Rev., 39,* 746 (1932).
[40] SHIMANOUCHI T., and MIZUSHIMA S., *J. Chem. Phys., 23,* 707 (1955).
[41] MEYER K.H., and MARK H., *Ber. deut. chem. Ges., 61,* 1932 (1928).
[42] SHIMANOUCHI T., and MIZUSHIMA S., *Kagaku, 17,* 24, 52 (1947); *Bull. Chem. Soc. Japan, 21,* 1 (1948); *Nature, 164,* 918 (1949).
[43] AMBROSE E.J., and HANBY W.E., *Nature, 163,* 483 (1949); AMBROSE E.J., ELLIOT A., and TEMPLE R.B., *Nature, 163,* 859 (1949).
[44] HUGGINS M.L., *Chem. Rev., 32,* 195 (1943).
[45] COREY R.B., and DONOHUE J., *J. Am. Chem. Soc., 72,* 2899 (1950).
[46] ASTBURY W.T., and STREET A., *Phil. Trans. Roy. Soc., A230,* 75 (1931).

[47] MIZUSHIMA S., SHIMANOUCHI T., TSUBOI M., SUGITA T., KATO E., and KONDO E., *J. Am. Chem. Soc., 73,* 1330 (1951).

[48] MIZUSHIMA S., SHIMANOUCHI T., TSUBOI M., SUGITA T., KÙROSAKI K., MATAGA N., and SODA R., *J. Am. Chem. Soc., 74,* 4639 (1952).

[49] MIZUSHIMA S., SHIMANOUCHI T., TSUBOI M., KURATANI K., SÙGITA T., MATAGA N., and SODA R., *J. Am. Chem. Soc., 75,* 1863 (1953).

[50] MIZUSHIMA S., TSUBOI M., SHIMANOUCHI T., SUGITA T., and YOSHIMOTO T., *J. Am. Chem. Soc., 76,* 2479 (1954).

[51] MIZUSHIMA S., TSUBOI M., SHIMANOUCHI T., and ASAI M., *J. Am. Chem. Soc., 76,* 6003 (1954).

[52] MIZUSHIMA S., SHIMANOUCHI T., TSUBOI M., and ARAKAWA T., *J. Am. Chem. Soc., 79,* 5357 (1957).

[53] TSUBOI M., SHIMANOUCHI T., and MIZUSHIMA S., *J. Am. Chem. Soc., 81,* 1406 (1959).

[54] HASSEL O., *Research, 3,* 504 (1950).

[55] KOZIMA K., SAKASHITA K., and MAEDA S., *J. Am. Chem. Soc., 76,* 1965 (1954).

[56] KOZIMA K., and YOSHINO T., *J. Am. Chem. Soc., 57,* 166 (1953).

[57] KOZIMA K., SAKASHITA K., and TAKEISHI Y., *Bull. Chem. Soc. Japan, 27,* 287 (1954).

[58] REEVES R.E., *J. Am. Chem. Soc., 72,* 1499 (1950).

[59] HAYASHI M., SHIMONOUCHI T., and MIZUSHIMA S., *J. Chem. Phys., 26,* 608 (1957).

[60] TSUBOI M., to be published.

[61] HAYASHI M., SHIMANOUCHI T., and MIZUSHIMA S., *Bull. Chem. Soc. Japan, 29,* 802 (1956).

[62] MIZUSHIMA S., SHIMANOUCHI T., KURATANI K., TSUBOI M., SÙGITA T., NAKAGAWA I., and KUROSAKI K., *Nature, 169,* 1058 (1952).

DISCUSSION

LIQUORI

We have used the matrix method proposed by you and SHIMANOUCHI [1] in order to carry out a semiquantitative conformational analysis on the stability of simple helical polymers [2].

The polymers which have been considered, contain either all chemically equivalent bonds (« one atom chain ») or pairs of chemically non equivalent bonds in the skeleton (« two atom chain »). Of course, polypeptide chains contain triplets of chemically non equivalent bonds in the backbone chain and should formally be considered as « three atom chain ».

Using the matrix methods in conjunction with a vector method proposed by HUGHES and LAUER [3] the cylindrical coordinates of all the atoms of the helix have been expressed as a function of one or two angles of rotation around the skeleton bonds for one and two atom chain helices respectively.

---

[1] T. SHIMANOUCHI and S. MIZUSHIMA, *J. Chem. Phys.*, *23*, 707 (1955).

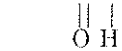
[2] F. DE SANTIS, E. GIGLIO, A.M. LIQUORI and A. RIPAMONTI, *International Symposium on Machomolecular Chemistry*, Montreal, 1961.

[3] R.E. HUGHES and J.L. LAUER, *J. Chem. Phys.*, *30*, 1565 (1959).

Choosing suitable potential energy function (4) to describe VAN DER WAALS interactions between non directly bonded atoms in the chain we have calculated with the aid of an electronic computer the conformational potential energy of the infinite helix as a function of the monomer repeat and the number of monomers per turn.

The two-dimensional energy diagrams as a function of the angles of rotation obtained for two atom chain helices such as polyisobutylene or isotactic polypropylene and the one dimensional diagram obtained for one atom helices such as Teflon contains several minima. It is very interesting that the deepest minima correspond to the helices found in the crystallized polymers by X-ray diffraction (5).

This method therefore appears thus far very successful for predicting the most stable helix to be expected for a given chain polymer. We are now attempting to extend this approach to polypeptide chains which may be treated for this purpose as a two atom chain since the angle of rotation around the -C-N- bond can be set equal to zero or to π.



One aim of such a calculation is to investigate whether there are other subsidiary minima in the conformational potential energy plot of an infinite helix besides that corresponding to the α-helix. Furthermore, we will try to carry out a systematic conformational analysis of polypeptide chains with a given sequence with the hope of obtaining some information about the influence of VAN DER WAALS interactions between non bonded atoms on the tertiary structure. Of course, the main difficulties we are trying to overcome are those connected with the choice of a suitable potential function to describe the hydrogen bonding.

(4) E.A. MASON and M.M. KREVOY, *J. Am. Chem. Soc.*, 77, 5808 (1955).
(5) A.M. LIQUORI, *Acta Cryst.*, *8*, 345 (1955); G. NATTA and P. CORRADINI, *J. Polymer Sci.*, *20*, 251 (1956); C.W. BUNN, *Nature*, *174*, 549 (1954).

MIZUSHIMA

It is interesting that the calculated VAN DER WAALS interactions between non-bonded atoms in the chain well explain the experimental results of polyisobutylene, etc. This means that in such a case the VAN DER WAALS force becomes much more significant as compared with the other forces. However, this is not always the case and for polyvinyl chloride the electrostatic interactions become also important (6). I think that for polypeptide chains internal hydrogen bond must be taken into account in addition to the electrostatic and the VAN DER WAALS forces.

E. KATCHALSKI

I am very much interested in the possible conformation of low and high molecular weight polypeptides containing prolylprolyl links. Poly-L-proline seems to exhibit conformations very different that those of other polyamino acids, and optical rotation data indicate that a considerable number of conformations may be obtained from poly-L-proline in solution under the suitable conditions.

It is perhaps pertinent to note that we have recently observed that the copolymerization of the N-carboxy anhydride (NCA) of L-proline with that of other amino acids leads as a rule to the formation of block copolymers containing shorter or longer polyproline block. Random copolymers are obtained only on copolymerizing the NCA of proline with the NCAs of O-acetyl-hydroxyproline or glycine. Polyglycine is known to attain form II similar in conformation to that of polyproline II.

LIQUORI

With reference to the optical rotation of polypeptide chains I would like to observe that you have shown (1) that an α-helix can be characterized by the repetition of a « conformational triplet » TGG, being T, G and G the equilibrium values of the angles of

(6) See SHIMANOUCHI and TASUMI, *Spectrochim. Acta*, *17*, 755 (1961).
(1) S. MIZUSHIMA, *Adv. Protein Chem.*, *9*, 299 (1954).

rotation ($T \sim 180°$ $G \sim 60°$) around the three consecutive skeleton bonds of an aminoacid residue.

Each monomer unit in a α-helix is thus frozen in an asimmetric conformation which should accordingly be associated with a large optical rotation apart from specific effects due to resonance exciton interactions of the transitions of the chromoforic electrons of the amide groups as implied in MOFFITT's theory (2).

A similar situation is encountered in the study of optically active vinyl polymers synthesized in our laboratory (3) where the unequal population of conformational isomers of a monomer unit is responsable for large exhaltations of the optical rotation.

The above type of analysis might prove suitable to explain in terms of conformation the main contribution to the optical rotation of polypeptide chains in a relatively simple fashion following for instance the approach outlined by KAUZMANN and others (4).

PUTZEYS

May I ask a question? Does that mean that in the MOFFITT equation b would be equal to 0?

LIQUORI

As commonly used, MOFFITT's treatment of the optical rotation data of polypeptides includes both conformational and specific electronic effects. I am suggesting that it might be possible to separate the conformational effects and that these might turn out to be dominant.

(2) W. MOFFITT, *Proceedings National Acad. Sci.*, *42*, 736 (1956).

(3) F. ASCOLI, V. CRESCENZI, M. D'ALAGNI, A.M. LIQUORI, *Int. Symposium of Macromolecular Chemistry*, Montreal (1961).

(4) W. KAUZMANN, F.B. CLOUGH and I. TOBIAS, *Tetrahedron*, *13*, 57 (1961).

RELATION BETWEEN STRUCTURE AND SEQUENCE OF HAEMOGLOBIN

M.F. PERUTZ

Medical Research Council - Laboratory for Molecular Biology
Cambridge (England)

When the tertiary structure of horse haemoglobin was first published, its amino acid sequence was still unknown, and the resolution of the X-ray analysis was too low to allow side chains to be seen [1]. Since then, the amino acid sequences in the two different chains of human haemoglobin and in one of the chains of horse haemoglobin have been elucidated [2], [3], [4], [5]. In addition, most of the structure of sperm whale myoglobin is now known in atomic detail [6]. Thanks to the close similarity in tertiary structure between sperm whale myoglobin and horse haemoglobin the position of almost any residue of a given sequence number can be found by looking up the co-ordinates of the corresponding residue in myoglobin and the transferring those co-ordinates to the model of haemoglobin. This procedure can be checked, because the positions of certain residues can be found directly in the electron density maps of haemoglobin, despite the low resolution, and can then be compared with those inferred by analogy with myoglobin. For instance, the cysteines are found by labelling them with mercurials; the indol rings of tryptophans are recognizable as bulges of high electron density protruding from the main chain;

by contrast glycine in non-helical regions sometimes shows up by causing a marked constriction in the roughly cylindrical cloud of high electron density which represents the polypeptide chain.

A year ago, WATSON and KENDREW published a preliminary comparison of the structure and sequence of haemoglobin, but at that stage only parts of the chemical sequence were known [7]. In the light of the further knowledge now available, many interesting new features reveal themselves.

In the discussion which follows the tertiary structure of equine and human haemoglobin will be assumed to resemble each other closely. This has not yet been proved, but is highly probable in view of the similarity in amino acid sequence of these two haemoglobins and in view of the correspondence between the tertiary structures of the different species of myoglobin and haemoglobin so far examined [9], [10].

CHEMICAL IDENTIFICATION OF CHAINS IN CRYSTALLOGRAPHIC MODEL

The two different chains of horse haemoglobin have the N-terminal sequences valyl-leucyl and valyl-glutaminyl, and are similar in composition to the valyl-leucyl and valyl-histidyl-leucyl chains of human haemoglobin [4], [5], which have become known respectively as the α- and β-chains. SMITH and PERUTZ have shown the valyl-glutaminyl chain to correspond to the black chain of the crystallographic model [8], from which it would follow that the black chain also corresponds to the β chain of human haemoglobin. We therefore have the following relations:

	N-terminal sequences		*crystallographic model*
	horse	*human*	
α-chain	valyl-leucyl	valyl-leucyl	White chain
β-chain	valyl-glutaminyl	valyl-histidyl-leucyl	black chain

Table I lists those parts of the sequences of human and equine haemoglobin which are discussed in this paper, while figure 1 shows the tertiary structures of the α- and β- chains with the positions of certain residues superimposed. The sequences of the α- and β-chains are similar, but to bring them into register all along their length appropriate gaps have to be left in certain positions. Thus a gap of one and another of five residues have to be assumed between residues no. 46 and 51 in the α-chain in order to keep its sequence in register with that of the β-chain. These gaps evidently correspond to a shortening of the α-chain which should be noticeable in the model. By looking up the positions of the corresponding residue in myoglobin the shortening can be traced to the non-helical region CD and the helical region D which lie in the loop at the top of the models in figure 1. It will be noted that in the white chain this loop is cut short compared to the black chain, showing that the shortening of the α-sequence as its corollary in the electron density maps.

Co-ordination of haem groups

The haem-linked residue in myoglobin has been identified as histidine. The α- and β-sequences each show a histidine in the corresponding position, thus confirming that histidine is the haem-linked residue in haemoglobin, as had long been predicted from physico-chemical data [12], [13], [14]. Kendrew and Watson have pointed out that a second histidine occurs on the distal side of the haem group in all three chains;

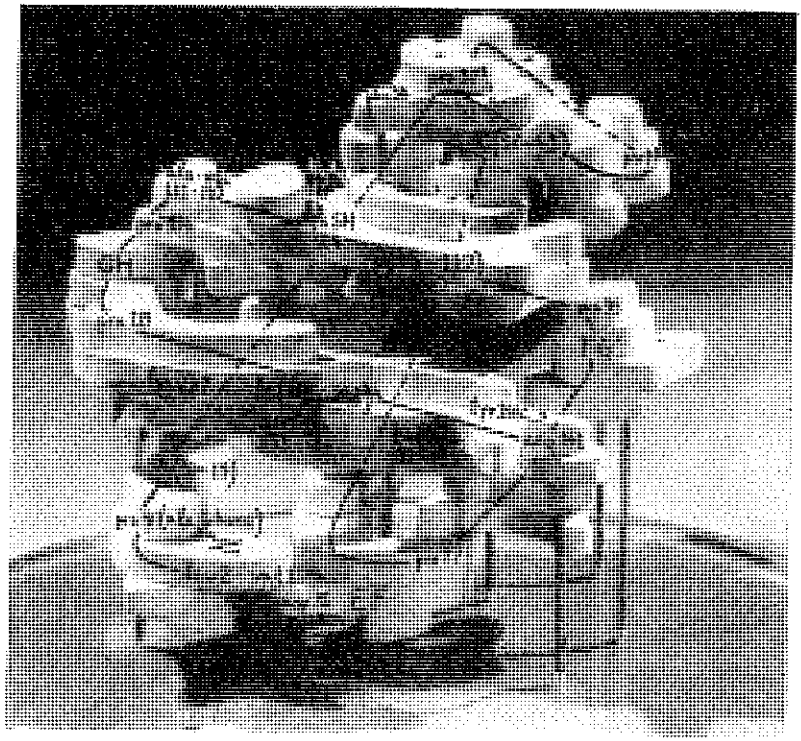


FIG. 1-*a* — A model of the cloud of high electron density representing the α-chain of horse haemoglobin. The letters A, C, G and H refer to various regions of chain identified as right-handed α-helices in myoglobin. The numbers in brackets give the sequential residue numbers associated with them. EF, FG, GH refer to corners or non-helical regions. $HgAc_2$ denotes a mercuric acetate group used to label cysteine 104_α. Except for proline 4_α all the residues marked in this figure occupy the same sequential positions in human and equine haemoglobin.

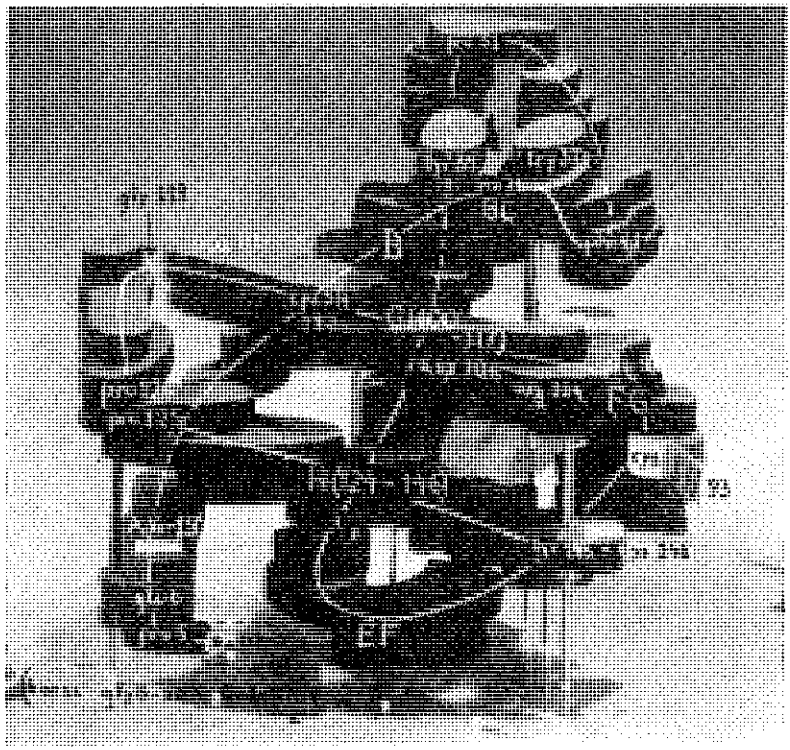


FIG. 1-*b* — The β-chain. The residues marked on the model are taken from the sequence of human haemoglobin, except for the N-terminal sequence, where the data for equine haemoglobin are also listed. Cysteine 112 is known to be replaced by another residue in equine haemoglobin; there may be some other differences between the two haemoglobins which affect this figure and which are not yet known. Note the position of glutamic acid 6, the residue which is replaced by valine in sickle cell haemoglobin, also the bulges marking the positions of tyrosine 35, trptophan 37 and of phenylalanines 41 and 42, and the gap which marks the position of glycine 119 in the non-helical region GH.

The haem group lies at the back of the models.

this histidine is not in contact with the iron atom, but is within hydrogen-bonding distance of the haem-linked water molecule, and is the residue which is replaced by tyrosine in two of the forms of haemoglobin M [15]. This may be the second histidine in close proximity to the iron atom, but not directly linked to it, whose presence has long been surmised [11], [12], [13], [16], [17]. It is most satisfactory to find these expectations confirmed by crystallographic analysis.

Position of cysteines

All vertebrate haemoglobins appear to contain at least one pair of cysteines whose sulphydryl groups are reactive in the native protein. Their positions can be determined by X-ray analysis, using mercurials as a label. From the X-ray results, combined with model building, it was possible to show that in horse haemoglobin the reactive cysteine is the residue immediately following the haem-linked histidine in the β-chain, and a cysteine in this position has since been found by chemical sequence work in human haemoglobin. The histidine and cysteine are part of an α-helix, and the side chain of cysteine points in the direction away from the iron atom.

Horse haemoglobin contains another pair of cysteines which is unreactive in the native protein [18], [19], but can be labelled with mercurials under appropriate conditions. This is attached to the helical regions G of the α-chains, and lies tucked away in the interior of the molecule. The position deduced by labelling with mercury corresponds exactly to that inferred for residue no. 104_{α}, the sequential position of the corresponding cysteine in human haemoglobin, by analogy with the myoglobin structure.

Adult human haemoglobin, but not the foetal form, contains a third pair of cysteines in position 112_{β}, whose position in the model is similar to that of 104_{α} in that it is part of the

[15] *Perutz* - pag. 6

helical region G and lies in the interior of the molecule. This pair is also unreactive in the native protein [19], [20]. Neither of the unreactive pairs has any contact with the haem group, nor are they in contact with each other or with residues of any chain other than their own. Thus their function is obscure. Their lack of reactivity in the native protein can be attributed to their situation in the interior of the molecule, though covering of the sulphydryl groups by their immediate neighbours may possibly be an additional factor.

Tryptophan

As has been mentioned above, the indol groups of tryptophan are recognizable as marked bulges of high electron density protruding from the main polypeptide chain, even at a resolution of only 5.5 Å. Tryptophan 14_α and 15_β appear in the exact positions occupied by trptophan 12 in myoglobin and are part of the helical region A. The indol ring lies in the interior of the molecule, sandwiched between helices A and E; it is parallel to the porphyrin ring, but separated from it by helix E, and lies at a distance of 14 Å from the iron atom. Since this is the only tryptophan which all chains of haemoglobin and myoglobin so far analysed have in common, it is probably the residue responsible for the photodissociation observed when carbonmonoxyhaemoglobin is illuminated at the ultraviolet absorption band of tryptophan. In this process a quantum must be transferred from the indol ring, across helix E, to the porphyrin ring. However, the photodissociation of carbon monoxide is unlikely to be of biological significance, and the common occurrence of tryptophan in this position suggests that it may actually play a part in the oxygen-combining activity in a manner which we cannot yet guess. The β-chain of human haemoglobin contains another tryptophan in position 38. This is also present in the slow component of horse

haemoglobin, but in the fast component it is replaced by lysine [21]. Myoglobin contains a second tryptophan in position 5 which is absent in haemoglobin. Evidently these tryptophans are not essential.

Prolines and corner turning

The two chains of human haemoglobin each contain seven prolines which all lie in corners or non-helical regions of the chain, as had been suggested by Watson and Kendrew [7]. However, there are only three positions (the corner CD, the N-terminal ends of helic G and the corner GH) where prolines occur both in myoglobin and in the α- and β-chains. The fourth proline of myoglobin occurs in a position where the haemoglobin chains have none, and curiously enough, the prolines which occur in each of the chains of haemoglobin do not all occur at the same corners. For instance, in the neighbourhood of the short helical region D the α-chain has no proline and the β-chain has two. On the other hand, the α-chain has a proline near the C-terminal end of the region C where the β-chain has none. Admittedly the regions C, CD, and D are the ones where the two chains also differ most markedly in structure, but there is also proline 77_α which occurs in the non-helical region EF and has no counterpart in the β-chain. In the γ-chain of human foetal haemoglobin which replaces the β-chain in the adult form, the prolines in position 5 and 51 are replaced by glutamic acid and alanine respectively [22]. The two chains of horse haemoglobin contain only six prolines each, and differ from the human ones by the absence of prolines 4_α and 5_β at the N-terminal end of helix A; these are replaced respectively by alanine and glycine. These results suggest that the corner-turning mechanism which depends on the presence of proline is readily replaced by others, and is therefore not of great evolutionary importance.

[15] *Perutz* - pag. 8

The correspondence between structure and sequence found in this work provides an independent check for the interpretation of the X-ray results. If the meaning of the electron density maps of either haemoglobin or myoglobin had been misinterpreted, the correspondence between the observed and predicted positions of cysteine and other residues could not have been found.

A full account of this work, which also gives a complete description of the methods and results of the X-ray analysis has been published by CULLIS, MUIRHEAD, NORTH, PERUTZ and ROSSMANN [23].

REFERENCES

[1] PERUTZ M.F., ROSSMANN M.G., CULLIS A.F., MUIRHEAD H., WILL G. and NORTH A.C.T., *Nature, 185,* 416 (1960).

[2] BRAUNITZER G., GEHRING-MULLER R., HILSCHMANN N., HILSE K., HOBOM G., RUDLOFF V. and WITTMANN-LIEBOLD B., *Hoppe-Seyl. Z., 325,* 283 (1961).

[3] KONIGSBERG W., GUIDOTTI G. and HILL R.J., *J. Biol. Chem., 236,* PC55 (1961).

[4] BRAUNITZER G. and MATSUDA G., *Hoppe-Seyl. Z., 324,* 91 (1961).

[5] SMITH D.B., *Private Communication.*

[6] KENDREW J.C., WATSON H.C., STRANDBERG B.E., DICKERSON R.E., PHILLIPS D.C. and SHORE V.C., *Nature, 190,* 663 (1961).

[7] WATSON H.C. and KENDREW J.C., *Nature, 190,* 663 (1961).

[8] SMITH D.B. and PERUTZ M.F., *Nature, 188,* 406 (1960).

[9] SCOULOUDI H., *Proc. Roy. Soc., A, 258,* 181 (1960).

[10] NORTH A.C.T. and GREEN D.W., *Unpublished results.*

[11] CONANT J.B., *Harvey Lectures, 28,* 159 (1932).

[12] WYMAN J., *J. Biol. Chem., 127,* 581 (1939).

[13] CORYELL C.D. and PAULING L., *J. Biol. Chem., 132,* 769 (1940).

[14] KEILIN J., *Nature, 187,* 365 (1960).

[15] GERALD P.S. and EFRON M.L., *Proc. Nat. Acad. Sci. U.S., 47,* 1758 (1961).

[16] HAUROWITZ F., *J. Biol. Chem., 193,* 443 (1951).

[17] KEILIN D. and HARTREE E.F., *Nature, 170,* 161 (1952).

[18] INGRAM V.M., *Biochem. J., 59,* 653 (1955).

[19] ALLISON A.C. and CECIL R., *Biochem. J., 69,* 27 (1958).

[20] CECIL R. and SNOW N.S., *Biochem. J., 82,* 247 and 255 (1962).

[21] BRAUNITZER G., *Private communication.*

[22] SCHROEDER W.A., SHELTON J.R., SHELTON J.B. and CORMICK J., *Proc. Nat. Acad. Sci. U.S., 48,* 284 (1962).

[23] CULLIS A.F., MUIRHEAD H., NORTH A.C.T., PERUTZ M.F. and ROSSMANN M.G., *Proc. Roy. Soc. A, 265,* 15 and 187 (1961-62).

[15] *Perutz* - pag.10

TABLE I

Abbreviated amino acid sequences of various haemoglobin chains, illustrating points mentioned in the text. For the α-chain of horse haemoglobin only those residues are shown which differ from the corresponding ones in human haemoglobin. In the β-chain of horse haemoglobin only the N-terminal sequence is shown which differs from human haemoglobin in three places. The vertical lines indicate tryptic peptides whose composition is either the same as that of the corresponding peptide of human haemoglobin or differs from it by only one amino acid residue. In the γ-chain only those residues are shown which differ from the corresponding ones of the β-chain or (bracketed) where only the composition, but not the sequence of the peptide is known.

human α		horse α	human β		horse β	human γ	Remarks
no.	residue		no.	residue			
			1	val		gly	These residues form a hook at the N-terminal and which has no counter-part in myoglobin.
1	val		2	his	gluN		
2	leu		3	leu		phe	
3	ser		4	thr	ser		
4	pro	ala	5	pro	gly	glu	
5	ala		6	glu			glu 6β is replaced by valine in haemoglobin S.
•	•	•	•	•			
14	try		14	try			tryptophan responsible for photodissociation effect.
15	gly		15	gly			
•	•		•	•			
36	phe		35	tyr			
37	pro		36	pro			proline in CD corner
38	thr		37	try			tryptophan replaced by lysine in « fast » horse haemoglobin.
•	•		•	•			
44	pro		43	glu		asp	
45	his		44	ser			
46	phe		45	phe			
|	|		46	gly			α-chain shortened by one residue
47	asp		47	asp			
48	leu		48	leu			
49	ser		49	ser			
50	his		50	thr		ser	
|	|		51	pro		ala	proline at N-terminal end of helix D in β-chain
|	|		52	asp		ser	
|	|		53	ala			α-chain shortened by five residues.
|	|		54	val		ileu	
|	|		55	met			
51	gly		56	gly			
52	ser		57	asp			
53	ala		58	pro			proline in corner DE in β-chain
54	gluN		59	lys			
55	val		60	val			
56	lys		61	lys			
57	gly	ala	62	ala			
58	his		63	his			histidine on distal side of haem
59	gly		64	gly			glycine essential for short contact between helices E and B
60	lys		65	lys			
61	lys		66	lys			three lysines in neighbourhood of distal histidine forming BRAUNITZER's « basic centre »
•	•		•	•			
76	met	leu	81	leu			
77	pro		82	lys			proline in non-helical region EF of α-chain
78	aspN	gly	83	gly			
•	•		•	•			
86	leu		91	leu			
87	his		92	his		his?	haem linked histidine
88	ala		93	cysH		cysH?	reactive cysteine in β-chain
89	his		94	asp		asp?	
•	•		•	•			
95	pro		100	pro			prolins in corner FG
•	•		•	•			
104	cysH		109	val		val?	unreactive cysteine in α-chain
•	•		•	•			
107	val	ser	112	cysH	not cys H	thr	unreactive cysteine in human β-chain
•	•		•	•			
114	pro		119	gly			proline in non-helical region GH
115	ala	asp	120	lys			
116	glu	asp	121	glu			
117	phe		122	phe			
118	thr		123	thr			
119	pro		124	pro			prolines at C-terminal end of helix H
120	ala		125	pro		glu	
•	•		•	•			
139	lys		144	lys		arg	
140	tyr		145	tyr			tyrosine forming hydrogen bond with main chain carbonyl in FG corner
141	arg		146	his			

DISCUSSION (*)

THEORELL

Well, I think we all agree that these two lectures have been highly fascinating bringing up such a tremendous lot of data, solving so many problems which we would scarcely have dreamt of seeing cleared up.

Now, as is well known, the myoglobins from all the animals investigated so far, show a certain degree of complexity or microheterogeneity. For instance, the horse myoglobin always shows the same pattern so far as we can have seen, which consists of the medium component making out some 65-70% of the total, a second one which may come to some 15-20% and another fraction making up to something like 10%. This does not seem to be due to the mode of preparation. We have put some efforts into trying to find out what might be the difference between these. As far as we have been able to see from our amino acid analysis, they seem to be identical in total amino acid composition, and the total amount of amide groups is the same in all three. Nevertheless they can be separated electrophoretically without too much effort. So, we asked ourselves, where does the difference come from? We made some experiments with INGRAM fingerprints and the expected number of spots came out (something between 20 and 25) and most of them were identical for the three myoglobins. But there were some small differences. One or two extra spots appeared. Now, how can such a thing be explained?

(*) This discussion refers also to Prof. KENDREW's lecture.

Since the tryptic degradations were made on denatured material, I would not expect alternative ternary structures to be the cause, and one hesitates to think that the same total amount of amino acids could be rearranged and change position one against the other. That could give rise to differences in the INGRAM fingerprints. Another more likely possibility is that the total amino acids and their sequence are the same in all three components, but that acetylation at different amino acid residues occurs in analogy with the presence of N-acetyl glycine in cytochrome *c* (MARGOLIASH and SMITH). This item is now investigated in collaboration with Å. ÅKESON.

There are so many questions that perhaps I should not take them all at the same time, but let me at once bring up one major problem, which I would very much like to hear Dr. KENDREW's and Dr. PERUTZ's opinion about. Of course, what makes hemoglobin and myoglobin function physiologically is chemically very difficult to understand, because if you make an imidazole hemochromogen and you add oxygen to it, oxygen will inevitably oxidize the ferrous ion to ferric. Of course, this would be disastrous for hemoglobin and myoglobin. That does not happen and I wonder why. Could that perhaps be connected with this second histidine which is not really coordinated but sitting a little bit apart? By the way, it is interesting to remember that old paper by PAULING and CORYELL, from 1940 I believe, where they concluded from magnetic and titrimetric data that on one side of the haemin disc there would be a histidine imidazole strongly bound to the iron and on the other side another histidine residue in an unfavorable position for coordinating with iron. Now both histidines are there in the predicted positions; very beautiful indeed: I wonder, could this arrangement be correlated with the very remarkable property of iron not oxidized by oxygen?

PERUTZ

J.H. WANG has shown that ferrous haem can combine with molecular oxygen, without becoming oxidized, if it is embedded in a non-polar medium. He discovered this before the situation of the

haem group in myoglobin and haemoglobin became known, and his prediction has been verified by the results of KENDREW and myself. WANG suggests that the non-polar medium supresses the ionization of the oxygen which would be required to oxidize the iron atom. As to the second histidine residue, I wonder if it could form a hydrogen bond with the oxygen molecule.

KENDREW

We have been aware for a long time of the inhomogeneity of myoglobin, as normally prepared. The two major components of sperm-whale myoglobin have been compared chemically by Dr. A. EDMUNDSON, and the only difference he can find is a single amide group, present in the one and absent in the other. We have crystallized these components separately; the crystals, and their X-ray patterns, are almost identical with one another and with crystals prepared from the mixed components.

TISELIUS

For future structure determinations on proteins by X-ray analysis, it would be interesting to hear from Drs. PERUTZ and KENDREW to what an extent the experiences made may facilitate the work and make it less time-consuming. If you have, for example, two closely related proteins, and if you know the structure of one of them, what hope will this furnish for the solution of the structure of the other one? Is there any possibility of a « differential » method?

KENDREW

Dr. HELEN SCOLOUDI, in her studies of seal myoglobin, has calculated a 2-dimensional projection of the unit cell which, like other projections of protein crystals, is uninterpretable as it stands. By comparing it with the known 3-dimensional structure of sperm-whale myoglobin, however, she was able to show that the two molecules were nearly identical at 6 A. resolution; this comparison was made by trial-and-error methods; it seems likely that a more

formal procedure will make it possible to solve related structures given a prior solution of one of them, provided the structures are rather similar.

PERUTZ

In haemoglobin GREEN and NORTH performed a similar experiment. They showed that electron density projections of ox haemoglobin were explicable on the basis of the three-dimensional electron density distribution of horse haemoglobin projected in certain directions.

JACOB

Does it not seem reasonable from your results on the compared structures of hemoglobin and myoglobin to conclude that most of the mutations, which result in a change of non essential amino-acids, will not affect deeply the structure and function of a protein and therefore will not be detectable? On the other hand, most of the mutations resulting in a change of the key amino-acids will bring about a loss or a decrease of function. These would correspond to the so-called hot spots observed in mutational analysis.

PERUTZ

I agree entirely with Dr. JACOB's remark.

FRAENKEL-CONRAT

In regard to the problem of the origin of protein conformations, I would like to draw attention to the renatured TMV protein which was first described by ANDERER and which we have studied quite a bit: apparently a single chain protein, even one lacking disulfide bonds, can and will after denaturation return to what appears to be its original tridimensional chain conformation. If one cannot yet predict the shape of the protein from its amino acid sequence, it would seem that this will in the future become possible. All one would have to do is to evaluate all possible interact-

ing bond strengths and calculate from such data the thermodynamically most stable conformation based on the greatest number of interactions. But the specific question I wanted to ask was the following: « There has come from study of the synthetic polyamino-acids the impression that certain residues (valine, threonine) will not fit into helical regions, and I wondered in how far the extensive helical regions in myoglobin are free from amino acids which are not supposed to be there? ».

PERUTZ and KENDREW

Both in myoglobin and in haemoglobin threonine, valine and serine occur in helical as well as in non-helical regions. Serine and threonine appear to be useful in stabilizing non-helical regions, but they do not inhibit helix formation. On steric grounds one would only expect valine or threonine to inhibit helix-formation if pairs of these residues lie four apart in the sequence; this situation does not occur in the helical segments of myoglobin.

PUTZEYS

There is no necessity to assume a change in oxygenation to explain a change in crystal structure. Carbonylhemoglobin of the horse shows two kinds of crystals, thin needles and rhomboidal plates (1). According to GREEN the last ones are formed in the neighbourhood of pH 6.6, the former ones at more alkaline reactions. We have found that needles are formed at very low ionic strength, while plates appear by salting out with ammonium sulfate. Furthermore, when the solid phase is equilibrated with a solvent as in solubility measurements, the needles are the stable phase at low ionic strength (Na_2SO_4) and these change to rhomboidal plates at an ionic strength of 1.4 or higher. Such a transition point appears to indicate a change in crystal structure rather than a mere change of habit. It may be noted that the solutions are kept saturated with pure carbon monoxide all the time and that the change is

(1) A.A. GREEN, *J. Biol. Chem.*, *93*, 495 (1931).

henceforth not due to a dissociation of the carbonylhemoglobin. On another point I would like to have information. If thiols are buried inside the hemoglobin structure, how is it possible to combine mercury with them?

PERUTZ

Many proteins, including haemoglobin, are polymorphic. A change in crystal structure can often be brought about by a change in the pH of the crystallization medium, and does not necessarily imply a change in molecular structure.

However reduced haemoglobin *always* crystallizes in forms which are quite different from those of oxy- carboxy- or methaemoglobin, and in this case a change in molecular structure occurs (MUIRHEAD and PERUTZ, unpublished).

It is sometimes possible to combine mercury with sulphydryl groups which are buried in the interior by using methods which tend to *open up* the molecule, such as high temperature, very low ionic strength or high pH.

LIAISONS STABILISANT LA STRUCTURE TERTIAIRE DES PROTEINES

PAUL PUTZEYS

Laboratoire de Biochimie, Université de Louvain, Louvain (Belgique)

La complexité des protéines et leur principal intérêt dépend fondamentalement du fait qu'elles sont composées d'atomes liés par des forces de puissance très différentes. Cette phrase, légèrement modifiée, empruntée à BERNAL [1] peut servir d'introduction à cet exposé consacré aux liaisons stabilisant la structure tertiaire des protéines. Depuis que l'on connaît la structure tertiaire de la myoglobine et de l'hémoglobine par les travaux de KENDREW [2, 3, 4, 5, 6], de PERUTZ [7] et leurs collaborateurs, il ne peut y avoir de doute que la conformation exceptionnelle et unique des polypeptides qui constituent ces protéines, ne saurait être maintenue que grâce à la collaboration de nombreuses liaisons de caractères très différents.

Cette conclusion s'était d'ailleurs imposée avant même que les rayons X n'aient démontré la réalité d'une structure que l'induction à partir des observations faisait prévoir, sans cependant pouvoir la préciser.

Il peut être utile de rappeler comment on y est arrivé. La structure fondamentale d'une protéine est une molécule en chaîne liée par des covalences, le polypeptide. Tout comme dans les molécules simples, les covalences ne saturent pas les

affinités résiduelles du polypeptide. Celles-ci donnent lieu à la formation de liaisons généralement faibles par rapport aux liaisons covalentes et leur influence sur les propriétés des molécules simples peut être considérée comme secondaire. Il n'en est pas de même dans le cas des polypeptides où leur influence se trouve amplifiée par accumulation, au point de devenir prépondérante. Ces affinités résiduelles cherchent à se saturer par la formation de laisons soit intramoléculaires, soit intermoléculaires, soit par la formation de liaisons avec les molécules d'un solvant approprié. On peut imaginer un solvant, telle une solution concentrée d'urée, dans lequel toutes les liaisons intra- et intermoléculaires auraient été rompues au profit de liaisons avec le solvant. (Le solvant est l'ensemble des substances présentes dans le système, le polypeptide excepté). Nous dirons alors que les polypeptides sont « libres ». Dans l'eau pure ou les solutions aqueuses diluées, les polypeptides ne restent pas libres; ils s'agrègent et finissent par précipiter. Ne faundrait-il en conclure que les affinités résiduelles ont tendance à se saturer par liaisons intermoléculaires plutôt que par liaisons avec l'eau?

Or les protéines au sens strict (protéines globulaires) sont solubles dans l'eau, elles sont homodispersées en solution comme SVEDBERG a eu le grand mérite de le démontrer à l'aide de son ultracentrifuge; elles ont de ce fait un poids moléculaire défini, elles sont relargables, cristallisables dans les conditions les plus variées de température, d'acidité, de force ionique, sans modification de leur poids moléculaire, toutes propriétés incompatibles avec celles de polypeptides libres. Force nous est de conclure que, le polypeptide constitutif d'une protéine a saturé la plus large part de ses affinités résiduelles par liaisons intramoléculaires. Ceci implique une conformation exceptionnelle et unique où la chaîne se trouve repliée sur elle-même et tend vers une forme compacte et sphérique.

C'est cette conformation existant en solution aqueuse que les rayons X retrouvent dans les cristaux de myoglobine et d'hémoglobine.

J'ai employée la dénomination « structure tertiaire » et cet usage m'oblige, avant de poursuivre, d'ouvrir une parenthèse concernant la nomenclature. Récemment WETLAUFER [8] a proposé d'abandonner les désignations: structure primaire, secondaire, tertiaire. Il a raison de dire que la structure d'un polypeptide est toujours complètement définie par deux données: la séquence des amino-acides et la conformation de la chaîne. Sans entrer dans une analyse et une discussion approfondie de ces propositions, constatons que la conformation d'un polypeptide peut présenter des variétés indéfiniment nombreuses et que nous n'avons pas de moyen pour indiquer la conformation dont il s'agit. Or la conformation d'un polypeptide dans une protéine est unique; « Structure tertiaire » désigne cette conformation exceptionnelle et en ce sens ne prête pas à confusion. La désignation « structure tertiaire » peut s'appliquer encore lorsqu'une molécule de protéine est formée de plusieurs polypeptides. Deux cas sont à considérer, soit que chaque polypeptide forme une sous-molécule ayant sa propre structure tertiaire comme dans le cas de l'hémoglobine, soit que les différents polypeptides s'entortillent en une seule structure tertiaire, s'étendant à tout la molécule.

Ce dernier cas doit être rare et je n'en connais pas d'exemple. On peut au contraire signaler de nombreux exemples du premier cas qui semble bien être la règle. Nous reviendrons plus loin sur cette question. En tout cas la désignation de structure quaternaire, comme l'a fait BERNAL [1] est superflue.

* * *

Les affinités résiduelles dues à la liaison peptide sont communes à tous les polypeptides quels que soient le genre et le nombre des amino-acides qui les forment. La saturation de

ces affinités par ponts hydrogène *intra*moléculaires est réalisée de façon parfaite par la conformation en hélice alpha. Supposons un polypeptide complètement sous cette forme. Les groupes latéraux sont alors à la surface du cylindre ainsi formé. En présence d'eau, cette conformation est instable. On avait pensé qu'elle se trouvait réalisée dans la fibroïne soluble dans l'eau (AMBROSE, BAMFORD, ELLIOTT, HANBYI [9] sur la foi de l'absorption à la fréquence 1660 cm.$^{-1}$ dans les films d'évaporation à froid. Mais il fut reconnu plus tard que cette fréquence n'est pas nécessairement associée à la présence d'hélices alpha (BAMFORD, ELLIOTT et HANBY) [10]. BAMFORD et al. concluent que la conformation de la fibroïne soluble ne nous est pas connue de façon précise, mais qu'il est probable que la plupart des ponts hydrogène ne contribuent pas à la formation de liaisons intercaténaires. La seconde proposition me semble pour le moins exagérée (ibid., p. 389). En effet les propriétés de la fibroïne en solution feraient plutôt penser le contraire. Il ne faut pas interprèter le mot « solution » comme signifiant moléculairement dispersé, lorsqu'il s'agit d'une solution aqueuse d'un polypeptide. Lorsqu'on se donne la peine de déterminer le poids moléculaire moyen, de préférence par diffusion moléculaire de la lumière, on constate que l'on est loin d'une dispersion moléculaire. Des expériences récentes nous ont montré que des solutions de fibroïne obtenues par dissolution dans la cupri-éthylène diamine et dialysées ensuite à fond contre de l'eau distillée, étaient en fait des suspensions colloïdales à micelles énormes, nonobstant leur apparente limpidité.

Comme l'a montré LONTIE [11], une chose semblable se produit pour les polypeptides de l'édestine. Après dénaturation en milieu acide, redissolution du précipité par l'acide chlorhydrique 0,1 m et dialyse, on obtient au bout de quelques jours une solution limpide dont le pH est voisin de 5. Le poids moléculaire moyen de ces solutions, très pauvres en électrolytes, peut varier de 700.000 à plus de 2 millions. Tout dépend

de l'acidité et de l'âge de la solution. Il est évident que ces solutions sont fortement agrégées. L'addition de petites quantités d'électrolyte augmente l'agrégation jusqu'à donner des précipités. Cette agrégation considérable implique nécessairement de nombreuses liaisons intermoléculaires dont un certain nombre doit être du type bêta, car en acidifiant de nouveau, on ne revient pas à l'état d'agrégation premier. La tendance à la formation des ponts hydrogène du type bêta est d'ailleurs très marquée puisqu'il suffit de l'aider mécaniquement pour la voir se réaliser très largement. C'est la conformation stable vers laquelle tend un polypeptide libéré en solution aqueuse; par conséquent, la conformation en hélice alpha est métastable par rapport à la première dans ces mêmes conditions.

Les travailleurs du laboratoire Carlsberg, sous la direction de LINDERSTROM-LANG, ont d'ailleurs apporté un ensemble de preuves théoriques et expérimentales de cette instabilité de l'hélice alpha en milieu aqueux [12, 13, 14]. Or l'hélice alpha est bel et bien présente dans les molécules des protéines. Il fallait s'y attendre puisque la possibilité d'existence de ces molécules implique, comme nous l'avons vu, la saturation intramoléculaire du plus grand nombre des affinités résiduelles, et il n'y-a sans doute pas de moyen plus radical d'attendre ce but, en ce qui concerne les affinités résiduelles de l'épine dorsale, que la formation de l'hélice α. Si le polypeptide n'est pas complètement sous cette forme, c'est qu'il y a d'autres affinités résiduelles à satisfaire encore, à part celles de l'épine dorsale.

Nous avons deux genres de preuves de la présence de l'hélice alpha dans la molécule de protéine.

En tout premier lieu les résultats de KENDREW et PERUTZ. Citons KENDREW [3]: « In the 2-Å synthesis all the straight segments of chain were found to be hollow cylindrical tubes of high density; projecting at intervals from the cylindrical core are dense regions of various shapes and sizes, which are clearly the amino acid sidechains. More detailed examination shows that, in fact, the cylindrical tubes are helices, consisting of a

single strand of high density with axial repeat about 5.4 Å.... the observed density follows the configuration of the alpha helix with remarkable precision. It is also found that the side-chains emerge at intervals of 1.5 Å, in conformity with the parameters of the alpha helix. Several indirect lines of evidence had led to the conclusion that parts of the polypeptide chain in globular proteins are in the form of alpha helices and in haemoglobin PERUTZ had found traces of an X-ray reflexion characteristic of this configuration; but these results are the first direct proof that alpha helices are present and, indeed, enable them to be seen directly for the first time ».

Voilà un bel hommage à PAULING et COREY [15] inventeurs de l'hélice alpha.

Cette hélice est une hélice à droite, car la direction de la valence reliant l'atome de carbone alpha à l'atome de carbone bêta est à l'opposé de la direction C→O du groupe carbonyle. Il se pourrait bien que l'hélice à droite ait l'avantage pour des raisons stériques (HUGGINS) [16].

Le second genre de preuves est fourni par les mesures de pouvoir rotatoire et de la dispersion de celui-ci. Nous renvoyons à YANG et DOTY [17], à SCHELLMANN [18] et à ELLIOTT, HANBY et MALCOLM [19] pour l'exposé de cette méthode. Un polypeptide libre a un pouvoir rotatoire fortement lévogyre tel qu'on le constate dans les solutions concentrées d'urée. Une protéine en solution aqueuse diluées possède un pouvoir rotatoire beaucoup moins lévogyre, la différence est due à la contribution de la structure en hélice alpha qui augmente le pouvoir rotatoire si elle est une hélice à droite.

Comme nous l'avons dit, il ne faut pas s'attendre à trouver le polypeptide entièrement sous cette conformation. KENDREW (op. cit.) estime à 65-72% la fraction de la longueur totale du polypeptide de la myoglobine qui se trouve sous forme d'hélice alpha. Le pouvoir rotatoire donne pour diverses protéines, selon la façon de faire le calcul, des fractions soit de 30 à 40%, soit d'environ 60%, malheureusement, comme on le voit, avec

[16] *Putzeys* - pag. 6

une incertitude sérieuse. Cela ne doit cependant pas faire oublier que les mesures de pouvoir rotatoire suggèrent l'existence d'un pourcentage important de conformation en hélice alpha.

* * *

Parmi les liaisons qui conditionnent la conformation d'un polypeptide, il faut faire une place spéciale à la liaison disulfure.

Liaison covalente, elle pourrait être considérée comme faisant partie de la structure primaire, mais dans le cadre de cet exposé, nous la considérerons comme un facteur de conformation.

Il faut noter, en premier lieu, que la liaison disulfure n'est pas indispensable à la structure tertiaire d'une protéine. La myoglobine n'en renferme pas et quoiqu'il y ait quatre thiols dans l'hémoglobine, il n'y a pas de liaison disulfure.

En second lieu, il y a des protéines pauvres en soufre, où la liaison disulfure n'a probablement qu'une influence réduite et localisée sur la structure tertiaire; telles sont l'ovalbumine, qui renferme une liaison disulfure par 46.000 g, et l'édestine, qui en renferme probablement une par 25.000 g. Le contraste est frappant avec l'ocytocine, l'insuline, la ribonucléase, la sérumalbumine qui en renferment dix fois plus en chiffres ronds.

	Un pont S-S par
Ovalbumine	46.000 g
Edestine	25.000 »
Ribonucléase	3.200 »
Sérumalbumine	2.000 »
Insuline	2.000 »
Ocytocine	1.000 »

Dans ce dernier cas, l'influence sur la structure et les propriétés doit être prépondérante. Sur la structure tout d'abord.

S'il est vrai que les ponts S-S peuvent stabiliser l'hélice alpha, il se peut aussi, et dans certains cas, il est manifeste qu'ils sont incompatibles avec cette conformation et qu'ils servent précisément à maintenir une conformation différente (BOYER) [20]. SCHERAGA [21] a proposé pour la ribonucléase, un modèle de structure formé de six segments d'hélice alpha, LINDLEY et ROLLETT [22] en ont présenté un, formé de deux hélices, pour l'insuline. Mais une remarque de PERUTZ [23] suggère que dans l'un comme dans l'autre cas la teneur en hélice alpha doit être faible ou nulle. Le dernier mot sera dit par les Rayons X. Attendons.

La haute teneur en ponts S-S permet aussi de comprendre que la structure tertiaire puisse être considérablement modifiée par changement de pH ou par l'urée de façon renversable, alors que, dans les mêmes conditions, pour les protéines pauvres en ponts S-S, la structure tertiaire est irrémédiablement perdue.

Il n'est par conséquent pas permis de généraliser et d'étendre à un groupe les conclusions tirées de l'étude de l'autre.

* * *

On pourrait appeler « vulgaires », les liaisons formées par les groupes hydrocarbures apolaires. Ils constituent la grosse masse des groupes latéraux. D'après le relevé de TRISTRAM [24], la teneur en groupes apolaires (phénylalanine, leucine, isoleucine, valine et alanine) exprimée en pour-cent du nombre total des groupes, est en moyenne de 40% pour nne bonne dizaine de protéines, avec une marge de variation de ±8%.

Les liaisons qu'ils forment relèvent des forces de cohésion, ou forces de Van der Waals, relativement faibles, de l'ordre de 1 à 2 Kcal.

Leur intérêt comme facteurs de structure tertiaire résulte de leur grand nombre et du fait qu'elles ne sont pas affaiblies par la présence d'eau (SCHEIBE) [25]. Leur efficacité dépend des

[16] *Putzeys* - pag. 8

possibilités de rapprochement à très courte distance et de leur accumulation en certains points, plutôt que de leur abondance totale. L'ovalbumine avec 47,5% de groupes « hydrophobes », ce qui est presqu'un maximum, n'est précisément pas un exemple de stabilité. L'hypothèse que les groupes non polaires seraient rassemblés à l'intérieur de la molécule et que les groupes polaires et ionisables se trouveraient en surface, n'est qne partiellement sontenne par la structure de la myoglobine (KENDREW et al., op. cit.); 45% seulement des groupes apolaires, parmi lesquels la phénylalanine [6] et la méthionine [2] sont à l'intérieur (KENDREW dit « ensevelis »), à peu près toutes les alanines se trouvent dans les 55% de groupes apolaires dirigés vers l'extérieur.

Et ceci nous amène une fois de plus à la conclusion que la conservation de la strncture tertiaire ne peut être attribuée à un genre de liaisons en particulier, mais qu'elle est le résultat de la collaboration des affinités résiduelles de tous les groupes.

* * *

Restent les groupes polaires. Parmi ceux-ci la fonction alcool (Sérine, Thréonine, Hydroxyproline) n'est pas négligeable sans être abondante. Elle représente généralement 10%, parfois un peu plus, du nombre total des restes d'aminoacides.

La fonction alcool ne s'ionise pas dans les conditions à considérer lorsqu'il s'agit de protéines, mais tant par sa faculté de former des ponts hydrogène que par ses possibilités d'estérification, elle pourrait intervenir en des points critiques de la structure tertiaire. La fonction phénol possède des propriétés semblables mais à cause de son caractère acide elle entre dans le cadre du paragraphe suivant.

* * *

Il faut attribner une importance particulière aux groupes à caractère acide-base dont résultent les propriétés d'ampholytes

et d'amphions qui font des protéines des polyélectrolytes très particuliers. Ce qui nous intéresse exclusivement ici, c'est la contribution que les groupes acidobasiques apportent au maintien de la structure tertiaire, et quelles que soint les divergences d'opinion à ce sujet, on ne peut nier le fait que les protéines peuvent être dénaturées par acidification et alcalinisation. De toute évidence, lorsque les liens, de quelque nature qu'ils soient, entre les groupes acidobasiques sont rompus, les autres liens ne suffisent plus à maintenir la structure tertiaire originale.

On a voulu attribuer la destruction de la structure tertiaire à un effet purement électrostatique de la charge *totale,* soit positive soit négative, qu'acquièrent les polypeptides aux pH extrêmes, et on cite comme exemple les protamines, dépourvues de structure tertiaire, polypeptides libres et déliés à cause de la répulsion du grand nombre de charges positives des groupes guanidinium. Il ne s'agit pas de nier l'importance de ces répulsions électrostatiques, mais de se demander si cette interprétation est suffisante et péremptoire.

Si les protamines sont un bon exemple de l'effet de l'accumulation de charges de même signe et très rapprochées, elles sont un mauvais modèle d'une protéine, car elles ne renferment très généralement que des groupes ionisables de même espèce, notamment des groupes acides cationiques, et la possibilité de former des liens entre groupes acidobasiques n'est pas réalisée

D'autre part, si l'hypothèse envisagée était vraie, il faudrait s'attendre à trouver un maximum de résistance à la dénaturation au pH isoélectrique, ce qui ne se vérifie pas. Par exemple l'amandine et l'édestine ont toutes deux une zone de stabilité limitée à pH 5 du côté acide. Pour la première, le point isoionique se situe un peu au delà de 6; pour la seconde, il doit se trouver entre pH 8 et 9 (Travaux non publiés de MELLES MAHIEU et RUTGEERTS). L'édestine a sa limite de stabilité un peu en deça de pH 9, l'amandine résiste jusqu'au delà de pH 10. On a l'impression que c'est la rupture d'un cer-

tain nombre de liaisons particulières qui initie la dénaturation plutôt qu'une charge totale portée à une valeur critique.

On peut d'ailleurs se demander si la charge totale qu'un polypeptide de protéine peut acquérir est suffisante pour rompre toutes les autres liaisons. D'après les données de BRAUNITZER et al. [26] la chaîne alpha de l'hémoglobine humaine peut acquérir un maximum de 25 charges positives ou de 23 charges négatives (s'il n'y a pas de carboxyles à l'état d'amide) sur un total de 142 restes d'aminoacides, soit environ une charge tous les six restes. Il en est pratiquement de même pour la chaîne bêta.

Enfin la possibilité de liaisons (d'interactions, si on préfère), entre groupes particuliers ne saurait être exclue a priori, d'autant moins que de nombreuses observations font pressentir leur présence.

Les plus anciennes indications de la présence de liaisons entre groupes ionisables ont été fournies par l'analyse des courbes de titrage.

Le nombre moyen de protons (ν) échangés par une protéine avec son solvant est donné par [27]:

$$\nu = q_1 \frac{H}{H + K_1 Z} + \frac{H}{H + K_2 Z} + \ldots \ldots - q'_1 \frac{K'_1 Z}{K'_1 Z + H} - q'_2 \frac{K'_2 Z}{K'_2 Z + H} \ldots$$

équation dans laquelle les symboles ont la signification suivante:

H activité des ions hydrogène,

q_1, q_2 .... q'_1, q'_2 .... nombre de groupes ionisables respectivement des acides cationiques et des bases anioniques,

K_1, K_2 .... K'_1, K'_2 .... constantes de dissociation intrinsèques respectivement des groupes q et q',

Z coefficient fonction de la force ionique et de la charge effective de la protéine.

Si tous les groupes ionisables et leur nombre exact sont connus, la courbe de titrage doit être représentée par la relation ci-dessus. A chaque groupe doit correspondre une constante intrinsèque déterminée, la même d'ailleurs pour toutes les protéines. Or on trouve en général pour le même groupe des constantes intrinsèques différentes dans les diverses protéines. Pis encore: pour rendre compte de la courbe de titrage d'une même protéine, on est parfois obligé d'attribuer à un même groupe deux constantes intrinsèques différentes. Ainsi pour la dissociation de l'imidazolium, le pK de la constante intrinsèque varie da 5,6 à 7; le pK d'un carboxyle, de 3 à 5. Il y a longtemps que COHN attira l'attention sur ce fait [28]. Les anomalies constatées ont été attribuées soit aux interactions de nature électrostatique entre les groupes, soit à la présence d'un certain nombre de groupes cachés, c'est-à-dire non immédiatement accessibles aux acides et aux bases servant au titrage. La question a été passée en revue par STEINHARDT et ZAISER [29] et par LINDERSTROM-LANG et NIELSEN [30].

L'origne de l'idée de groupes cachés se trouve dans la constatation suivante. Lorsqu'on titre une protéine en retour, on obtient un résultat supérieur à celui du titrage direct, si celui-ci a été poursuivi suffisamment loin. On titre également plus de groupes lorsqu'on dénature au prélable la protéine par une acidification ou une alcalinisation suffisante. La dénaturation révèle donc des groupes que le titrage direct et ménagé ne parvient pas à atteindre. On pourrait être accusé de byzantinisme si on posait la question de savoir si la révélation des groupes cachés est la cause ou la conséquence de la dénaturation. Un exemple particulièrement net est fourni par l'étude que STEINHARDT et ZAISER [31] ont faite de la courbe de titrage de l'hémoglobine de cheval. Le titrage en retour révèle 36 groupes basiques, qui ne sont pas immédiatement titrables dans l'opé-

ration directe. La différence se manifeste à partir de pH 4,5, limite de la zone de stabilité. La dénaturation et la révélation de groupes cachés vont donc de pair.

On peut rapprocher de cette observation la remarque de KENDREW à propos de la structure de la myoglobine. Presque tous les restes de lysine, d'arginine et d'acide glutamique se trouvent à l'extérieur de la molécule; il faut donc penser que les restes de l'histidine et de l'acide aspartique se trouvent tous ou presque tous à l'intérieur. Si cette remarque est encore valable pour l'hémoglobine, ce sont l'histidine et l'acide aspartique qui sont responsables des 36 groupes cachés. Précisément l'hémoglobine renferme 36 molécules d'histidine. Supposons l'imidazolium lié à un carboxylate; la combinaison ainsi réalisée est un nouveau groupe amphotère donnant lieu aux équilibres suivants:

$$(1)\quad R_1COOH + Im^+H\text{-}R_2 + H_2O \rightleftharpoons [R_1\text{-}COO^- \dashrightarrow Im^+H\text{-}R_2] + H_3O_+$$

$$(2)\quad [R\text{-}COO^- \dashrightarrow Im^+H\text{-}R_2] + H_2O \rightleftharpoons R_1\text{-}COO^- + Im\text{-}R_2 + H_3O^+$$

A ces deux équilibres correspondent deux constantes de dissociation K_1 et K_2 qui sont reliées aux constantes intrinsèques des groupes libres par

$$K_1 = K_{COOH}\, e^{-\frac{\Delta G^\circ}{RT}}$$

$$K_2 = K_{Im^+H}\, e^{\frac{\Delta G^\circ}{RT}}$$

où ΔG° est l'enthalpie libre de liaison.

Si $\triangle G^{\circ} = -1500$ cal, il y aura une différence d'une unité entre le pK d'un groupe libre et le pK de ce même groupe lié, le pK du carboxyle étant diminué, celui de l'imidazol augmenté. La valeur numérique choisie pour $\triangle G^{\circ}$ convient généralement pour l'interprétation des courbes de titrage: elle correspond aussi à l'énergie électrostatique si on suppose que l'oxygène du carboxylate et l'azote de l'imidazolium sont à la distance 2,8 Å, la constante diélectrique étant celle de l'eau.

Il est important de noter que l'existence des équilibres ci-dessus implique que les deux groupes soient maintenus par la structure tertiaire dans le voisinage immédiat l'un de l'autre, leur distance pouvant seulement varier entre certaines limites. Notons encore que l'état représenté par le premier membre de l'équation 1 est nonliant tout comme l'état figuré par le second membre de l'équation 2. Pour indiquer que ces groupes sont contraints de rester dans le voisinage l'un de l'autre nous les appelerons « pairés ».

Supposons encore, pour simplifier, que tous les imidazols soient pairés avec des carboxyles. On voit alors ce qui se produira au cours du titrage de l'hémoglobine par un acide. Nous disposons de courbes de titrage de la carbonylhémoglobine de cheval, récemment établies avec beaucoup de soin par M.me BRONCKERS-DE JAER. Considérons le titrage en solution 0,1 m de chlorure potassique et à 25°C. Le point isoionique se trouve alors à pH 7,02. Un abaissement du pH d'une unité fixe 10,8 protons sur la molécule. Il est évident que seuls des imidazols peuvent être rendus responsables de la neutralisation dans la région entre pH 7 et 6. C'est par conséquent la réaction 2 dans le sens de droite à gauche qui entre en jeu. L'acidification ne rompt pas de liaisons, dans le cas présent elle en forme. Les protons seront d'autant plus facilement fixés que l'imidazol pairé est une base plus forte que l'imidazol libre. Bientôt tous les groupes pairés seront liés; il faudra un abaissement de pH jusqu'en dessous de 4,5 avant que cette liaison ne commence à être rompue et passe à l'état non-liant du premier mem-

[16] *Putzeys* - pag. 14

bre de l'équation 1. Avant cela des groupes carboxylate libres, notamment ceux situés en surface, auront pris part au titrage, et au total on devra obtenir la courbe de titrage immédiat de STEINHARDT et ZAISER, correspondant au pH atteint immédiatement (3 secondes) après l'addition d'acide.

En dessous de 4,5, le pH varie lentement en fonction du temps; cette variation est d'autant plus rapide que le pH est plus bas et devient instantanée vers pH 2,0.

Cette région de variation lente correspond précisément au domaine d'ionisation des groupes liés, dont le pK doit se situer vers 3,6. Si ces liaisons sont critiques pour la conformation tertiaire, et leur nombre [36] justifie cette hypothèse, on peut s'attendre à une dénaturation d'autant plus rapide que le nombre moyen de liaisons rompues sera plus grand, c'est-à-dire que l'acidité sera plus forte. Seulement, dans les molécules dénaturées, les groupes en question ne sont plus pairés, la constante de dissociation des carboxyles revient à sa valeur normale et, à pH égal, le nombre des protons fixés sera plus grand. Les groupes « cachés » sont d'après celà, 36 groupes carboxyles pairés; l'hypothèse d'un empêchement stérique est superflue.

Contre l'idée de groupes pairés, de graves objections ont été soulevées par JACOBSON et LINDSTROM-LANG [32] sur la base de mesures dilatométriques. Celles-ci permettent en principe de décider si un groupe acide ou basique est libre ou lié. Ainsi la neutralisation des carboxylates de la bêta-lactoglobuline s'accompagne d'un incrément dilatométrique de +11 à +12 ml par équivalent d'acide neutralisé; on retrouve également cette valeur pour des composés simples où il n'est pas question de liaison. Une discussion approfondie de la méthode dilatométrique dépasserait le cadre de cet exposé; contentons-nous de constater que dans le cas de l'hémoglobine, on devrait s'attendre à un incrément d'environ −3 ml par équivalent d'acide neutralisé si les imidazols étaient aussi libres que dans l'histidine isolée. Or on trouve (M.me BRONCKERS-DE JAER et M.lle VANDORMAEL, résultats non publiés) des incréments de +3 à +5. Cette

constatation n'est certainement pas favorable à l'hypothèse de groupes libres et elle est compatible avec celle de groupes liés.

* * *

L'importance de liaisons entre groupes acido-basiques se révèle encore dans l'étude des *zones de stabilité* auxquelles SVEDBERG, il y a déjà bien longtemps, a consacré une attention particulière [33].

Les zones de stabilité révèlent deux choses: tout d'abord, que beaucoup de protéines se dissocient en sous-molécules lorsque le pH dépasse certaines limites; ensuite, que cette dissociation — parfois réversible — se produit entre des limites de pH très étroites.

Sans vouloir généraliser prématurément, ces anciennes observations comme celles, plus récentes, de PERUTZ, nous font penser que les protéines qui comprennent deux ou plusieurs polypeptides sont constituées de sous-molécules, chacune d'elles formée d'un seul polypeptide ayant sa structure tertiaire propre. Si tel est le cas, l'assemblage doit être assez compact pour former un nombre suffisant de liaisons entre les sous-molécules et réaliser un ensemble robuste. La dissociation de cet ensemble aux limites de la zone de stabilité montre nettement que les liaisons entre groupes acido-basiques sont déterminantes.

Il ne reste disponible pour ces liaisons que les groupes situés à la surface des sous-molécules. Dans la myoglobine, ce sont les groupes de l'arginine, de la lysine et de l'acide glutamique. Du côté acide de la zone de stabilité, entrent en ligne de compte les carboxylates liés aux groupes à charge positive, ce qui situe la limite de la zone de stabilité vers pH 4,5. C'est le cas pour beaucoup de protéines. Du côté alcalin, si ce sont les groupes aminés du radical de la lysine qui sont pairés, on doit s'attendre à une limite vers pH 10 à 11. Ce serait le cas de l'amandine.

[16] *Putzeys* - pag. 16

Cependant il faut éviter de généraliser. Les hémocyanines nous fournissent des exemples de zones de stabilité beaucoup plus étroites. L'hémocyanine d'Helix pomatia [34] se scinde en sous-molécules à partir de pH 4,2 d'un côté, à partir de pH 7 de l'autre. De même, plusieurs globulines végétales [35] ont une limite acide un peu au-dessus de pH 5 et une limite basique située à pH 8,9. La nature des groupes reste à définir, mais on peut affirmer qu'ils doivent être sensibles aux ions hydrogène.

On a voulu voir dans ces dissociations des cas de simple répulsion électrostatique. Cela me semble difficile à admettre car la dissociation de l'hémocyanine d'Helix pomatia se produit à quelques dixièmes d'unité pH du point isoélectrique où la charge totale doit être voisine de zéro.

La complexité de ces phénomènes est mise en évidence par l'observation de BROSTEAUX [36] que la présence d'une faible concentration en ions calcium ou magnésium (0,01 m) empêche l'hémocyanine (*Helix pomatia*) de se dissocier à pH 7. La zone de stabilite s'étend alors jusqu'à pH 9. La limite acide n'est pas affectée par la présence de Ca ou de Mg. Les deux sous-molécules pourraient être liées par la formation de complexes entre les ions de Ca ou de Mg et les carboxylates superficiels, à la manière des complexes avec l'éthylène-diamine-tétracétate.

* * *

Aux réflexions qui précèdent, on pourrait en ajouter bien d'autres, telles celles qui concernent l'agrégation, la dénaturation et la stabilisation par substances étrangères. Elles rentrent dans l'image générale que l'on peut actuellement se faire de la structure de la molécule de protéine, que les travaux de KENDREW, de PERUTZ et de leurs collaborateurs nous donnent maintenant la satisfaction de « voir » dans leur réalité.

La structure tertiaire de la molécule de protéine est le résultat de la collaboration de tous les genres de liaisons réalisés par

les affinités résiduelles du polypeptide et, par ce fait même, elle est conditionnée par la nature, l'abondance et la séquence de ses aminoacides. La connaissance de cette structure tertiaire, très imparfaite dans la plupart des cas mais relativement satisfaisante dans l'ensemble, soulève de nouveaux problèmes, tels ceux de sa synthèse cellulaire et de ses propriétés enzymatiques. Mais cela, pour employer une phrase chère à Rudyard Kipling, c'est une autre histoire.

BIBLIOGRAPHIE

[1] BERNAL J.D., *Structure arrangements of Macromolecules*. « Disc. Faraday Soc. », *25*, 7 (1958).

[2] KENDREW J.C., BODO G., DINTZIS H.M., PARRISH R.G., WYCKOFF H. and PHILLIPS D.C., *A three-dimensional model of the Myoglobin molecule obtained by X-Ray analysis*. « Nature », *181*, 662.

[3] KENDREW J.C., DICKERSON R.E., STRANDBERG B.E., HART R.G., DAVIES D.R., PHILLIPS D.C. and SHORE V.C., *Structure of Myoglobin*. « Nature », *185*, 422 (1960).

[4] EDMUNDSON A.B. and HIRS C.H.W., *The Amino-acid Sequence of Sperm Whale Myoglobin. Chemical Studies*. « Nature », *190*, 663 (1961).

[5] KENDREW J.C., WATSON H.C., STRANDBERG B.E., DICKERSON R.E., PHILLIPS D.C. and SHORE V.C., *The Amino-acid Sequence of Sperm Whale Myoglobin. A Partial Determination by X-Ray Methods and its Correlation with Chemical Data*. « Nature », *190*, 666 (1961).

[6] WATSON H.C. and KENDREW J.C., *Comparison between the Amino-acid Sequences of Sperm Whale Myoglobin and of Human Haemoglobin*. « Nature », *190*, 670 (1961).

[7] PERUTZ M.F., ROSSMANN M.G., CULLIS A.F., MUIRHEAD H., WILL G. and NORTH A.C.T., *Structure of Haemoglobin*. « Nature », *185*, 416 (1960).

[8] WETLAUFER D.B., *Terminology of Protein Structure*. « Nature », *190*, 1113 (1961).

[9] AMBROSE E.J., BAMFORD C.H., ELLIOTT A. and HANBY W.E., *Water-soluble Silk: An α-Protein*. « Nature », *167*, 264 (1951).

[16] *Putzeys* - pag. 18

[10] BAMFORD C.H., ELLIOTT A. and HANBY W.E., *Synthetic Polypeptides*. New York, 1956, p. 388.

[11] LONTIE R., *Foids moléculaire et zones de stabilité des globulines végétales*. III. *L'édestane*. « Bull. Soc. Chim. Biol. », *29*, 875 (1947).

[12] SCHELLMAN J.A., a) *The Thermodynamics of Urea Solutions and the Heat of Formation of the Peptide Hydrogen Bond*. b) *The Stability of Hydrogen-bonded Peptide Structures in Aqueous Solution*. « Comptes-rendus Carlsberg », *29*, n. 14-15, 223 (1955).

[13] HARRINGTON W.F. and SCHELLMN J.A., *Evidence for the Instability of Hydrogen-bonded Peptide Structures in Water, based on Studies of Ribonuclease and Oxidized Ribonuclease*. « Comptes-rendus Carlsberg », *30*, n. 3, 21 (1956).

[14] SCHELLMAN J.A., *The Optical Rotatory Properties of Proteins and Polypeptides*. « Comptes-rendus Carlsberg », *30*, n. 21-27 (1956).

[15] PAULING L. and COREY R.B., *Atomic Coordinates and Structure Factors for two helical Configurations of Polypeptide Chains*. « Proc. Nat. Acad. Sci. U.S. », *37*, 235 (1951).

[16] HUGGINS M.L., *Polypeptide Helices in Proteins*. « J. Am. Chem. Soc. », *74*, 3963 (1952).

[17] YANG Y.T. and DOTY P., *The Optical Rotatory Dispersion of Polypeptides and Proteins in Relation to Configuration*. « J. Am. Chem. Soc. », *79*, 761 (1957).

[18] SCHELLMAN J.A., *The Optical Rotatory Properties of Proteins and Polypeptides*. « Comptes-rendus Carlsberg », *30*, n. 21-27 (1956).

[19] ELLIOTT A., HANBY W.E. and MALCOLM B.R., *Optical Rotation and Infra-red Spectra of some Polypeptide and Protein Films*. « Disc. Faraday Soc. », *25*, 167 (1958).

[20] BOYER P.D., *Disulfide and Sulfhydryl Groups in Protein Structure*. Brookhaven Symposia in Biology n. 13. Protein Structure and Function. Upton, New York, 1960.

[21] SCHERAGE H.A., *Protein Structure*. New York, 1961.

[22] LINDLEY H. and ROLLETT J.S., *An Investigation of Insulin Structure by Model Building Techniques*. « Biochim. Biophys. Acta », *18*, 183 (1955).

[23] PERUTZ M.F., *Brookhaven Symposia in Biology*, n. 13. Protein Structure and Function. Upton. New York, 1960, p. 86.

[24] TRISTRAM G.R., *Amino Acid Composition of Purified Proteins*, in « Advances Protein Chemistry », *5*, 83 (1949).

[25] SCHEIBE G., *Konstitution und physikalisch-chemisches Verhalten von Eiweissmolekülen*, in « Zwischen molekulare Kräfte », Karlsruhe, 1949.

[26] BRAUNITZER G., HILSCHMANN N., RUDLOFF V., HILSE K., LIEBOLD B. and MÜLLER R., *The Haemoglobin Particles. Chemical and Genetic Aspects of their Structure*. « Nature », *190*, 480 (1961).

[27] PUTZEYS P. and BOUCKAERT L., *The Ionisation of Proteins. I. Kon. VI. Acad. Wetensch.* (1942), IV, n. 3. *The Ionisation of Proteins*. II. Ibid. (1954), XVI, n. 1.

[16] *Putzeys* - pag. 19

[28] COHN E.J. and EDSALL J.T., *Proteins, Amino Acids and Peptides*. New York, 1943.

[29] STEINHARDT J. and ZAISER E.M., *Hydrogen Ion Equilibria in Native and Denatured Proteins*, in « Advances Protein Chemistry », *10*, 151 (1955).

[30] LINDERSTROM-LANG K. and NIELSEN S.O., *Acid-Base Equilibria of Proteins*, in « M. *Bier*. Electrophoresis », New York 1959, p. 35.

[31] STEINHARDT J. and ZAISER E.M., *Masking of Acid-binding Groups in Native Horse Carbonylhemoglobin*. « J. Biol. Chem. », *190*, 197 (1951).

[32] JACOBSEN C.F. and LINDERSTROM-LANG K., *Salt Linkages in Proteins*. « Nature », *164*, 411 (1949).

[33] SVEDBERG T., *The pH Stability Regions of the Proteins*. « Trans. Faraday Soc. », *26*, 740 (1930.

[34] ERIKSSON-QUENSEL I.B. and SVEDBERG T., *The Molecular Weights and pH-Stability Regions of the Hemocyanins*. « Biol. Bull. », *71*, 498 (1936).
PUTZERS P. and BROSTEAUX J., *Light Scattering and the Molecular Weight of the Proteins*. « Med. Kon. Vl. Acad. Wetensch », *2*, n. 1 (1941).

[35] BEECKMANS M.L. et LONTIE R., *Poids moléculaire et zones de stabilité des globulines végétales*. II. *L'Edestine*. « Bull. Soc. Chim. Biol. », *18*, 509 (1946).

[36] BROSTEAUX J., *Über den Einfluss der Ca- und Mg-Ionen auf die Stabilität des Hämocyanins*. « Naturwiss. », *25*, 249 (1937).

[16] *Putzeys* - pag. 20

DISCUSSION

LINDQVIST

You have correctly emphasized the importance of intermolecular hydrogen bonds to other molecules. If you look at all structures known, not only the simple amino acids, but also metal complexes such as the one I showed yesterday, you will find that there is no atoms in them having alone electron pair, which is not used in some way as Professor PUTZEYS said here. There is always formed some kind of hydrogen bond, if necessary another water molecule is introduced in the structure to bind it together so that even very slightly acidic hydrogen is used, even hydrogen on amino groups might be used for hydrogen bonding to external molecules. And of course, the difference between those hydrogen bonds which are very strong and the ionic interaction, i.e., the difference between strong hydrogen bond and a complete proton transfer is almost none.

SCHRAMM

The proteins containing many disulphide bridges seem to be more stable than the other ones. But there are other factors which can influence the stability; for instance TMV itself is very stable against heat denaturation (denaturation point between 90-95°C), on the other hand the free protein is much more instable (denaturation point about 37°C). Therefore, we can assume that the attachment of the RNA stabilizes the protein structure. I should also mention that WIELAND and others have found that enzymes are stabilized by combination with the substrates.

PUTZEYS

Disulphide bridges pertain to the proper structure of the protein molecule; one may consider them as pertaining to the primary structure, because they are covalent bonds. The properties of molecules having many disulphide bridges are different from those that have only a few or none. The « stabilization » they bring about is of a very different kind from the stabilization due to all kinds of substances which may be associated with a protein molecule, or may simply influence activity coefficients.

INVESTIGATIONS ON MOLECULAR BIOLOGY

SILVIO RANZI

Istituto di Zoologia dell'Università di Milano, Milano (Italia)

PROTEIN DIFFERENTIATION DURING EMBRYONIC AND LARVAL DEVELOPMENT

It is possible to investigate some aspects of protein differentiation during embryonic and larval development by means of DERRIEN et al. method (salting out), slightly modified, using ammonium sulphate added to a protein solution in WEBER and EDSALL fluid [7], or by dilution with water (salting in) of a protein solution in 1M KCl [6].

In such a way, a gradual transformation in protein composition may be easily followed from the egg as far as to adult stage. The salting out (or salting in) diagrams of animals belonging to a same systematic group are similar for the adults, for the eggs and for the corresponding embryonic stages of the animals considered (fig. 1).

Adult antigens were detected in fractions isolated in salting out or salting in experiments. As soon as at cleavage stage of toad (*Bufo bufo* and *B. viridis*) eggs it was possible to demonstrate the appearance of new antigens. By considering the hybrids of the two species, it was seen that some antigens, which appear at cleavage stage, have a maternal character while some others behave as paternal (fig. 2) [7].

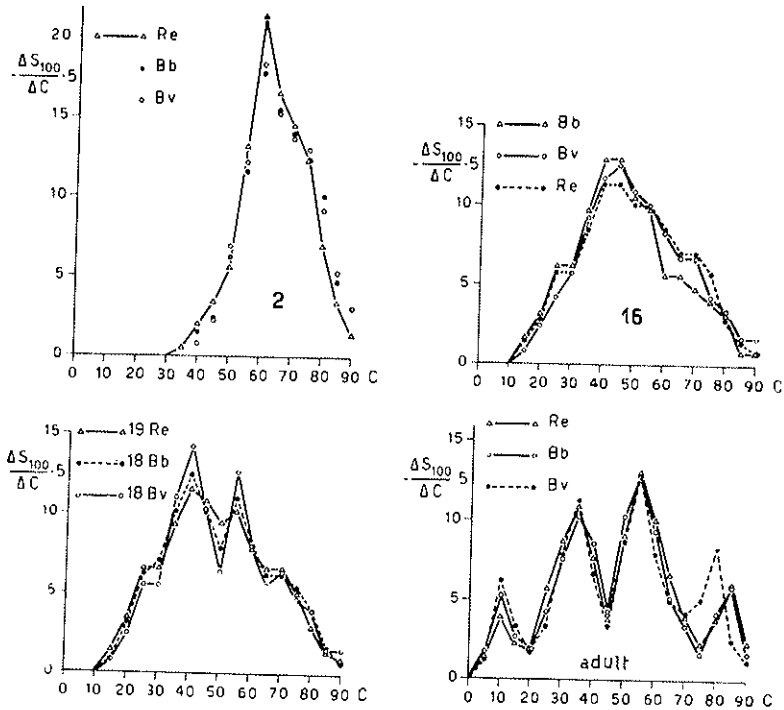


FIG. 1 — Salting out diagrams of the proteins extracted from different stages of development and from adult animals of *Rana esculenta* (Re); *Bufo bufo* (Bb) and *B. viridis* (Bv): 2, Fertilized egg; 16, Neural tube stage; 18, Muscular response stage; 19, Heart beat stage; *Adult*, Adult animals; (numbers of the stages following RUGH). (From RANZI, CITTERIO and SAMUELLI, 1960).

During the embryonic development different changes occur in these antingens; at first some exhibiting one reactive group appear, while in the following stages only the antigens with two reactive groups can be detected [9]. Therefore there is a real molecular ontogenesis, in some way comparable to the transformation described by BON in the α-chrystallin [1].

At least some of the adult antigens appear very early in the embryo. In sea urchin embryos some of them are syn-

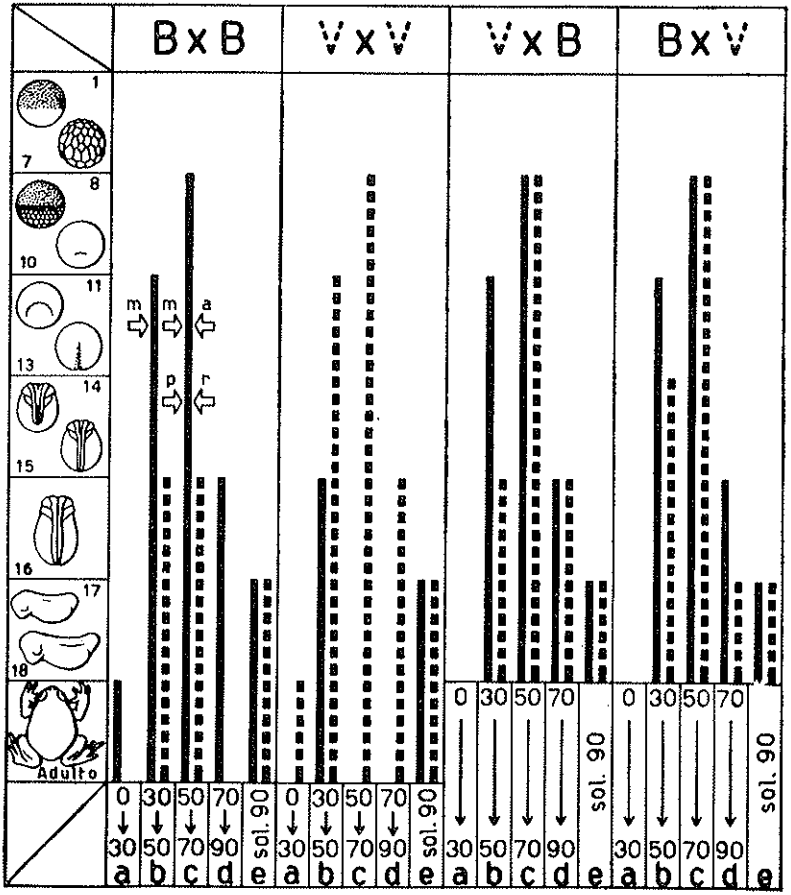


FIG. 2 — Antigens of toads and of their hybrids. The different developmental stages are on the ordinate: they are represented by sketches and the numbers of the stages (following RUGH). B × B: *Bufo bufo*; V × V: *Bufo viridis*; V × B: lethal cross of *B. viridis* F. × *B. bufo* M.; B × V: viable cross of *B. bufo* F. × *B. viridis* M. The numbers on the abscissa indicate the per cent of ammonium sulphate saturation corresponding to concentrations between which it is possible to precipitate the antigens. The precipitation with anti-*Bufo bufo* serum is the uninterrupted column. The precipitation with anti-*Bufo viridis* serum is the interrupted column. The arrow *a* represent the start of the anti-actin precipitation; the arrow *m* the start of anti-myosin precipitation; the arrows *p* and *r* the start, in our experiments of precipitation with anti-plasma and anti-red cells sera. (From RANZI, CITTERIO and SAMUELLI, 1960).

thesized at swimming blastula stage, when primary mesenchyme is formed, while in frog and chick embryos the muscle protein antigens appear just before the end of gastrulation, when the muscle rudiments are not yet detectable (at primitive streak stage in chick embryo, following EBERT) [6]. These transformations of the early embryo proteins are associated to the formation of osmotically active substances which increase the osmotic pressure of the embryo causing a water uptake. Salts dissolved in water also enter into the embryo and are employed in the formation of the skeleton as well as for the synthesis of metalloproteins.

Working on sea urchin and Amphibian (frogs, toads and interspecific crosses of the last) embryos, my collaborators and I could recognize some modifications which occur in the proteins during embryonic and larval development, and could get some knowledge of the corresponding transformations occuring during silk worm, chick, guinea-pig and other animal development [2]. We could, for instance, demonstrate the presence of two proteins in the electric organ of the *Torpedo,* one exhibiting the same precipitation behaviour as muscle myosin and precipitating with the antisera prepared against this protein, the other precipitating in the same way as muscle actin and reacting with the antisera prepared against this protein. This phenomenon accounts for the presence in the electric organ of muscles antigens: in fact during the phylogenesis, the electric tissue was derived from the same rudiments of the muscles [3].

PROTEIN DENATURATION AND EMBRYONIC DETERMINATION

The investigations illustrated above have essentially a descriptive character, but they are part of a larger experimental work carried out in order to study the process of embryonic development at a molecular level. After a research on various developmental malformations which may be obtained by expe-

rimentation and are known since long ago, we could state that in the embryo the first step of some determination phenomena is a protein denaturation [5, 6, 7, 8]. These phenomena are: ectoderm formation and origin of other organs from the animal pole in sea urchin embryos; notochord formation in Amphibian embryos; nervous system appearance in these embryos. The evidences for this conclusion were the following: 1) proteins extracted from eggs and embryos, or even well characterized proteins from adults (serum albumin, myosin, actin and so on) are denatured by treatment with substances which induce a hyperdevelopment of the above mentioned rudiments (the viscosity of their solutions varies; more reactive groups can be detected; some proteins, like lipovitellin, show at the ultracentrifuge the formation of subunities); 2) the proteins extracted from embryos which exhibit a hyperdevelopment of the above mentioned rudiments, are predisposed to a further denaturation induced by urea or are more easily digested by proteolytic enzymes (like trypsin and papain); 3) the above embryos are less resistant to the substances considered; 4) urea (the classical denaturing agent of proteins) can produce all the determination phenomena which, as we could demonstrate, are based on protein denaturation.

Since long ago it is known that other substances produce some alterations in the determination, called by the early investigators « inhibitions of development ». These substances, in the three cases mentioned above, account for the inverse alterations, i.e.: entoderm determination in sea urchin embryos; mesoderm (instead of notochord) determination in Amphibian embryos; reduction of nervous system. LiCl is the best known substance used in these experiments. It was evident to us that this phenomenon is, at a molecular level, a stabilization of protein molecules, which preserves from denaturation. This conclusion was derived from the following experiments: 1) proteins extracted from eggs and embryos or well characterized proteins from adults (serum albumin, myosin, actin, and

so on) are not denatured by treatment with LiCl or other substances having the same morphological effect on the embryonic development (the viscosity of solutions of fibrous proteins increases; the flow birefringence does not vary; the reactive groups do not increase; subunities are never seen in the ultracentrifugation analyses); 2) proteins extracted from embryos which exhibit the characteristic malformations produced by LiCl are more resistant to the denaturing effect of urea and less easily digested by proteolytic enzymes (like trypsin and papain). In this way the lethal cross *Bufo viridis* ♀ × *B. bufo* ♂ is very instructive because the hybrids of this cross exhibit the same malformations as LiCl-treated embryos: the proteins extracted from these embryos behave like those extracted from LiCl-treated embryos; 3) LiCl-treated embryos are more resistant to the action of urea and proteolytic enzymes. We may interpret the action of LiCl as follows: Li ion bounds to the protein molecule and enlarges its aqueous coat because of its very high hydration, so protecting it from denaturation and the action of proteolytic enzymes, consequently inhibiting its metabolism. This, at a molecular level, is the development inhibition of the old experimental embryologists.

These conclusions, supported by experimental data, induce us to imagine the above mentioned phenomena in the following way: the formation of ectoderm and other organs from the animal pole in sea urchin embryos, the notochord formation in Amphibian embryos, even the formation of neural tube, are based on an intense protein synthesis, which implies the rapid metabolism of preexisting molecules (fig. 3). This process follows a denaturation which may be an extensive one, so that the rudiments here considered will be larger, or inhibited by the formation of the aqueous coat by Li, in which instance the rudiments are reduced in extension or even completely disappear (transformation of notochord into somites). Our experiments put also into evidence an auto-limiting mechanism of

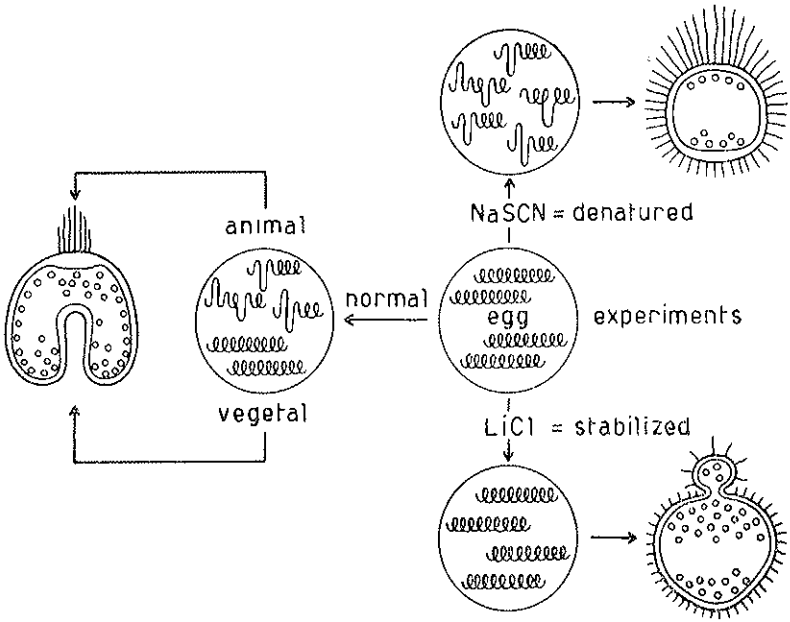


FIG. 3 — Tentative interpretation of the behaviour of protein molecules (indicated by helixes) in normal development of sea urchin and in experimental conditions when they are denatured by NaSCN or stabilized by LiCl

denaturation caused by an excess of some metabolites. The same process evidently regulates the dimensions of the various rudiments in normal development.

RIBONUCLEOPROTEINS AND CYTODIFFERENTIATION

The protein denaturation mentioned above represents a necessary step to the synthesis and preparation of material essential to the embryo, but there is something more. The genetic information embodied in the DNA is passed on RNA, and in contact with RNA, specific proteins are formed. Niu [4] succeeded in extracting ribonucleoproteins active in conditioning

the organ development in Amphibian embryos; RANZI, GAVAROSI and CITTERIO [10] performed some experiments, extracting ribonucleoproteins from liver and heart (1).

The presumptive cardiac area of 16-hour chick embryo was incubated in the presence of these ribonucleoproteins. After 79 hour incubation the heart volumes expressed in $\mu^3 \times 10^6$ (mean and standard error) were 2.4 ± 0.342 for the controls and 4.2 ± 0.627 for the embryos incubated with heart ribonucleoproteins. There is a significant difference ($0.05 > P > 0.02$), the increase amounting to 75.0%. The heart of embryos incubated with liver ribonucleoproteins did degenerate. No degeneration was detectable on the contrary if the liver ribonucleoproteins were denatured by heating.

In a second series of experiments [11] ribonucleoproteins from liver, heart, muscles and kidney of adult chick were extracted and a sponge made up of coagulated egg albumen was soaked with these ribonucleoproteins and then grafted in

(1) The organs, as soon as collected from animals after 48-hour starvation were put in a cold 0.25 M sucrose solution and all the following operations were performed in a cold room at 0°C. The organs were minced and carefully homogenized in a glass Waring Blendor with a 0.9% NaCl solution (100 g of organ for each 400 ml of saline). Afterwards the homogenate was filtered on a thin gauze and stirred for 1 hour, during which the pH was made 7.8 with a few 1 N NaOH. The solution was then centrifuged in the Servall centrifuge at 5000 r.p.m. for 30 min, the precipitate discharged and the supernatant carefully adjusted at pH 4.5 with 1 N HCl and finally centrifuged at 5000 r.p.m. for 10 min. The precipitate thus obtained was redissolved in saline and the pH of this solution adjusted at 7.2. After stirring for 30 min, the undissolved particles were eliminated by centrifugation (at 5000 r.p.m. for 10 min), therefore by adding 2 vol. of 95% ethanol a new precipitate was obtained which was redissolved in saline at pH 7.2. Purification was obtained by centrifuging at 5000 r.p.m. for 10 min and then another 2 vol. of 95% ethanol were added to the supernatant. The precipitate thus formed was redissolved in saline at pH 7.2 and purified by centrifugation at 5000 r.p.m. for 30 min. Potassium acetate was then added to this solution up to 2% final concentration and this solution was finally diluted with 2 vol. of 95% ethanol. The new precipitate was separated by centrifuging at 5000 r.p.m. for 10 min and redissolved in Gey fluid at pH 7.2. A last purification was performed by centrifuging at 3000 r.p.m. for 5 min and the supernatant dyalized against Gey fluid. This solution contained the ribonucleoproteins which were grafted in our experiments.

the chorio-allantois of the embryo (fig. 4). It was possible to obtain the formation in the chorio-allantois respectively of glandular tissue (fig. 5), of multinucleate elements, with fibrillar cytoplasm (fig. 7), of another kind of elements (different from those induced by heart ribonucleoproteins for the peripherical nuclei) (fig. 8), of tubules and other formations tentatively interpreted as kidney glomeruli (fig. 6). These experiments deal with a phenomenon at a molecular level; the molecules of heart proteins (as well as of liver, skeletal muscles or kidney) are synthesized on an information given by heart ribonucleic acid (or respectively liver, skeletal muscles or kidney); therefore hy grafting ribonucleoproteins of different organs it is possible to

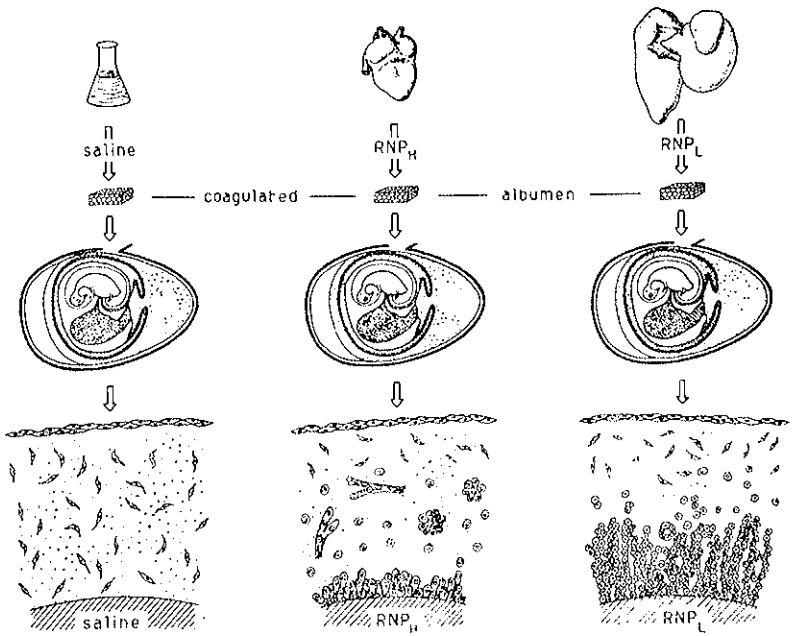


FIG. 4 — Experimental procedure in grafting ribonucleoproteins extracted from adult heart (RNP) or liver (RNP) in the chorio-allantois. The sketches at the bottom indicate the obtained inductions.

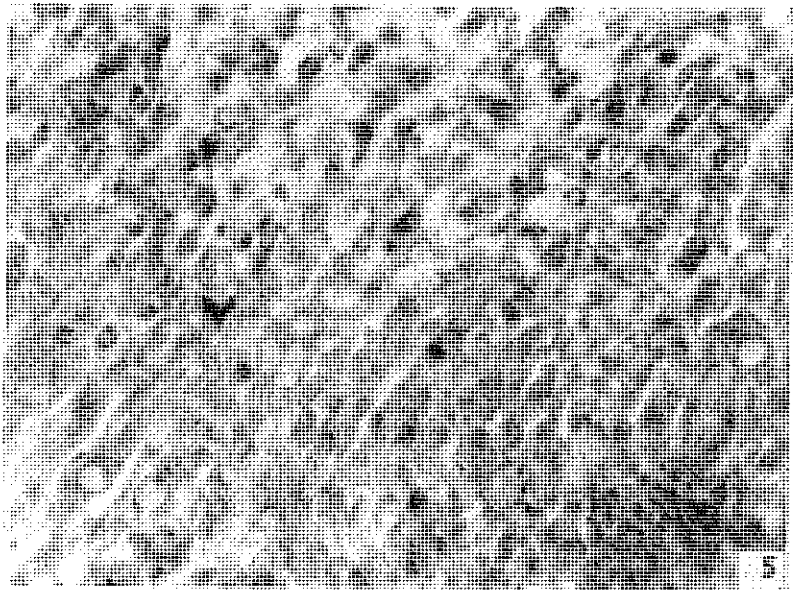


FIG. 5 — Glandular tissue induced by liver ribonucleoprotein (× 790).

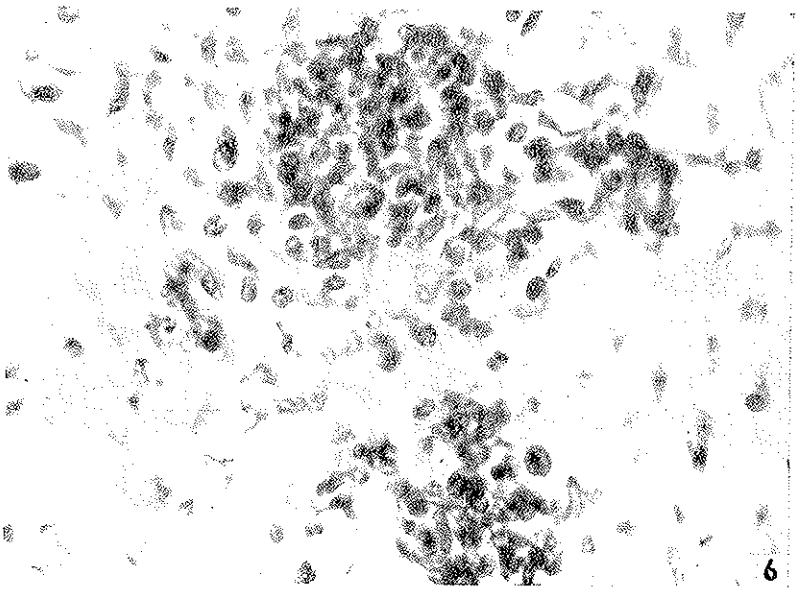


FIG. 6 — Structure tentatively interpreted as glomeruli induced by kidney ribonucleoprotein (× 790).

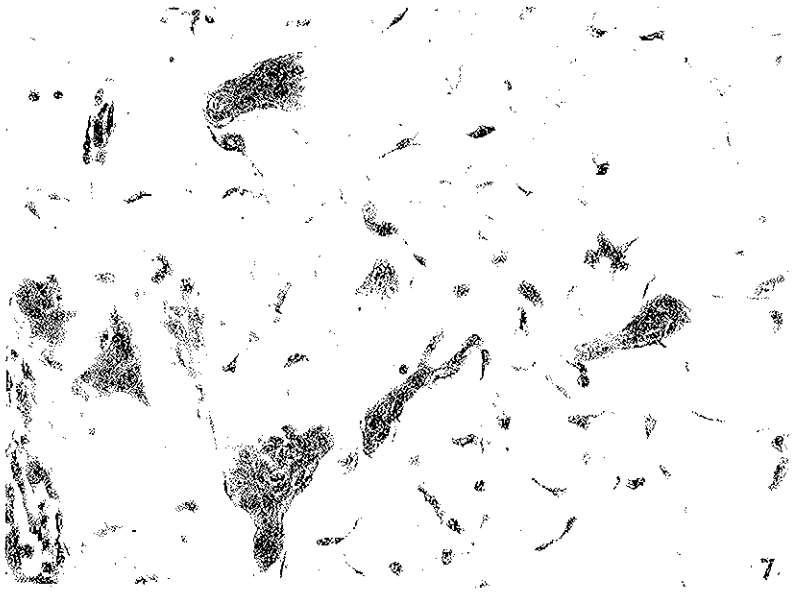


FIG. 7 — Muscular elements induced by heart ribonucleoprotein (× 385).

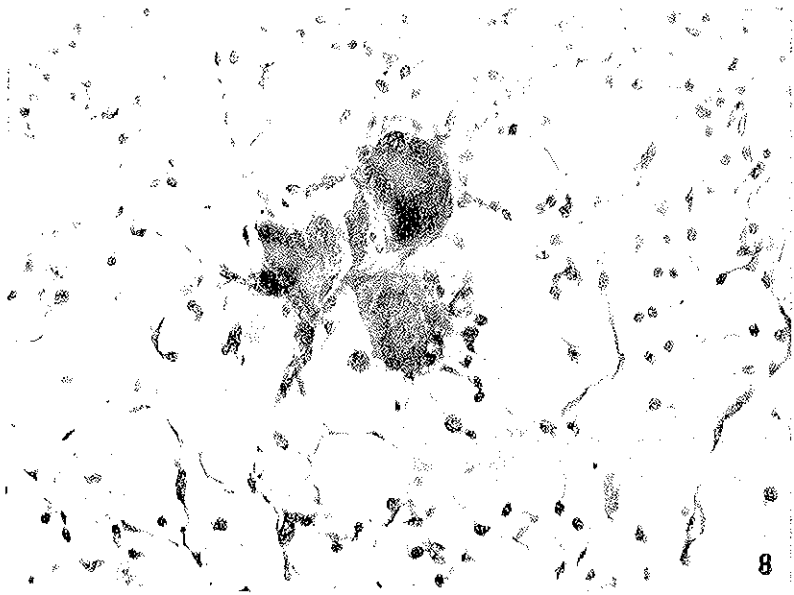


FIG. 8 — Muscular elements induced by skeletal muscle ribonucleoprotein (× 385).

obtain, from cells evidently not yet differentiated, cells in some way corresponding to the origin of the ribonucleoproteins.

After having imbibed a piece of albumen with chick RNP, we heated it in order to denaturate the RNP, putting the whole thing in a test tube in boiling water for 10 minutes before grafting it in the chorio-allantois. The amount of induced tissues was enormously reduced and only few or no cells were transformed.

By extracting ribonucleoproteins from frog heart, it is possible to induce, in chick chorio-allantois, the same kind of tissue induced by chick heart ribonucleoproteins while ribonucleoproteins from frog liver induce the formation of the same cells induced by chick liver ribonucleoproteins. The organ differentiation is, therefore, associated to a ribonucleoprotein differentiation, which makes these proteins similar in the same organ of the chick and of the frog.

In order to progress in the study of this process, we tried to graft in the chorio-allantois: microsomes, ribosomes and ribonucleoproteins extracted from microsomes. In all these experiments very little (some times with microsomes) or no induction at all was observed.

At present we (RANZI, GAVAROSI, CITTERIO and NECCHI) are trying to graft purified total RNA prepared following LASCOV (1) in the chorio-allantois but from the results up to this moment no specific action RNA is revealed.

(1) Organs were collected from animals after a 48 hour starvation and immediately frozen by means of liquid air or dry ice. This material (1000 g) was then ground in a mortar and a mixture of 300 ml EDTA 10^{-4} M pH 8 and of 300 ml 90% freshly redistilled phenol was added to this powder. The suspension thus obtained was homogenized in a glass homogenizer for 2 min, gently stirred for 1 hour at 20°C and finally centrifuged at 10.000 × g for 3 min at 0°C. The upper aqueous phase was eliminated without disturbing the interfacies with the underlaying phase. The remainder, phenol containing, was extracted once with 200 ml 10^{-4} M EDTA pH 8 and then centrifuged at 10,000 × g for 3 min at 0°C. The upper aqueous phase was added to the former and centrifuged at 10,000 × g for 20 min at 0°C in order to eliminate phenol. To these so-called aqueous phases NaCl was added up to 3 M final concentration. This solution was then centrifuged at

REFERENCES

[1] BON F.W., *Electrophoretical and ultracentrifugal investigation of the structural changes of α-crystallin during development.* « Pubbl. Staz. Zool. Napoli », *32*, 104 (1961).

[2] CIGADA LEONARDI M., *Ancora sulle variazioni di alcune frazioni proteiche nel corso dello sviluppo del riccio di mare e della rana.* « Ist. Lombardo (Rend. Sc.) », B, *92*, 453 (1958).

[3] CITTERIO P., MALDACEA L. and RANZI S., *On the presence of actin and myosin in the electric organ of the Torpedo.* « Pubbl. Staz. Zool. Napoli », *29*, 434 (1957).

[4] NIU M.C., *Thymus ribonucleic acid and embryonic differentiation.* « Proc. Nat. Ac. Sci. », *44*, 1264 (1958).

[5] RANZI S., *Early determination in Development under Normal and Experimental Conditions: in Tyler, von Borstel and Metz.* « The Beginnings of Embryonic Development », AAAS, p. 291, 1957.

[6] RANZI S., *Les changements protéiques au cours du développement embryonnaire et larvaire.* « Année Biol. », *33*, 523 (1957).

[7] RANZI S., CITTERIO P. e SAMUELLI C., *Proteine e antigeni durante l'accrescimento embrionale e larvale di due specie di rospi (Bufo bufo e Bufo viridis) e dei loro incroci.* « Acta Embryol. et Morph. Exp. », *3*, 65 (1960).

[8] RANZI S., *The proteins in the embryonic and larval development.* « Advances in Morphologenesis » *2* (in press) this paper is a complete review of the research on protein differentiation during growth and determination).

[9] RANZI S., CITTERIO P. and SAMUELLI C., *Molecular Ontogenesis and Molecular Genetics in the Crosses of Toads.* « Experientia », *17*, 16 (1961).

[10] RANZI S., GAVAROSI G. e CITTERIO P., *Ribonucleoproteine e citodifferenziamento.* « Ist. Lombardo (Rend. Sc.) », B, *94*, 254 (1960).

[11] RANZI S., GAVAROSI G. and CITTERIO P., *Cytodifferentiation Induced by Ribonucleoproteins.* « Experientia », *17*, 395 (1961).

6000 × g for 3 min. The precipitate containing denatured proteins, glycogen and phenol crystals was discharged. After an overnight storage at 0°C the remaining phenol crystals were eliminated by rapid filtration. The filtrate was stored for 1-3 days at 0°C as long as it became cloudy and then centrifuged at 10,000 × g for 5 min in order to precipitate the ribonucleic acid as a gel. This was dissolved in 15-50 ml 10^{-4} M EDTA pH 8 and dyalized against 2 liters 10^{-3} M NaCl twice renewed during 18 hours (The tubes used for the dyalisis were always washed with 10^{-2} M EDTA pH 8 during 2 or 3 days). If some precipitate appeared, it was discharged. The solution was then lyophilized. The lyophilized powder was then dissolved in Gey fluid.

DISCUSSION

CHANTRENNE

Do you know whether your material comes from the isolated nuclei or from the cytoplasm?

RANZI

Our material comes from the entire organ, that is cell nucleus and cytoplasm.

TUPPY

Have the induced cells and tissues been characterized morphologically only, or enzymatically too?

RANZI

It is possible to detect glycogen in the cells induced by liver ribonucleoproteins. Consequently an enzymatic indnction is present too.

A HYPOTHESIS CONCERNING THE FOLDING OF DNA

ALEXANDER RICH

Department of Biology, Massachusetts Institute of Technology, Cambridge, Mass. (U.S.A.)

ABSTRACT

Several polymers are now know which form a condensed phase by folding back and forth upon themselves with a constant folding length. In this form they produce flat lamellae which are stacked in layers to form a three dimensional crystal. Materials which crystallize in this fashion include not only structurally simple polymers such as polyethylene, but, also hydrogen bonding polymers such as nylon and helical polymers such as polypropylene. The polymer chains characteristically fold back and forth with a length of 70-200A°, depending upon the temperature and other factors. Results obtained recently from an X-ray diffraction study of crystals of the T7 bacterial virus provide some evidence which supports the hypothesis that the DNA inside the virus may also be folded back on itself with a spacing near 105A°. It is also possible to interpret some of the X-ray diffraction results on DNA-histone as indicating a similar folding arrangement with a folding period of 110A°. The implications of this hypothesis are discussed with respect to virus and chromosomal structure.

Although we know a great deal about the detailed molecular structure of deoxyribonucleic acid (DNA), we know comparatively little about the manner in which the DNA molecule is folded to form the compact units which are seen in viruses and chromosomes. There is at the present time relatively little experimental data on this subject. Nonetheless, the problem is a very important one and is being worked on in many laboratories. The purpose of this paper is to present a hypothesis that sets forth one way in which DNA may form a condensed phase which might he found in certain chromosomes or in some DNA containing viruses. The hypothesis stems from experimental work in the field of polymer chemistry carried out during the past few years. This work has made us aware of a remarkable mechanism leading to the formation of a stable configuration in which polymer molecules fold back and forth to form flat lamellae which can build up into crystalline aggregates.

Crystals of Folded Polymers

In 1955, Keller [1] was studying single crystals which were prepared from linear polyethylene by cooling a dilute solution of this material in xylene. In the electron microscope he observed the formation of a numher of very thin lamellar crystals which were lozenge shaped. These flat crystals had a number of terrace steps on them and by shadowing with metal he could measure the step height. Remarkably, he found that they all had the same height, close to 100 A°. This represented an unusual type of regularity to arise from a linear polyethylene solution, since the solution was composed of molecules which were heterogeneous in length. Indeed the molecular weight of the polyethylene was such that the molecules were thousands of angstroms long. This paradox was compounded by the observation that the molecular axis of the molecule was perpendicular to the flat, lozengeshaped crystal. This was

shown by allowing the electron beam to go through the crystal and registering the electron diffraction pattern. The pattern revealed that the direction of the molecules, the *c*-axis, was perpendicular to the plane of the crystal and that the *a* and *b* axes of the unit cell were along the short and long diagonal of the lozenge. Thus the growing face of the crystal was (110).

These results were somewhat paradoxical and could only be expiained by assuming that the molecules folded back and forth to make a layer or sheet about 100 A° in thickness. Thus the crystal was built up by stacking several of these sheets to build the observed terraces, as shown schematically in figure 1. When this interpretation was initially presented in 1955, it generated some scepticism, but further work, carried out by KELLER and several other investigators, has served to confirm the correctness of his initial interpretation. These investigations [2, 3] have revealed many other aspects of this type of polymer folding, and several features are worth discussing here.

Since the initial discovery, it has been demonstrated that this type of folded lamellar phase is not unique for polyethylene, but rather is found in a very broad class of polymer

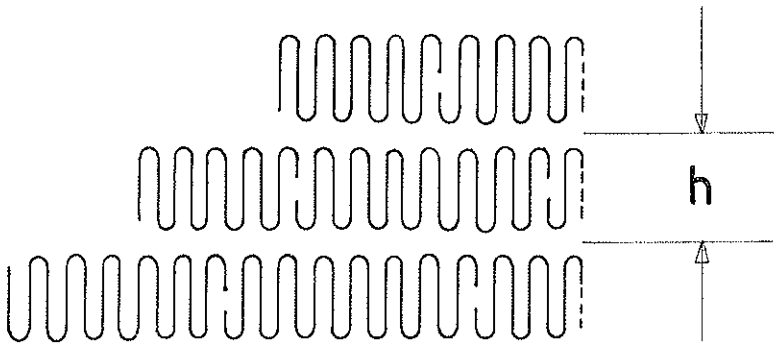


FIG. 1 — A diagram illustrating the folding of polymer chains to form a lamellar sheet. The sheets of constant height (*h*) are piled up to build a three-dimensional lattice.

molecules. In Table I several different types of polymers are listed which have been demonstrated to form a condensed phase, based upon the back and forth folding of polymer chains. The group is fairly heterogeneous. Some of the molecules are very simple, such as polyethylene or its helical homologue, polytetrafluoroethylene. On the other hand some of the molecules have fairly complicated side chains such as poly-4-methyl-pentene-1, or cellulose triacetate. Some are helical such as polypropylene. In addition a large number of hydrogen honding polymers are included such as all of the nylons in Table I. The polymers which form these folded phases are not structurally simple; in many cases the chains have a polarity. Because of the rapidly increasing number of polymers which are being discovered in this folded form, there is every indication that the phenomenon is a general feature of polymer interactions.

TABLE I (*)

LIST OF POLYMER MOLECULES WHICH FORM FOLDED LAMELLAE

Polyethylene (linear and branched)	Polyesters, such as Poly1, 10-decanol-succitate, and others
Polypropylene	
Polybutene	Polyhexamethylene terephathalate
Poly-4-methyl-pentene-1	Polyoxymethylene
Nylon 6	Polyethylene oxide
Nylon 7	Polydioxolane
Nylon 8	Polytetrafluoroethylene
Nylon 6-6	Cellulose triacetate
Nylon 6-10	Polychlorotrifluoroethylene
Isotactic Polystyrene	

(*) Data found in references 2, 3, 4, 5, 6 and also from P GEIL (personnal communication).

Much work has been done on characterizing the step height in the folded phase [3, 4, 5]. This aspect of the subject has best been approached through low angle X-ray diffraction studies. If, for example, one gently filters a solution containing the flat lamellar crystals, they will form a thin pad which can be readily removed from the filter paper. This sample can be mounted for low angle X-ray diffraction studies. The diffraction patterns show a strong reflexion near 100 A° which corresponds to the step height which was first seen in the electron microscope. The reflexion is not a powder ring, but rather has an orientation perpendicular to the flat crystals. Many of the polymer molecules which have been discovered in a folded, condensed phase of this type have been detected in the course of low angle X-ray diffraction studies taken from fibers of the material. Although most of the molecules in the fiber are fully unfolded and oriented with their molecular axes parallel to the fiber, some of them persist in the form of microcrystalline inclusions which maintain the folded arrangement. This means the polymer molecules have not fully unravelled to form part of the fiber. The relevance of this will be noted in our discussion of fibers prepared from DNA-histone preparations.

Several observers have shown that the height of the folding period in the polymer is temperature dependent [3, 4, 5]. For example, if one prepares polyethylene single crystals at different temperatures from an xylene solution, the step height is found to be a sensitive function of the temperature. Materials which are crystallized near room temperature have a step height of 92 A°, while those crystallized at 70°C have a step height of 110 A°. Crystal formation at 90°C produces a step height of 140 A°. In a general way it is easy to understand why the folding period would depend upon the crystallization temperature. The stability of the configuration may be discussed in terms of two positions, the linear interior and the turn. In the interior of the fold, the polyethylene molecules are stabilized by VAN DER WAALS interaction between adjoining polymer

molecules. On the other hand, in the fold where the molecule is turning about, there is a loss of the VAN DER WAALS stabilization and, because the packing may be less efficient in that region, the fold may be regarded as a destabilizing site. Thermal agitation increases as the temperature is raised, and this produces a decrease of the energy stabilization due to the VAN DER WAALS interaction in the interior of the fold. Consequently at a higher temperature it would require a longer fold period to stabilize the extended portions relative to the destabilization of the turns at the ends. A qualitative explanation of this type is the agreement of the observed facts. Some attempts have also been made to derive quantitative theories to explain the folding mechanism [7].

The step height dependence upon temperature is not unique with polyethylene, but is also characteristic of the other polymers listed in Table I.

THE FOLDING OF DNA

The necessity for an organized systematic folding of DNA is most clearly seen by a simple consideration of the geometry of the bacterial viruses. The T2 virus head which contains the DNA is about 1100 A^0 long and has a hexagonal cross section with about 800 A^0 between opposite flat faces. If we assume that the membrane has a thickness of about 100 A^0, we can calculate the internal volume of the head as approximately 2.3×10^8 cubic angstroms. The DNA which has been isolated from the T2 virus has a molecular weight of 120 million and is in the form of a single continuous strand. The molecular volume of this molecule is 1.9×10^8 cubic angstroms and if it were closely packed together in hexagonal array, it would occupy 2.1×10^8 cubic angstroms. Hence, even if we ignore the minor components such as polyamines and the internal virus protein we can see from these figures that the

DNA cannot be found inside the viral head in a more or less random form, since there appears to be very little wasted volume.

We do not have evidence at the present time which shows that the DNA inside the bacterial virus head is in a folded form stacked together in lamellae. Nonetheless, we do have some suggestive evidence which is compatible with that interpretation. We have recently initiated a programm of X-ray diffraction studies on DNA containing virus [8]. These studies have suggested that a significant part of the DNA in T2 virus has its molecular axis parallel to the tail. More recently we have developed a very gentle technique for preparing virus samples, with a minimum of manipulative procedure. This is done by using a fine capillary tube with a funnel shaped ending which will take about 0.2 ml. of a virus solution. The capillary is filled with the virus solution containing DNA-ase to digest any disrupted virus. The capillary tube is then mounted in a centrifuge and spun to centrifuge the virus down to the end of the capillary tube, where it gently packs to form a gel. The overlying solution is removed and the capillary sealed and mounted directly in the X-ray beam. Using this mild procedure we have been able to obtain some preliminary low angle diffraction photographs from T2 phage which show a moderately strong reflexion in the region near 92 A° [9]. This is a very recent finding and one which needs to be elaborated by further studies. In particular it will be necessary to demonstrate that the reflexion does not arise from the tail structure. Nonetheless it a compatible with some other observations which have been made on the T7 virus.

In the course of studying T7 virus gels, it was observed that the virus would crystallize and produce a diffraction pattern which shows two axes at right angles to each other with spacings of 570 A° and 470 A° [8]. Along the 570 A° direction almost a dozen orders can be seen, and from the photograph it is possible to visualize an envelope of the diffracted intensity

in the central region. Of interest here is the fact that a large band of intensity is seen in the region of 105 A°. In the absence of a complete analysis of the structure of the virus and the mode of packing of the viral particles in the unit cell, it is impossible to draw any definite conclusions from a finding of this type. Nonetheless it is possible that these strong reflexions result from a folded lamellar arrangement of DNA within the virus with a folding period near 105 A°.

Both of these findings on the T2 and T7 virus suggest that strong reflexions occur in a region near 100 A°. These should be regarded as suggestive at the present time, since reflexions of this type could, of course, arise from other structures.

Several investigators have recently carried out low angle X-ray diffraction investigation on DNA-histone preparations. LUZZATI and his coworkers [10, 11, 12] have found evidence for three phases existing in nucleo-histone solutions, each of which occurs at a different concentration. These phases are characterized by a series of low angle diffraction maxima. One of the phases has the same diffraction pattern as that produced by the fowl erythrocyte nucleus. It shows sharp reflections at a 110 A° spacing, as well as the next two higher orders at 55 and 37 A°. This may be evidence for a lamellar structure within the chromosome, and it is possible that these are related to the reflexions described above in the virus preparations.

WILKINS and his coworkers [13, 14] have been studying oriented fibres of nucleo-histone, and they find in addition to the DNA X-ray diffraction pattern, some meridional reflexions at 55 A° and 35 A°. These arise at reflexions which are very close to the second and third orders of the spacings which LUZZATI observed, and it is possible that these represent the higher orders of a 110 A° repeat. Further work will have to be done on this before there is enough evidence to conclude that a folded lamellar organization of this type occurs in nucleo-histone fibers. However, it is worth noting that a somewhat similar situation is seen in some of the synthetic polymers which

are known to form lamellar sheets by back and forth folding. In those cases most of the molecules are extended fully to form the backbone of the fiber, but some of them persist in the folded state and give rise to the low angle reflexions. This might also be the case in the work on the nucleo-histone fibers.

Is There a Regular DNA Folding in Chromosomes?

Very little is known about the structure of chromosomes. A large variety of electron microscope studies have been carried out, but many of them suffer from the fact that staining methods are not well developed, and artifacts occur quite frequently. However, a relatively constant feature of certain types of chromosomes, is the so-called « core » structure which was found by Moses [15]. The chromosomal core consists of a series of lamellar-like structures found in the center of the chromosome which run more or less parallel to its axis. In figure 2 we have reproduced an electron micrograph from Moses' work on the primary spermatocyte prophase of the crayfish [15]. The left side of figure 2 shows a longitudinal section from the core of the chromosome and there we see a series of 6 or 8 dense lines, 150 A° wide, separated by clear regions about 250 A° across. In the cross section shown at the right of figure 2, it can be seen that the flat regions seem to be about 500 A° in width and piled on top of each other. The lamellar construction of the core seems to extend through the entire length of the chromosome. Although nothing is known about the molecular organization of the chromosomal core, nonetheless, it is possible that these lamellae could be formed by DNA molecules folded back and forth on each other. Core structures such as these have been seen in a variety of chromosomes including plant cells.

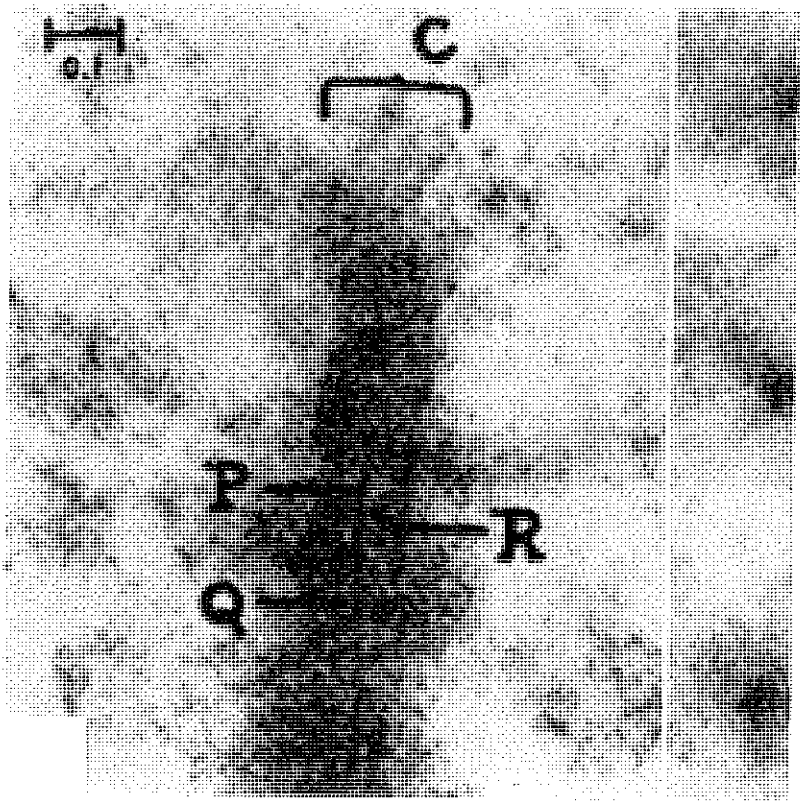


FIG. 2 — Electron micrograph of a section through a chromosome in the primary spermatocyte of the crayfish during prophase. The « core » is shown in longitudinal sections at the left. P points to one of the 150 A° wide structures which is surrounded by less dense areas (Q and R) about 250 A° wide. On the right are shown cross sections of the core. From MOSES (15).

The purpose of this article was to present a hypothesis. The hypothesis specifically is that the DNA molecule is capable of forming a condensed flat lamellar formed by folding back and forth upon itself with a period in the range 90 A° to over 200 A° which is similar to the fold periods which have now

been found in a variety of synthetic polymers. The reasons for believing the DNA molecule capable of this type of folding arrangement are in part due to the external simplicity of the molecule as it has a rounded cross section. The molecule is also simple in that it does not have directionality, so that from the outside the molecule looks the same in adjacent folds. It is obvious that the DNA molecule cannot form such a structure by itself. It is a negatively charged polyanion, and there must be cationic species present to keep the molecule in this configuration. In the case of the bacterial viruses, the polyamines may act as cross-linking stabilizing agents, in addition to cations. In the chromosomal organization the positively charged histones would play a corresponding stabilizing role. We might inquire about the organization of the DNA molecule at the bend. In order to bend back upon itself in this fashion, two or possibly three base pairs have to be broken. However, the liberated bases could still lie almost perpendicular to the axis of the molecule, and thus retain some of the stabilizing van der Waals energy due to stacking of the bases.

Even though a folding arrangement of this type may be found in some viruses, it is quite obvious that it need not be found in all viruses. In a similar way there may exist some chromosomal structures, such as the core, in which the DNA has this folded lamellar form. However, it is quite unlikely that all of the DNA in the chromosome is in this form, since many of the chromosomal loops which bave been studied in the electron microscope do not show any evidence which at all suggests a lamellar structure.

It is apparent that while this idea is plausible, it is nothing more than an attractive hypothesis at the present time and further experimental work will have to be carried out before we can evaluate its merits.

REFERENCES

[1] KELLER A., *J. Polymer Sci.*, *17*, 291 (1955).
[2] — *Phil. Mag.*, *9*, 1171 (1957).
[3] KELLER A. and O'CONNOR A., *Disc. Farad. Soc.*, *25*, 114 (1958).
[4] STATTON W.O. and GEIL P.A., *J. Applied Polymer Sci.*, *3*, 357 (1960).
[5] GEIL P.H., *J. Polymer Sci.*, *47*, 65 (1960).
[6] GEIL P.H., SYMONDS N.K.J. and SCOTT R.G., *J. Applied Physics*, *30*, 1516 (1959).
[7] FISCHER F.W., *S. Naturforschung*, *41*-a, 584 (1959).
[8] NORTH A.T.C. and RICH A., *Nature*, *191*, 1242 (1961).
[9] These are experiments carried out with L. KATZ.
[10] LUZZATI V. and NICOLAIEFF A., *Compt. Rend.*, *248*, 1426 (1959).
[11] LUZZATI V. and NICOLAIEFF A., *J. Mol. Biol.*, *1*, 127 (1959).
[12] NICOLAIEFF A., MAZEN-KNOBLOCK A., VENDRELY R. and LUZZATI V., *Compt. rend.*, *248*, 2805 (1959).
[13] WILKINS M.H.F., ZUBAY G. and WILSON H.R., *Trans. Farad. Soc.*, *55*, 497 (1959).
[14] WILKINS M.H.F., ZUBAY G. and WILSON H.R., *J. Mol. Biol.*, *1*, 179 (1959).
[15] MOSES M.J., *J. Biophys. Biochem. Cytol.*, *2*, 215 (1956).

[18] *Rich* - pag. 12.

DISCUSSION

JACOB

Would you comment on the electron micrographs of chromosomes. At what stage of the division cycle were they taken? Are they perpendicular or parallel to the axis of the chromosome?

RICH

I believe that the « core » is oriented parallel to the axis of the chromosome. In the photograph which I showed, the long striations go along the length of the chromosome. The core can be seen only when the chromosome is in a condensed form, and they are not visible at all during interphase.

LIPMANN

Do you know anytbing about folding of s-RNA?

RICH

Some evidence has been presented by ZUBAY and BROWN which suggests that the s-RNA is folded back on itself to form a DNA-like structure. Hower, if this were a regular configuration, it would imply that the base ratios in an individual s-RNA molecule are complementary, or approximately so, considering that the ends may not be completely matched. The recent work of HOLLEY shows that the base ratios of two different s-RNA's are far from complementary, however, so we must be cautious in interpreting the suggestive structural work.

RUBIO HUERTOS

You have said that you have crystallized some T_7 phages. Are these crystals formed by the complete virus particles or only DNA?

RICH

The entire phage particle is crystallized, including the protein shell. We would like to obtain information about the way that the DNA is packed inside the virus protein shell, but this has not proceeded very far as yet.

CHARGAFF

What is known about particles of DNA in the polyhedral viruses which are considered to be crystalline?

RICH

It is interesting that electron microscope studies of viruses and virus infected cells have shown that two types of polyhedral aggregates can be visualized. One of these is a « ghost », that is, the protein shell by itself which has the polyhedral form even when it is not filled with nucleic acid. However, other aggregates consist of the DNA alone before they are covered with the protein coat. Thus, both the DNA and the protein can assume a polyhedral configuration independent of each other. The DNA is not « folded » by the protein, but is able to fold up independent of its enclosing shell. However, as far as I am aware, nothing is known abont the way in which the DNA is folded inside the virus.

ULTRASTRUCTURE OF CRYSTALLINE INCLUSIONS DETERMINED BY TWO PLANT VIRUSES

M. RUBIO HUERTOS

Instituto « Jaime Ferrán » de Microbiologia, Madrid (España)

Different plant viruses are known to produce intracellular inclusions in the cells of their hosts. Only a few among them produce both amorphous X bodies and crystalline inclusions at the same time. Petunia ringspot belongs to the latter group of viruses and the study of the morphology, formation and histochemical reactions of its inclusions have already been done by the author and it was found that the crystalline inclusions develop from the X bodies as in the case of tomato aucuba mosaic described by SHEFFIELD.

In this paper we study the formation and ultrestructure of the crystalline inclusion bodies by ultrathin sections in the electron microscope and its comparison with those well known of tobacco mosaic virus.

The virus was maintained on young *N. tabacum. N. glutinosa, Petunia hybrida and Vicia faba* plants in the green house, and periodically epidermical strips of the infected leaves were stained with 1 per 100 phloxin without previous fixation, mounted in water and observed under the microscope. The inclusions both amorphous and crystalline are stained in a very bright red color as well as the nucleoli; the nuclei are stained in pink color. Chloroplasts and plastids remained unstained.

For electron microscopy, pieces from *Vicia faba* infested leaves were fixed for 3 hours in veronal-buffered 1% osmic acid according to STRUGGER. Others were fixed for 48 hours in the same fixation. The preparations were dehydrated in ethyl alcohol, embedded in methyl (2 parts) butyl (8 parts) methacrylate and sectioned with a glass knife in a Porter-Blum microtome. The sections from the blocks fixed in osmic acid during 3 hours were treated for 3 hours with aqueous 1% uranyl acetate to enhance the contrast (STRUGGER) [10] sections fixed during 48 hours in osmic acid gave very good contrast without previous uranyl acetate treatment. Some of the sections were observed directly under the electron microscope, others were previously treated with benzene to dissolve the methacrylate and shadowed with rodium-platinum.

The observations were made in a Siemens-Elminskop I electron microscope.

RESULTS

The crystalline inclusions vary in shape from hexagonal plates to octahedral forms being predominant the latter all belonging to the cubic system (fig. 1) and often are embedded in the amorphous inclusions (fot. 1, 2). Their size is approximately 3-10 microns μ. They are very refringent and are stained in a bright red color by 1 per 100 phloxine without previous fixation, after fixation they can be stained with DELAFIELD'S hematoxilin, light green etc. They are not found in epidermical tissues of healthy *Nicotiana tabacum, N. glutinosa, Petunia* and *Vicia faba* leaves.

Neither type of inclusions was soluble in water or in 1 × 100 HCL and whereas the amorphous inclusions are also not soluble in 1 × 100 OHNa the crystalline ones are.

3 different specific features were observed in the ultrathin sections of infected cells: bundles of virus particles in the

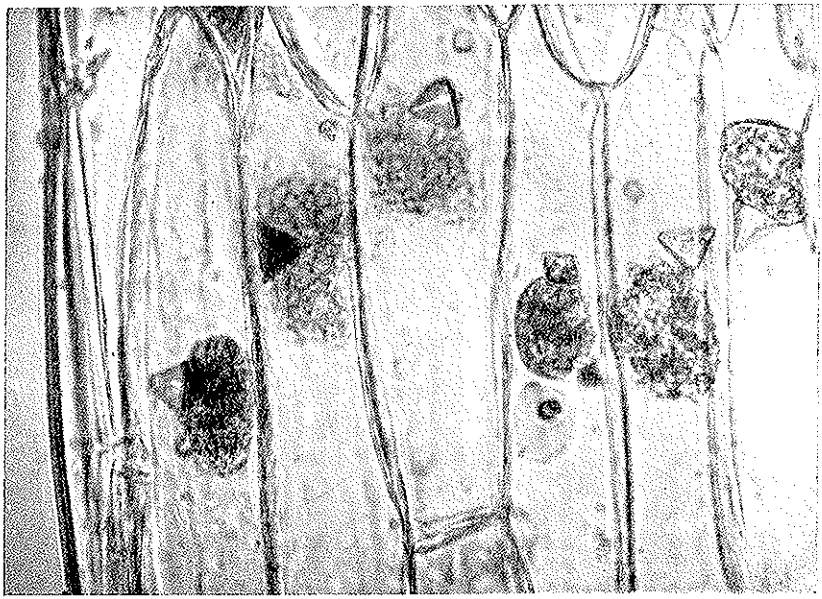


FIG. 1 — Amorphous and crystalline inclusions in *V. faba* × 700.

cytoplasm; the crystalline inclusions and the amorphous × bodies.

Sections of the crystalline inclusions showed that they are formed by virus particles arranged in a tridimensional order. Longitudinal sections showed that the particles aggregated end-to-end forming a continuous particle length with no unions shorwing, as it happens in Tobacco mosaic virus crystalline plates which clearly showed a layer structure (STEERE, 1956; WEHRMEYER, 1957); cross sections of the crystalline inclusions showed a definite honeycomb pattern formed by the hexagonal cross-sections of the individual virus particles arranged in tridimensional order (figs. 2, 3, 4).

The amorphous inclusions were formed by a material different from the normal cytoplasmic material. In some instances this material seem to be aggregated in a pattern

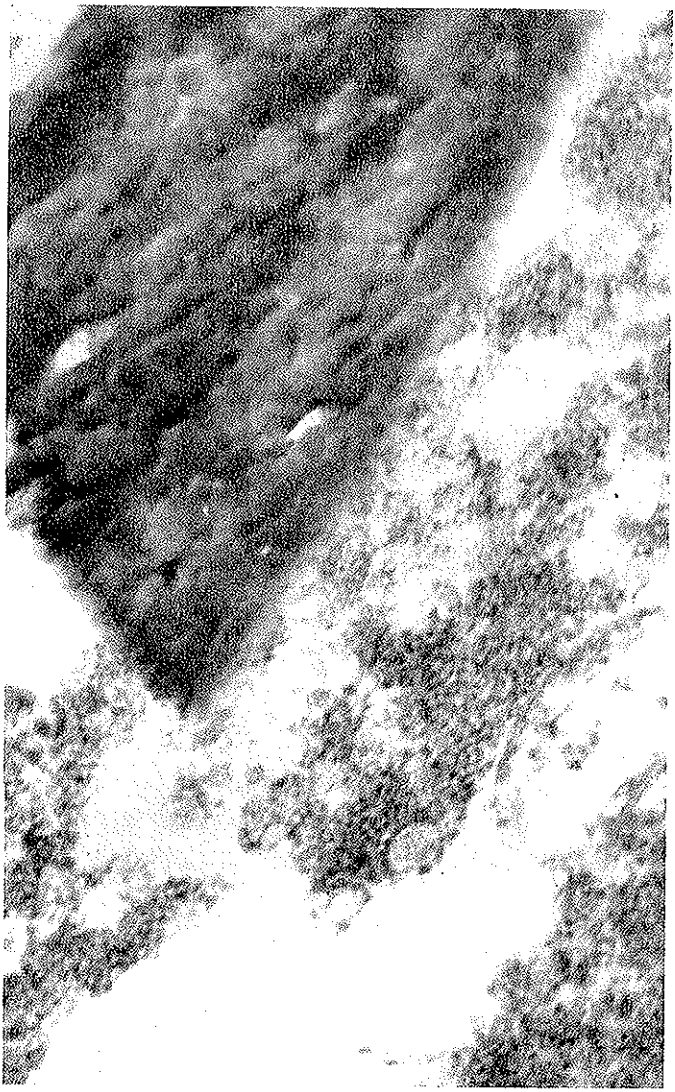


FIG. 2 — Longitudinal section of an X body and a crystallaine inclusion.

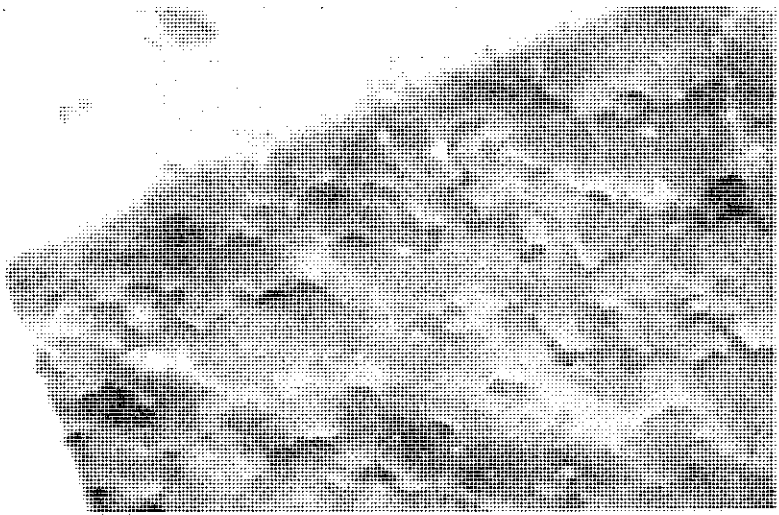


FIG. 3 — Cross-section of a crystal showing the honey-comb pattern.

which suggests it have some structural relationship with virus particles (fot. 2), although no complete virus particles were found in it.

Mitochondria were always found associated with X bodies.

Some plant viruses can crystallize from purified solutions and some of them can also form espontaneous crystalline inclusions in the cells of their hosts.

Crystalline inclusions induced by T.M.V. were shown by STEERE and WILLIAMS and later by us using a different technique to be formed by virus particles. Petunia ringspot virus crystals are also formed of virus particles although the crystalline structure differs in both viruses. Also electron mycroscope studies of virus crystals obtained from purified virus solutions

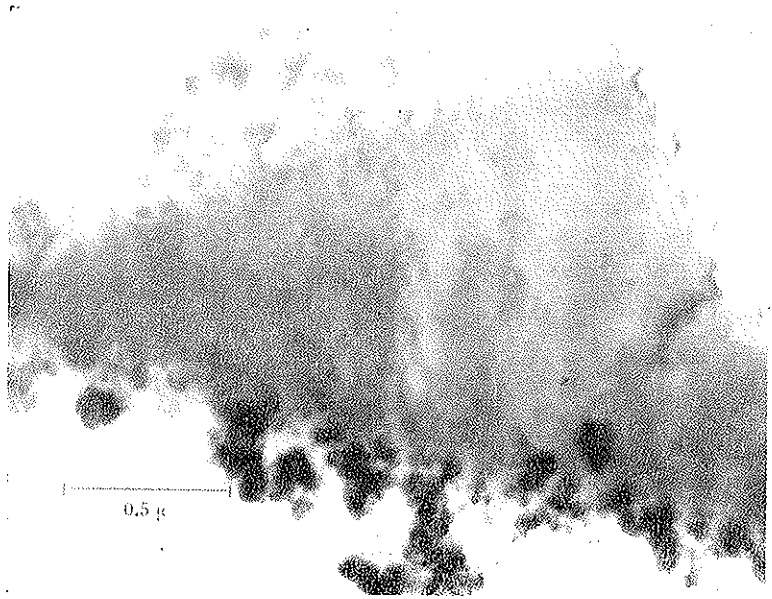


FIG. 4 — Longitudinal section of a Crystal. The virus particles do not *show their unions as in T.M.V.*

have been made by WYCKOFF et al. and they have shown beautiful pictures of the arrangement of the virus particles to form the crystals; so both types of crystals are formed by virus particles arranged in a tridimensional order.

If we consider that viruses may be living entities because they reproduce and mutate, it seems odd that they are able to form crystals as any simple non living protein, but the active part of virus nucleo protein is the nucleic acid as it has been proved by FRAENKEL-CONRAT et al. the protein being just a sort of a protective shell or coat; this shell confers to tbe virus particle all the external physicochemical properties of a pure protein, crystallization among them, provide they are in a suitable environment, and the behaviour of resting virus part-

icles is the same as it formed by protein alone. The case of turnip yellow mosaic is a very clear one.

When a T.Y. virus purified suspension is let to stand for 4h or longer separation in two layers can be observed, the particles in both layers observed being alike in their shape and size and indistinguishabled in the electron microscope Analysis of the particles of both layers showed that the bottom layer is formed by nucleoprotein whereas the top layer is formed only by protein. Infectivity tests showed that only the particles from the bottom layer were infective. Later it was demonstrated by electron microscopy that the protein particles had an internal cavity which in the particles from the bottom layer is filled by the nucleic acid. Both types of particles can form identical crystals separately. So it seems it is protein responsible for crystallization.

Although we have not been able to observe a complete sequence of events leading to the formation of virus particles from the amorphous material collected in the bodies there is some evidence of morphological relationship between the od of material in the X bodies and the complete virus particles which later will form the crystals, so it seems that the X bodies are one of the « situs » for virus multiplication.

We want to enfasize the interesting features of this virus that would perhaps permit, with more detailed work, to determine in the future the sequence of virus particle formation.

REFERENCES

McWHORTER F.P., *Isometric crystals produced by Pisum Virus 2 and Pha-*

McWHORTER F.P., *Isometric crystals produced by Pisum Virus 2 and Phaseolus Virus 2*. « Phytopathology », *31*, 760-761 (1941).

RUBIO HUERTOS M., *Estudios sobre inclusiones intracelulares producidas por virus en las plantas*. « Microbiol. Espanola », *3*, 207-232 (1950).

— *Rapid extraction of intact crystalline inclusions from the cells of plants infected with tobacco mosaic virus*. « Nature London », *174*, 313 (1954).

— *Origin and composition of cell inclusions associated with certain tobacco and crucifer viruses*. « Phytopathology », *46*, 553-556 (1956).

RUBIO HUERTOS M. and VAN SLOGTEREN D.H.M., *Light and electron microscopy of X-bodies associated with broad-bean mottle virus and Phaseolus Virus 2*. « Phytopathology », *46*, 401-402 (1956).

— *Nota previa sobre un virus de tipo « ringspot » encontrado en petunia hybrida vilm*. « Microbiol. Española », *4*, 325-329 (1959).

SHEFFIELD F.M.L., *The formation of intracellular inclusions in solanaceous hosts infected with aucuba mosaic of tomato*. « Ann. Appl. Biol. », *18*, 471-493 (1931).

STEERE R.L. and WILLIAMS R.C., *Identification of crystalline inclusion bodies extracted intact from plant cells infected with tobacco mosaic virus*. « Amer. J. Bot. », *40*, 81-84 (1953).

STRUGGER S., *Die Uranylacetat-Kontrastierung für die elektronenmikroskopische Untersuchung von Pflanzenzellen*. « Naturwiss », *43*, 357 (1956).

WEHRMEYER W., *Darstellung und Strukturordnung eines Tabakmosaikvirus-Einschlusskörpers in der Zelle*. « Naturwiss. », *44*, 519 (1957).

DISCUSSION

SCHRAMM

For X-ray investigations, good virus crystals are desirable. Unfortunately, the crystalline inclusions of TMV cannot be isolated from the cell, since the crystals disappear just after touching the cell. Is it possible to isolate the crystalline inclusions in your case?

RUBIO HUERTOS

Yes. These Petunia virus crystals are insoluble in water, they are soluble only in weak alkali. It is possible to break-up the cell walls with fine needles without damaging the crystals. We have isolated several crystals in this way for electron microscope observation. I think it would be very interesting to study them in the X-rays and electron microscope diffraction.

CHANTRENNE

Does anybody know why cells which are invaded by a virus, plant virus, eventually die? I mean is it an inhibition of protein synthesis, of nucleic acid synthesis, or something similar to what is known in the phage?

RUBIO HUERTOS

Normally cells infected with a plant virus do not show pathological changes besides the formation of intracellular inclusions and

diminution of the number and size of chloroplasts; only in some cases, depending of the virus and host plant, some cells became necrotic and die. This processus is not well understood. More than a competition between the virus and the cell for basic materials, it seems to be a reaction of the host with production of toxic substances.

[19] *Rubio Huertos* - pag. 10

MOEGLICHKEITEN ZUR GEWINNUNG GENETISCH AKTIVER NUCLEINSAEUREN

G. SCHRAMM
Max-Planck-Institut für Virusforschung, Biochemische Abteilung,
Tübingen (Deutschland)

Die erste biologisch aktive Nucleinsäure wurde 1944 VON AVERY [1] und Mitarbeitern hergestellt. Sie gewannen aus Pneumococcen eine Desoxyribonucleinsäure (DNS), die imstande war, bei anderen Pneumococcenstämmen eine genetisch stabile Änderung hervorzurufen. Durch enzymatische Abbauversuche wurde bewiesen, dass tatsächlich die DNS und nicht etwa Beimengungen von Ribonucleinsäure (RNS) und Protein genetisch aktiv sind. Die Arbeit schliesst mit dem Satz: « Nucleic acids must be regarded as possessing biological specificity the chemical basis of which is as yet undetermined ». Es wird auch die Möglichkeit diskutiert, dass es sich um ein « transmissible mutagen » handeln könnte. Sie wagten noch nicht, den weiter gehenden Schluss zu ziehen, dass die DNS die Erbsubstanz sei. Heute wissen wir, dass dies tatsächlich der Fall ist. Es darf als weitgehend gesichert gelten, dass bei allen Organismen die Gene aus DNS bestehen. Bei den Viren dagegen existieren solche, bei denen DNS, und solche, bei denen RNS der Träger der genetischen Information ist. Der sicherste Beweis für die genetische Wirkung der Nucleinsäuren ist stets die Reindarstellung und die Untersuchung der reinen Substanzen nach dem klassischen Beispiel VON AVERY.

Im folgenden Referat möchte ich daher die verschiedenen Möglichkeiten zur Gewinnung genetisch aktiver Nucleinsäuren schildern, wobei ich neben Methoden, die sich bereits praktisch bewährt haben, auch solche erwähnen möchte, die vielleicht erst in Zukunft Bedeutung gewinnen werden.

Darstellung genetisch aktiver Ribonucleinsäuren

Die erste genetisch aktive RNS wurde aus Tabakmosaikvirus (TMV) hergestellt und zwar unabhängig voneinander von zwei Arbeitskreisen. Fraenkel-Conrat [2] und Mitarbeiter benützten zur Darstellung ein Netzmittel, Dodecylsulfat. In unserem Laboratorium [3] wurde die Phenolmethode entwickelt. Durch zahlreiche Kontrollversuche konnten wir beweisen, dass allein die RNS für die Infektion und damit für die genetisch bedingte Struktur des Virus verantwortlich ist [4].

Die Phenolmethode ist besonders einfach anzuwenden und wurde daher zur Darstellung zahlreicher biologisch aktiver Nucleinsäuren benützt. Das Wesentliche bei dieser Methode ist, dass durch das Phenol die Enzyme sofort aus der wässrigen Phase entfernt und denaturiert werden, so dass die empfindliche Nucleinsäure nicht angegriffen wird. Doch auch bei wiederholter Phenolextraktion ist es nicht leicht, die letzten Enzymspuren zu entfernen. Dies lässt sich viel leichter erreichen, wenn man nach dem Vorschlag von Fraenkel-Conrat [5] Bentonit zusetzt, das ein starkes Adsorptionsvermögen für Ribonucleasen hat. Mit Hilfe der Phenolmethode konnte auch die verhältnismässig niedermolekulare s-RNS (lösliche RNS), die für die Incorporierung der Aminosäuren in die Proteine notwendig ist, rein dargestellt werden [6]. Die genetische Information für den Aufbau der Proteine scheint jedoch nur in den Nucleinsäuren höheren Molgewichts enthalten zu sein. Hierfür sprechen vor allem die Versuche an Virusnucleinsäuren, deren Molgewicht stets über 1.10^6 liegt, meist bei etwa 2.10^6. In der

folgenden Tabelle 1 sind die RNS-haltigen Viren zusammen. gestellt. Mit der Phenolmethode konnte aus fast allen diesen Viren die Nucleinsäure in infektiöser Form dargestellt werden [Literatur 7, 8, 9, 10, 11, 12]. Eine Ausnahme bilden nur die Viren der Influenza-und Parainfluenza-Gruppe. Auch hier wurden zum Teil positive Ergebnisse mitgeteilt, doch konnten diese nicht reproduziert werden [13]. Die Gründe, warum die Phenolmethode bei den Myxoviren versagt, sind noch nicht klar. Bei den erst kürzlich aufgefundenen RNS-haltigen Bacteriophagen [14] liegen noch keine Untersuchungen vor.

Die infektiöse RNS ist als Vorstufe des kompletten Virusteilchens anzusehen. ENGLER und SCHRAMM [15] zeigten, dass nach Infektion mit TMV der Anstieg an infektiöser RNS früher erfolgt als die Bildung der fertigen Virusteilchen. Ähnliche Versuche liegen bei zoopathogenen Viren vor.

Die Arboviren sind gegen Phenol in der Kälte beständig. Um das Protein von der Nucleinsäuren abzutrennen, muss eine Temperatur von etwa 40°C gewählt werden. WECKER [16] fand, dass aus infizierten Geweben auch mit kaltem Phenol eine infektiöse RNS extrahiert werden kann. Es kommen also zwei Formen von infektiösen Partikeln vor, die kompletten Virusteilchen, die gegen kaltes Phenol beständig sind, und eine Form, die die RNS frei oder nur lose gebunden enthält. Ähnliche Beobachtungen wurden von SANDERS [17] und von BELLET [18] et al. am Mäuse-Encephalomyocarditis-Virus gemacht. Die vegetative Form des Virus besteht wahrscheinlich nicht aus freier Nucleinsäure, sondern aus einem Ribonucleoprotein, das als Virosom bezeichnet wird.

Bei gewissen Pflanzenviren kommen Formen vor, bei denen das infektiöse Prinzip ausschliesslich aus einer derartigen freien oder locker gebundenen RNS besteht. Beispiele hierfür sind das von SCHLEGEL [19] untersuchte Cucumber Mosaic Virus und das von BRANDENBURG und SÄNGER [20] untersuchte Tabak Rattle Virus. Bei diesen Viren lassen sich durch Extraktion mit Puffer nur sehr wenig infektiöse Partikel gewin-

nen, da in diesem Medium durch die Enzyme das infektiöse Prinzip zerstört wird. Durch Extraktion mit Phenol lassen sich dagegen wirksame RNS-Präparate herstellen, da in diesem Fall die Wirkung der Enzyme ausgeschaltet wird.

Diese Versuche zeigen, dass der wesentliche Bestandteil der Viren eine genetisch aktive Nucleinsäure ist, die im Verlaufe des Infektionsprozesses an verschiedene Proteine gebunden wird. Was wir für gewöhnlich als Virusinfektion bezeichnen, ist der Übergang einer mehr oder weniger geschützten Nucleinsäure in eine fremde Wirtszelle. Jede übertragbare Nucleinsäure könnte demnach als Virus bezeichnet werden. So berichteten kürzlich HOTCHKISS und OTTOLENGHI [21] von spontanen Transformationen. Beim Zerfall von Bakterienzellen könnten die frei werdenden Desoxyribonucleinsäuren in andere Bakterienzellen eindringen. Es besteht also kein grundsätzlicher Unterschied zwischen normaler und Virusnucleinsäure. Wir müssen uns vor Augen halten, dass in der Biologie alle von Menschen getroffenen Definitionen künstlich sind und das Leben selbst keine scharfen Grenzen kennt.

DARSTELLUNG GENETISCH AKTIVER DESOXYRIBONUCLEINSÄUREN

Bei den Organismen ist es noch nicht gelungen, das gesamte Genom in Form eines DNS-Moleküls zu isolieren. Es ist auch keinesfalls sicher, ob das gesamte Genom in einem durchlaufenden DNS-Faden enthalten ist, wie dies z.B. bei der RNS des TMV der Fall ist. In den Transformationsexperimenten lassen sich jeweils nur bestimmte aktive Teile des Genoms in Form von DNS-Molekülen nachweisen. Nur bei kleineren DNS-haltigen Viren gelingt es, die gesamte genetische Substanz in Form isolierter DNS zu erhalten. Allerdings sind Molgewichtsbestimmungen und Charakterisierung der genetisch aktiven DNS noch nicht so vollständig wie bei den RNS-haltigen Viren. Zur Darstellung wurde meist die Phenolmethode benützt,

[20] *Schramm* - pag. 4

daneben auch die Hitzedenaturierung und die Behandlung mit Netzmitteln. Wie aus der Tabelle 2 hervorgeht, konnte aus Shope Papillom [22, 23] und aus Polyoma-Virus [24, 25] infektiöse DNS gewonnen werden. Bei den Adeno- und Pocken-Viren sind weder positive noch negative Resultate bekannt. GERSHENZON [26] berichtet, dass es nicht gelungen sei, aus dem Polyedervirus von Bombyx mori eine infektiöse DNS zu isolieren, doch konnte mit Phenol aus infizierten Raupen ein RNS-Präparat hergestellt werden, das die Produktion des kompletten DNS-haltigen Virus auslöst. Das bedeutet, dass RNS in der Lage ist, die Bildung von DNS in der Zelle auszulösen, doch sind vielleicht auch andere Interpretationen möglich.

Aus dem Phagen ∅X174 wurde von mehreren Seiten [27, 28, 29] infektiöse DNS isoliert, die sich interessanter Weise als Einzelstrang herausstellte und nur ein Molgewicht von $1{,}7.10^6$ besitzt [30]. Aus dem Lambda-Phagen [31] wurde ein DNS-Doppelstrang isoliert, der sich für Protoplasten als infektiös erwies. Bei den T_2-Phagen wurden ebenfalls Extrakte erhalten, die für Protoplasten infektiös sind, doch war die Isolierung reiner aktiver DNS bisher nicht möglich. Der Grund hierfür liegt wahrscheinlich in dem hohen Molgewicht der Phagen DNS. HERSHEY [32, 33] zeigte, dass schon geringe Scherkräfte, wie sie etwa beim Pipettieren auftreten, genügen, um die äusserst langen DNS-Moleküle zu zerbrechen. Diese Bruchstücke sind nicht mehr in der Lage, ein intaktes Phagenteilchen zu erzeugen, doch werden sie bei Recombinationsexperimenten in das Phagengenom miteingebaut [34]. Jedenfalls besteht kein Zweifel, dass DNS als Einzel- wie als Doppelstrang in Wirtszellen eindringen und dort genetische Wirkungen ausüben kann.

Chemische Änderungen an Nucleinsäuren

Durch Extraktion aus Viren lassen sich also genetisch aktive RNS und DNS als einheitliche Moleküle isolieren. Diese Reihe genetisch wirksamer Moleküle lässt sich noch dadurch erweitern, dass man auf chemischem Wege die isolierten Nucleinsäuren abändert. Um eine Mutation an Nucleinsäuren ausserhalb der Zelle zu bewirken, muss das Basenmuster abgewandelt werden, ohne dass die Kette gespalten wird, denn Gierer [35] zeigte, dass ein Kettenbruch stets zur Inaktivierung führt. Die Verhältnisse liegen hier also schwieriger als bei der Mutationsauslösung in der lebenden Zelle, wo Kettenbrüche wieder geheilt werden können. Schuster und Schramm [36-*a*] fanden, dass solche mutagene Änderungen in vitro durch Desaminierung mit salpetriger Säure möglich sind. Es gelingt mit diesem Reagenz, die Aminogruppen des Cytosins, Adenins und Guanins in Hydroxylgruppen umzuwandeln, ohne dass hierbei die Kette gespalten wird. Die ersten Versuche wurden an der Ribonucleinsäure des TMV durchgeführt. Eine Desaminierung je 3000 Nucleotide genügt, um das Molekül zu inaktivieren. Da das gesamte Molekül 6000 Nucleotide enthält, wurde vermutet, dass die Desaminierung bestimmter Nucleotide nicht zu einer Inaktivierung, sondern zu einer Mutation führt. Dies wurde dann experimentell von Gierer und Mundry [36-*b*] bewiesen. Um genetisch stabile Mutanten des TMV zu erzeugen, genügt ebenfalls die Abänderung eines einzigen Nucleotids. Die mutagene Wirkung der salpetrigen Säure beim TMV wurde inzwischen auch von anderer Seite bestätigt [37]. Auch bei den Viren der Poliomyelitis [38] und der Newcastle disease [39] sowie bei den DNS-haltigen Phagen [40] wurde nach Nitritbehandlung eine hohe Mutationsrate festgestellt. Es war nun interessant festzustellen, wie sich die Änderung des Basenmusters in den Nucleinsäuren in der Aminosäure-Zusammensetzung des davon abhängigen Proteins äussert. Das geeignetste Objekt

für derartige Untersuchungen ist wiederum das TMV, da hier die Aminosäurezusammensetzung des gesamten Proteins bekannt ist [41, 42]. Umfassende Versuche von WITTMANN [43] ergaben, dass bei einer grossen Anzahl der durch Nitrit erzeugten TMV-Mutanten die Aminosäurezusammensetzung des Proteins verändert ist. Jedoch wurden auch Mutanten gefunden, die sich biologisch vom Ausgangsstamm unterscheiden, jedoch keinen Unterschied in der Proteinzusammensetzung aufweisen. Hieraus kann man den Schluss ziehen, dass die RNS des TMV nicht nur die Synthese des Virusproteins kontrolliert, sondern auch noch andere Funktionen ausübt. Zur vollständigen Entzifferung des Nucleotidcodes [44] ist es notwendig, weitere chemische Reaktionen zu entwickeln, die möglichst spezifisch wirken, dh. jeweils nur ein bestimmtes Nucleotid verändern. In unserem Institut wurde von SCHUSTER [45] die Wirkung des Hydroxylamins untersucht. Dieses reagiert spezifisch mit den Pyrimidinresten, die unter Spaltung des Ringsystems aus der Nucleinsäurekette eliminiert werden. Die Reaktion ist insofern spezifisch, als Cytosin bei pH 6 etwa 30mal schneller reagiert als Uracil, während bei pH 9 die Verhältnisse umgekehrt liegen. Nür die Veränderung des Cytosins führt zu einer Mutation. Hydroxylamin wirkt auch bei ruhenden Phagen Teilchen mutagen [46]. FREESE [47] untersuchte daneben auch noch eine Reihe weiterer mutagemer Agentien bei Phagen.

Die Chemie ist also durchaus in der Lage, dem Biologen verschiedenartig modifizierte genetisch wirksame Nucleinsäuren zur Verfügung zustellen. Chemische Umsetzungen an der Nucleinsäuren, die zu einer Inaktivierung oder Mutation führen, haben auch praktische Bedeutung bei der Herstellung von Virusvaccinen [48].

Versuche zur chemischen Synthese von Nucleinsäuren

Es ist ein sehr verlockendes Ziel, neben der Abänderung der natürlichen Nucleinsäuren Verfahren zur Totalsynthese neuer Nucleinsäuren aus den Nucleotiden zu entwickeln. Natürlich bedeutet die Verknüpfung von vielen 1000 oder 100.000 Nucleotiden in einer ganz bestimmten Reihenfolge ein äusserst schwieriges Problem. Wird es dem Chemiker gelingen, diese Synthese, die die Natur ständig mit grosser Sicherheit vollführt, im Reagenzglase nachzuahmen? Mir scheint dieses Problem nicht so unlösbar, wie man im ersten Augenblick denken sollte. In der Zelle erfolgt die Synthese einer neuen Nucleinsäure nach einer Art Abdruckverfahren an der Oberfläche einer bereits bestehenden Matrize. Dieses Abdruckverfahren kann bereits bis zu einem gewissen Grade mit Hilfe spezifischer Enzyme im Reagenzglas nachgeahmt werden. Kornberg [49] und Mitarbeiter fanden, dass die Synthese hochmolekularer DNS mit Hilfe von DNS-Polymerase unter Benützung einer DNS-Matrize möglich ist. Die Zusammensetzung der neu gebildeten Nucleinsäure wird hierbei durch die Nucleotidsequenz der Matrize bestimmt [50]. Von Weiss und Nakamoto [51] und von Ochoa et al. [52] wurde ein Enzym gefunden, das in Gegenwart einer DNS-Matrize eine spezifisch zusammengesetzte RNS bildet und von Reddi [53] ein solches, das eine RNS-Matrize benötigt, um RNS zu synthetisieren. Wenn auch in keinem dieser Fälle bisher eine genetisch aktive Nucleinsäure gewonnen wurde, so ist dies vielleicht mehr methodisch als grundsätzlich bedingt. Auf jeden Fall zeigen diese Versuche, dass es auch im Reagenzglas gelingt, die Verknüpfung von Nucleotiden durch eine Matrize zu steuern. Die Synthese einer aus 6.000 Nucleotiden bestehenden RNS wäre unmöglich, wenn man gezwungen wäre, 6.000 einzelne Schritte hintereinander durchzuführen. Unter Benützung einer geeigneten Matrize sollte es aber grundsätzlich möglich sein, eine spezifische Polymerisa-

tion in einem Schritt durchzuführen, um eine genetisch aktive Nucleinsäure zu erhalten.

Die chemische Synthese der Nucleinsäuren lässt sich demnach in zwei Teilprobleme zerlegen. Zunächst muss eine Reaktion gefunden werden, die es gestattet, Nucleotide zu hochmolekularen Polynucleotiden zu kondensieren und zweitens muss geprüft werden, ob es auch gelingt, diesen Polymerisationsprozess durch Zugabe von Nucleotid-Matrizen zu beeinflussen. Zusammen mit meinen Mitarbeitern GRÖTSCH und POLLMANN haben wir in der letzten Zeit recht vielversprechende Erfolge erzielt [54]. Wir haben hiernach unser Ziel sogar noch etwas weiter gesteckt, indem wir versuchten, bei der Nucleosid-Synthese direkt von den Zuckern und den Basen auszugehen. Es gelang uns eine neuartige Nucleosid-Synthese auszuarbeiten und nach Phosphorylierung diese zu hochmolekularen Polynucleotiden zu kondensieren. Im Rahmen dieses Vortrags muss ich mich darauf beschränken, die wesentlichen Grundzüge darzustellen. Bei der Verknüpfung eines Zuckers mit einer heterocyclischen Base kommt es darauf an, die acetalische OH-Gruppe am C_1 durch einen geeigneten Substituenten so zu lockern, dass ein Ersatz dieses OH durch das N-Atom einer Base leicht vor sich geht. Für diese Aktivierung benützt die Natur Phosphat-haltige Gruppen. So erfolgt z.B. die enzymatische Synthese der Nucleoside entweder über Ribosyl-Phosphat oder Pyrophosphat. Es gelang uns nun, diese enzymatische Synthese nachzuahmen, indem wir Ribose mit Polyphosphorsäureester (PPE) behandelten. Hierbei bildet sich eine aktive Zwischenverbindung, die mit Adenin zu Adenosin umgesetzt werden kann. Durch Variation der Zucker und der basischen Komponenten lassen sich eine ganze Anzahl verschiedener Nucleoside darstellen (siehe Tabelle 3).

Bei diesen Reaktionen handelt es sicht also um die Aktivierung der Carbonylgruppe monomerer Zucker. Wir fanden nun, dass als Ausgangsmaterial auch hochomolekulare Kohlenhydrate mit freier Aldehydfunktion verwendet werden können.

Von CHARGAFF [55] und anderen wurde gezeigt, dass es in saurer Lösung gelingt, aus DNS die Purinreste zu entfernen und zu einer Apurinsäure zu kommen, in der also die Aldehydgruppen der DNS teilweise in freier Form vorliegen. Die Darstellung dieser Apurinsäuren wurde genau studiert und die Hydrolysenkonstanten sowohl für die glycosidische Bindung als auch für die Phosphatbindung quantitativ gemessen [56]. In der DNS erfolgt die Abspaltung der Purinreste etwa 120mal so schnell wie die Spaltung der Phosphatbrücken, so dass ein grosser Teil der Basen ohne Veränderung der Kettenlänge entfernt werden kann. Wird nun eine derartige Apurinsäure in Gegenwart von PPE mit Adenin behandelt, so werden die freien Aldehydgruppen wieder quantitativ besetzt. Durch Hydrolyse und anschliessende Charakterisierung der Produkte überzeugten wir uns davon, dass bei dieser Resynthese wieder die richtige sterische Konfiguration erhalten wird. Anstelle von Adenin können auch eine Reihe anderer Basen eingebaut werden. Um die Einbauraten zu messen, verwandten wir radioaktiv markierte Basen. Aus Tabelle 4 geht hervor, dass der Einbau der Pyrimidine langsamer vor sich geht als der der Purine, dass aber im Prinzip auf diesem Wege ein Ersatz der Purinreste durch Pyrimidinreste möglich ist. Es öffnet sich hier also ein Weg für die Herstellung einer grossen Anzahl verschiedenartiger Nucleinsäuren. Die Resynthese kann auch in Gegenwart von Wasser durchgeführt werden. Es sind zur Zeit Versuche im Gange, durch einen derartigen Basenaustausch mutagene Änderungen an verschiedenen Nucleinsäuren durchzuführen.

Mit Hilfe von PPE gelang uns schliesslich auch die Totalsynthese von hochmolekularen Nucleotiden aus monomeren Nucleotiden und Nucleosiden. Die Polymerisation der Nucleotide wurde schon in vielen Laboratorien versucht, doch wurden nur Oligonucleotide bis zu einem Polymerisationsgrad von 7-10 in sehr kleinen Ausbeuten erhalten [57, 58, 59]. Nachdem sich PPE bei der Aktivierung der acetalischen OH-Gruppe so bewährt hat, benützten wir dieses Reagenz auch zur Aktivie-

rung der sekundären Phosphatgruppen. Wir erwärmten die Nucleotide in Gegenwart von PPE für einige Stunden auf 50° und erhielten in guten Ausbeuten Polynucleotide mit einem Molgewicht von etwa 20-50.000 (siehe Tabelle 5). Es ist bemerkenswert, dass auch die empfindlichen Desoxyribonucleotide, wie z.B. die Thymidylsäure glatt polymerisiert werden können. Verschiedene Versuche sprechen dafür, dass diese auf chemischem Wege gewonnenen Polynucleotide den auf natürlichem Wege entstandenen Polynucleotiden zumindest sehr ähnlich sind. Auf elektronenmikroskopischen Aufnahmen der synthetischen Polyadenylsäure beobachtet man lange Fäden wie bei der natürlichen RNS. Es zeigt sich, dass die Polymerisation im wesentlichen linear verläuft und unter Umständen zu sehr hohen Molgewichten führen kann. Abbauversuche mit Enzymen zeigen, dass zumindest überwiegend die natürliche 3′-5′-Konfiguration erhalten wird. Die synthetischen Ribonucleotide werden durch Ribonuclease weitgehend abgebaut, soweit sie Pyrimidinreste enthalten. Synthetische Poly-Adenylsäure wird durch die von EGAMI [60] hergestellte Ribonuclease bis zu 70% gespalten. Da dieses Enzym sehr spezifisch ist und nur 3′-5′-Bindungen angreift, kann man schliessen, dass die Zahl der « falschen » Bindungen gering ist. Alle Präparate zeigen einen hohen Hyperchromie-Effekt, wie er für hochmoleculare Nucleinsäuren typisch ist. Mischungen der synthetischen Poly-Adenylsäure mit Poly-Uridylsäure geben eine starke Erniedrigung der Extinktion durch Bildung einer Doppelhelix, wie es von RICH et al. [61] für die enzymatisch hergestellte Polyadenylsäure und Polyuridylsäure beobachtet wurde. Es ist überraschend, dass mit dieser einfachen Methode synthetische Polynucleotide erhalten werden können, die den natürlichen sehr ähnlich sind. Ein wesentlicher Unterschied besteht allerdings darin, dass unsere synthetischen Produkte eine zufallsmässige Anordnung der Nucleotide zeigen, während die natürlichen eine spezifische Sequenz der Nucleotide besitzen.

Der nächste Schritt wäre nun zu prüfen, ob es möglich ist, durch Zugabe von Matrizen den Verlauf der Polymerisation zu lenken. Die ersten Vorversuche hierzu verliefen positiv. Es gelang uns zu zeigen, dass die Polymerisation von Uridylsäure zu Polyuridylsäure in Gegenwart von Poly-Adenylsäure bis auf das 10 fache beschleunigt wird. Wie in der Zelle begünstigt also die Anwesenheit des einen Strangs die Polymerisation des dazu komplementären. Wir hoffen also, ein einfaches Modellsystem in der Hand zu haben, mit dem wir die natürliche Entstehung der Nucleinsäuren wenigstens zu einem gewissen Grade nachahmen können. Es ist hervorzuheben, dass die von uns angewendeten Reaktionsbedingungen so einfach sind, das sie im Laufe der Erdgeschichte auch durchaus spontan eingetreten sein können. Durch sorgfältige kinetische Messungen wird es vielleicht möglich sein, eine experimentell gut fundierte Theorie über die Entstehung selbstvermehrungsfähiger Systeme in der Natur zu entwickeln.

TABELLE I

RNS-Viren	infektiöse RNS	Molgew. der RNS in 10^6
PFLANZENVIREN		
sphärische	+	2
stäbchenförmige	+	2
TIERVIREN		
Naniviren		
Maul u. Klauenseuche	+	2
Entero	+	2
Rous-Sarkom	+	(9)
Influenza	—	2
Parainfluenza	—	32
BAKTERIENVIREN		
RNS-Phagen	?	?

[20] *Schramm* - pag. 12

TABELLE II

DNS-Viren	infektiöse DNS	Molgew. der DNS in 10^6
Papillom	+	3
Polyoma	+	?
Adeno	?	?
Pockenviren	?	200
Polyederviren . . .	—	40
Bakteriophagen X 174	+	1,7
T_2	—	130

TABELLE III

Synthese von Nucleosiden mit Polyphosphorsäureester

Ribose + Adenin ⟶	Adenosin (+ 2′, 3′-Adenosinphosphat
Desoxyribose + Adenin ⟶	Desoxyadenosin
Fuctose + Adenin ⟶	Fructosyladenin
N-Acetylglucosamin + Adenin ⟶	N-Acetylglucosaminyladenin

TABELLE IV

Incorporation von $_{14}C$-Basen in Apurinsäure mit Polyphosphorsäureester bei 37°C in Dimethylformamid

Base	Einbau %	Zeit h
Adenin	96	24
Guanin	40	24
Orotsäure	40	24
Thymin	5	48
Uracil	8	48

TABELLE V

Synthese von Polynucleotiden

	S_{20}	$[\eta]_v$	MG	Hyperchromie
Poly-A	2,5	25	21.000	37%
Poly-C	2,0	23	15.000	16%
Poly-G	2,9	28	28.000	28%
Poly-U	4,0	35	50.000	20%
Poly-T	2,3	24	18.000	47%

[20] *Schramm* - pag. 14

LITERATURVERZEICHNIS

[1] AVERI O.T., MACLEOD C.M., MCCARTY M., *J. Expt. Med., 49,* 137 (1944).

[2] FRAENKEL-CONRAT H., SINGER B., WILLIAMS R.C., *Biochim. Biophys. Acta, 25,* 87 (1957).

[3] SCHUSTER H., SCHRAMM G., ZILLIG W., *Naturforschung, 11-b,* 339 (1956).

[4] GIERER A., SCHRAMM G., *Naturforschung, 11-b,* 138 (1956); *Nature, 177,* 702 (1956).

[5] FRAENKEL-CONRAT H., SINGER B., TSUGITA A., *Virology, 14,* 54 (1961).

[6] HOAGLAND M.B., ZAMECNIK P.C., STEPHENSON M.L., *Biochim. Biophys. Acta, 24,* 215 (1957).
MONIER R., STEPHENSON M.L., ZAMECNIK P.C., *Biochim. Biophys. Acta, 43,* 1 (1960).

[7] SCHUSTER H. in *The Nucleic Acids III, Herausgeber Chargaff u. Davidson,* Academic Press, New York, 1960.

[8] KOZLOFF L.M. in *Annual Review of Biochemistry, 29* (1960). *Biochemistry of Viruses.*

[9] HERRIOTT R.M., *Science, 134,* 256 (1961).

[10] LACOUR P.u.J., HAREL J., HUPPERT J., *Natl. Cancer Inst., 24,* 301 (1960).

[11] CRAWFORD L.V. u.E.M., *Virol, 13,* 227 (1961).

[12] WEIL R., BENDICH A., SOUTHAM C.M., *Proc. Soc. Expt. Biol. Med., 104,* 670 (1960).

[13] Vergleiche BENEDICT A.A., LEE D., WEST B., *Nature, 188,* 911 (1960).

[14] LOEB T., ZINDER N.D., *Proc. Natl. Acad. Sci., 47,* 282 (1961).

[15] SCHRAMM G., ENGLER R., *Nature, 181,* 916 (1958).
ENGLER R., SCHRAMM G., *Nature, 183,* 1277 (1959).

[16] WECKER E., *Virology, 7,* 241 (1959).

[17] SANDERS F.K., *Nature, 185,* 802 (1960).

[18] BELLET A.J.P., BURNESS A.T.H., *Nature, 190,* 235 (1961).

[19] SCHLEGEL D.E., *Virology, 11,* 329 (1960).

[20] BRANDENBURG E., SÄNGER H.L., *Naturwissenschaften, 48,* 391 (1961).

[21] HOTCHKISS R.D., OTTOLENGHI E., *Moskauer Kongress,* 1961.

[22] WATSON J.D., LITTLEFIELD J.W., *J. Mol. Biol., 2,* 161 (1960).

[23] ITO Y., *Virology, 12,* 596 (1960).

[24] SMITH J.D., FREEMAN G., VOGT M., DULBECCO R., *Virology, 12,* 185 (1960).

[25] WEIL R., *Virology, 14,* 46 (1961).

[26] GERSHENZON S.M., KOK I.P., VITAS K.I., DOBROVOLSKAYA G.N., SKURATOVSKAYA I.N., *Moskauer Kongress,* 1961, S. 92.

[27] GUTHRIE G.D., SINSHEIMER R.L., *J. Mol. Biol., 2,* 297 (1960).

[20] *Schramm* - pag. 15

[28] HOFSCHNEIDER P.H., *Naturforschung, 15-b,* 441 (1960).
[29] SEKIGUCHI M., TAKETO A., TAKAGI Y., *Biochim. Biophys. Acta, 45,* 199 (1960).
[30] SINSHEIMER R.L., in *Nucleic Acids III, Herausgeber Chargaff u. Davidson,* Academic Press, New York, 1960.
[31] MEYER F., MACKAL R.P., TAO M., EVANS E.A., *J. Biol. Chem., 236,* 1141 (1961).
[32] RUBENSTEIN I., THOMAS C.A., HERSHEY A.D., *J. Mol. Biol., 3,* 458 (1961).
[33] BÜRGI E., HERSHEY A.D., *J. Mol. Biol., 3,* 458 (1961).
[34] VAN DE POL J.H., VELDHUISEN G., COHEN J.A., *Biochim. Biophys. Acta, 48,* 417 (1961).
[35] GIERER A., *Nature, 179,* 1297 (1957).
[36-a] SCHUSTER H., SCHRAMM G., *Naturforschung, 13-b,* 697 (1958).
[36-b] GIERER A., MUNDRY K.W., *Nature, 182,* 1457 (1958).
[37] SIEGEL A., *Virology, 11,* 156 (1960).
[38] BOEYÊ A., *Virology, 9,* 691 (1959).
[39] GRANOFF A., *Virology, 13,* 402 (1961).
[40] VIELMETTER W., WIEDER C.M., *Naturforschung, 14-b,* 312 (1959).
[41] ANDERER F.A., UHLIG H., WEBER E., SCHRAMM G., *Nature, 186,* 922 (1960).
[42] TSUGITA A., GISH D.T., YOUNG J., FRAENKEL-CONRAT H., KNIGHT C.A., STANLEY W.M., *Proc. Natl. Acad. Sci., 46,* 1463 (1960).
[43] WITTMANN H.G., *Virology, 12,* 609 (1960).
[44] GIERER A., *Naturwissenschaften, 48,* 283 (1961).
[45] SCHUSTER H., *J. Mol. Biol., 3,* 447 (1961).
[46] FREESE E., BAUTZ-FREESE E., BAUTZ E., *J. Mol. Biol., 3,* 133 (1961). BAUTZ E., FREESE E., *Fed. Proc., 20,* 352 (1961).
[47] FREESE E., *Übersicht, Moskauer Kongress,* 1961.
[48] SCHRAMM G., *Zentralblatt für Bakteriologie, 184,* 96 (1962).
[49] LEHMANN I.E., BESSMAN M.J., SIMMS E.S., KORNBERG A., *J. Biol. Chem., 233,* 163 u. 171 (1958).
[50] JOSSE J., KAISER A.D., KORNBERG A., *J. Biol. Chem., 236,* 874 (1961).
[51] WEISS S.B., NAKAMOTO T., *J. Biol. Chem., 326,* PC 18; *Proc. Natl. Acad. Sci, 47,* 694 (1961).
[52] OCHOA S., BURMS D.P., KRÖGER H., WARREN R.C., WEILL J.D., *Proc. Natl. Acad. Sci., 47,* 749 (1961).
[53] REDDI K.K., *Science, 133,* 1367 (1961).
[54] SCHRAMM G., GRÖTSCH H., POLLMANN W., *Angewandte Chemie, 73,* 619 (1961) und *74,* 53 (1962).
[55] TAMM CH., HODES M.E., CHARGAFF E., *J. Biol. Chem., 195,* 49 (1952).
[56] POLLMANN W., SCHRAMM G., Naturforschung, *160,* 673 (1961).
[57] MICHELSON A.M., *J. Chem. Soc.* London, 1959, 1371 u. 3655.
[58] TENER G.M., KHORANA H.G., MARKHAM R., POL E.H., *J. Amer. Chem. Soc., 80,* 6223 (1958).
[59] CRAMER F., *Angewandte Chemie, 73,* 49 (1961).
[60] EGAMI, *J. Biochem. Japan, 46,* 757 (1959).
[61] FELSENFELD G., RICH A., *Biochim. Biophys. Acta, 26,* 457 (1957).

[20] *Schramm* - pag. 16

DISCUSSION

FRAENKEL-CONRAT

The question is raised as to whether the RNA polymerisation reaction described by Professor SCHRAMM is sufficiently specific in terms of 3′-5′ linkages, and sequence, to justify its consideration as a means for the synthesis of genetically active RNA.

SCHRAMM

NIRENBERG and MATTHAEI (1) have shown that synthetic ribonucleic acids prepared by enzymes are active in controlling protein synthesis in cell free systems. Studies are in progress to find out whether our synthetic products are also active in this synthesis.

THEORELL

Is there a physical meaning of « molecular weight » of filamentous particles as judged only from sedimentation constants? The figures given in one of your tables seem to fix to globular proteins rather than to fibers of RNA. I think there is no correlation whatsoever between sedimentation constants and molecular weights in your case.

SCHRAMM

You have to consider that the shape of ribonucleic acid molecules is depending on the ionic strength of the medium. Therefore, we

(1) *Proc. Natl. Acad. Sci.*, *47*, 1580, a. 1588, 1961.

observe filaments in the electron microscope since these specimens are salt free, whereas the sedimentation and viscosity measurements are made in solutions of high ionic strength to avoid electrostatic effects. Under these conditions the ribonucleic acid molecules are coiled.

LIPMANN

Is any intermediary formed, i.e. can one isolate any phosphorylated intermediary? The action of the polyphosphate ester in polynucleotide synthesis from mononucleotides is that of a condensing agent regarding the end product, but one might assume that an intermediary phosphorylation product in 3′ position of the ribose forms and is displaced by the 5′ phosphate of the nucleotide that elongates the chain.

SCHRAMM

The nature of the intermediary phosphorylation product is not known as yet.

TISELIUS

I think that one must agree that the determination of molecular weights of all nucleic acids as a whole is an extremely difficult problem and that we do not know actually as well as most people believe the exact values for the molecular weights. I think that perhaps the most valuable method is the one by HASCHER and his collaborators, where they count the rays from the radioactive atoms. But, in connection with that work, I think they have just emphasized how unsatisfactory the previous methods — in particular the ultracentrifugations — have been in this field and also, as Professor THEORELL pointed out, with these filamentous molecules the centrifugation methods breaks them down some time.

SCHRAMM

It is true that the determination of the molecular weights is difficult for desoxyribonucleic acids having a molecular weight of

more than 10.10^6 (1). These rigid and elongated molecules can be disrupted by the shear forces arising during normal handling of the solutions. In that case autoradiography is useful. The molecular weights of our synthetic products are much lower and can be determined by the normal methods. This is evident from many studies of tobacco mosaic virus ribonucleic acid (molecular weight 2.10^6) and other well known substances.

RICH

I think the fact that these synthetic polymers combine, implies that if there are 3′ and 2′ linkages, they are unlikely to be there at random. They may be there in the form of a block copolymer, and the 3′-5′ block links are combining. One could very well have a situation where part of the chain is combining and part of it is not. But from a stereochemical point of view. I think it is rather unlikely that a random chain of 3′ and 2′ links would be able to combine.

SCHRAMM

This is a very interesting statement. Of course, besides combination experiments we have to apply other methods to find out the amount of unnatural 2′-3′ linkages.

(1) New literature: EIGNER, SCHILDKRAUT, and DOTY, *Biochimica et Biophysica Acta*, *55*, 13, 1962.

THE BIOLOGICAL ROLE OF PHOSPHOPROTEINS

N. SILIPRANDI

Istituto di Chimica Biologica dell'Università di Padova, Padova (Italia)

In spite of their ubiquity in tissues phosphoproteins have ben scarcely studied for their chemical structure and biological function. This was mainly due to an undervaluation of their biological role: their abundance in the growing tissues and in some peculiar aliments, such as milk and eggs, suggested in fact that they simply represent a storage for phosphorus and aminoacids. This hypothesis is not consistent, however, with the very fast turnover of P^{32} in phosphoproteins.

More interest in phosphoproteins arose from the BURNETT and KENNEDY's results [1], which efforded the first direct evidence of the enzymatic phosphorylation of a protein substrate. BURNETT and KENNEDY demonstrated that casein is phosphorylated by an enzyme present in liver mitochondria, the « protein phosphokinase », and they put forward the hypothesis that phosphoproteins may function by cyclic phosphorylation and dephosphorylation of their serine residues.

This was later demonstrated by RABINOWITZ and LIPMANN [2]. Using a protein phosphokinase free of myokinase and ATPase activity, these Authors were able to prove the reversibility of the reaction between ATP and phosphoproteins (phosvitin and casein). This important finding indicates that

the energy potential of phosphoryl groups in phosvitin should not be far below that in ATP. On the other hand the phosphate groups in phosvitin, as well as in other phosphoproteins, are esterified with monoester bond. This is supported by the finding that on pretreatment with phosphodiesterase from snake venom, subsequent incubation with prostate phosphatase releases the same amount of inorganic P as with the enzyme alone. Furthermore no acid labile P is present in phosvitin. On the other hand WILLIAMS and SANGER [3] found that phosphoserine is present in phosvitin and in casein in blocks of four or more units in a row. RABINOWITZ and LIPMANN attributed to this situation a particular importance in connection with the potential energy of the esterified phosphate.

All these observations concern phosvitin and casein, which can be regarded as foreign proteins in respect to the living cell machinery.

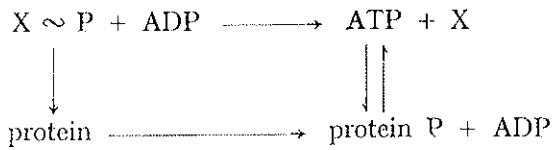
Therefore we have considered interesting to see if an endogenous cellular phosphoprotein retains in its structure the same phosphorylserine blocks as phosvitin and casein. In the positive case it would appear much more likely that these blocks are active sites in the cellular metabolism.

From the EHRLICH tumor cells, incubated with P^{32}, we have purified, as far as we were able, the phosphoproteins and we submitted their acid hydrolysis products to paper electrophoresis in 20% formic acid. At this pH (1.5) the carboxyl groups of peptides are uncharged and, unless they carry the strongly acidic phosphate groups, peptides are positively charged and move towards the cathode.

The electrophoretic pattern we obtained is very similar to that obtained with phosvitin and casein by WILLIAMS and SANGER [3] and by us in parallel experiments. Submitting the same hydrolysate to column electrophoresis, more consistent amounts of the same fractions were ohtained. After further hydrolysis some of these fractions proved constituted by blocks of several phosphorylserine residues. Hence the peculiar se-

quence in phosphoproteins from EHRLICH cells is very similar, if not identical, to that found in phosvitin and casein.

As regards the biological role of phosphoproteins it is of great importance the very recent work of AHMED and co. [4]. These Authors, who previously demonstrated that mitochondrial contraction, which is an enzymic event, and protein phosphorylation are closely associated, give evidence that phosphoproteins may accept P not only from ATP but also directly from some component of the oxidative phosphorylation system. In particular they showed that Benadryl, an antihistaminic drug, inhibits the mitochondrial phosphoproteins turnover by 60-80%, while it does not affect P^{32}-ATP exchange reaction nor oxidative phosphorylation process. A possible mechanism which explains such a result might be the following:



from which it appears that the phosphorylation of protein is partly dependent upon the protein phosphokinase reaction, as shown by BURNETT and KENNEDY [1] and by RABINOWITZ and LIPMANN [2], and partly connected with a shunt reaction alternative to the classical final reaction of the oxidative phosphorylation.

This metabolic process opens the opportunity of investigating if some cellular function, for instance nucleic acid turnover, is affected by Benadryl, which inhibits the phosphorylation of proteins without affecting the ATP production.

Furthermore this reaction gives additional value to the RABINOWITZ and LIPMANN observation [2] that phosvitin can accept phosphorus from ATP only if not completely dephosphorylated. It is probable that in the cell the phosphorylation

of protein through phosphokinase mechanism must be preceded by the shunt reaction shown in the above scheme.

Important contribution to the knowledge of metabolically active sites in phosphoproteins will derive by the resolution of phosvitin and casein fragments mixture, which was found by RABINOWITZ and LIPMANN [2] still active in phosphotransferase reaction and in testing the possible reversible phosphorylation, at mitochondrial level, with synthetic polymer of serine.

BIBLIOGRAPHY

[1] BURNETT G. and KENNEDY E.P., *J. Biol. Chem.*, *211*, 969, 1954.
[2] RABINOWITZ M. and LIPMANN F., *J. Biol. Chem.*, *235*, 1043, 1960.
[3] WILLIAMS J. and SANGER F., *Biochim. Biophys. Acta*, *33*, 294, 1959.
[4] AHMED K., JUDAH J.D. and GALLAGHER C.H., *Nature*, *191*, 1309, 1961.

DISCUSSION

KATCHALSKI

Professor SILIPRANDI, can you describe in some greater detail the composition and structure of the phosphopeptides which you have isolated?

SILIPRANDI

Phosphopeptides, we have isolated from phosphoproteins of the EHRLICH tumor cells, consist of phosphorylserine, glutamic and (or) aspartic acid in different reciprocal ratio; for iustance: ser P - ser P - glut - ser P. or ser P - ser P - ser P - aspart, and so on. In any case phosphorylserine is the main component of cellular phosphopeptides, being the other one of the two dicarboxylic acids.

Phosphorylserine blocks $(ser\text{-}P)_5$, or $(ser\text{-}P)_6$, previously found in typical phosphoproteins such as phosvitin and casein, have not been isolated from cellular phosphoproteins.

KATCHALSKI

In connection with the paper presented it might be of interest to recall some of the chemical properties of a seryl residue. In strong acid serine peptides are known to undergo an $N \rightarrow O$ acyl shift. In the presence of alkali C-phosphoseryl derivatives as well as other O-seryl derivatives can undergo a β-elimination reaction leading to the formation of the corresponding dehydroalanyl derivatives. Furthermore, our recent findings suggest that optically active seryl residue racemize in alkali more readily than hitherto assumed. These findings should be borne in mind while treating phospho-

proteins with acid or base, since the above chemical changes might affect also the secondary and tertiary structure of the protein.

The experiments of Professor LIPMANN as well as those of Professor SILIPRANDI indicate that some of the O-phospho-links present in phosphoproteins are energy rich bonds. Since monomeric O-phosphoserine is rather stable, it would be most desirable to determine the factors responsible for the activation of the phosphate in phosphoproteins such as phosvitin and casein. Neighbouring groups, as well as the secondary structure of the peptide backbone should be considered in this connection. Anyhow, I feel that at this stage it would be most desirable to investigate the chemical and physical chemical properties of various O-phosphoseryl containing polypeptides.

SILIPRANDI

It is of course much desirable to investigate the chemical and physical chemical properties of various O-phosphoseryl containing polypeptides, especially of those which, possibly, retain the property of a reciprocal transphosphorylation with ATP. Up to now it has not been possible, however, to isolate any definite phosphopeptide retaining such a property.

CHANTRENNE

There is one point, may be an important one, that I have missed What is the evidence you have for a direct phosphorylation of the protein without going through ATP? I didn't exactly see the point.

SILIPRANDI

There is not a direct evidence for an ATP independent phosphorylation of phosphoproteins. However, such a possibility appears reasonable on the basis of the finding of AHMED and coworkers [1] that Benadryl inhibits phosphoprotein turnover by 60-80% without appreciably affecting P^{32}-ATP exchange and the efficiency of oxidative phosphorylation.

---

[1] *Nature*, *191*, 1309, 1961.

THE MOLECULAR MECHANISM FOR TRANSCRIBING GENETIC INFORMATION (*)

S. SPIEGELMAN

Department of Microbiology, University of Illinois, Urbana, Ill. (U.S.A.)

I. Introduction: Nature of the Problem

The essence of the dogma, so acceptable to our time, can be briefly summarized by the following familiar diagram:

$$\text{DNA} \longrightarrow \text{RNA} \longrightarrow \text{Protein.}$$

Here, the arrows are meant to indicate flow of information. The presumption is that the genetic information coded in the base sequence of DNA is ultimately transcribed into the amino acid sequence of protein via a polyribonucleotide intermediary. It is the primary purpose of the present paper to focus attention on the first step in this chain of events.

The simplest imaginable transcription mechanism one can propose would suggest the synthesis of complementary RNA copies of the DNA. The RNA so formed would mimic the base ratio of its parental DNA providing one of the two following conditions was satisfied.

(*) This investigation was aided by grants in aid from The U. S. Public Health Service, National Science Foundation and The Office of Naval Research.

a) Both strands of the DNA are employed as templates for complementary RNA synthesis, or

b) The over-all base composition of the two complementary strands of DNA are the same or nearly so.

Comparison of the base compositions of the total RNA amongst bacteria of widely differing DNA compositions does not encourage the belief that a major proportion of the cellular RNA is analogous to its homologous DNA. Examination of the principal RNA molecular species reveals a surprisingly uniform picture amongst rather unrelated organisms. Three major components are universally found which are distinguishable by their sedimentation constants (23S, 16S and 4S). The 23S and 16S varieties have the same base compositions, are found in the ribosomes (KURLAND, 1960) and constitute approximately 85% of the total RNA of the cell. The 4S variety has, in general, a different base ratio and is found in the supernate or « soluble » fraction.

Thus far none of the three major components have exhibited any detectable correlation in their base ratio with the DNA composition of the cells from which they are derived. This is illustrated in Fig. 1 with respect to ribosomal RNA. Fig. 1 reproduces BELOZERSKY's (1959) relative plot of total RNA base composition (open circles) versus homologous DNA and includes as solid circles data on ribosomal RNA (23S+16S) obtained from *D. pneumoniae,* and *M. lysodeikticus* (WOESE, 1961), *E. coli, Ps. aeruginosa* and *B. megaterium* (HAYASHI and SPIEGELMAN, 1961). Although the information on ribosomal RNA is still scanty, the five samples available do derive from organisms which extend over a considerable portion of the GC range of DNA base composition. Nevertheless, the ribosomal (G+C/A+U) ratios fall within rather narrow limits and show no correlative tendency. Such data suggest that the comparatively slight correlation established by BELOZERSKY and his co-workers between bulk RNA and homologous DNA

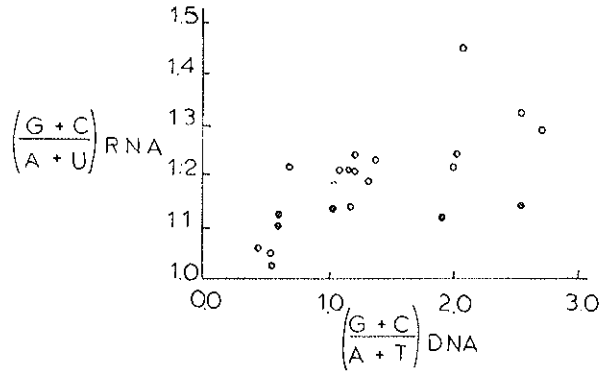


FIG. 1 — Relative Plot of Base Composition of RNA and DNA from the Same Organisms. Open circles are BELOZERSKY's (1959) data on total RNA from 22 bacteria. Closed circles are data on ribosomal RNA from five bacteria (WOESE, 1961; HAYASHI and SPIEGELMAN, 1961).

is to be ascribed to a small fraction of the cellular RNA. If, therefore, hope is to be retained for the existence of the simple transcription mechanism proposed above, it will have to be assumed that a quantitatively minor component of the RNA is involved in transferring information from DNA to the protein synthesizing machine. From the viewpoint of such considerations, the normal cell would not appear to represent ideal material for the initial search for the RNA which programs protein synthesis.

Several years ago my colleagues and I decided to turn our attention to the virus infected cell. The bacterial-virus complex possessed a number of obvious attractive advantages for the experimental analysis of the relations amongst DNA, RNA and protein. In the first place, the data accumulated over the past 15 years suggested that infection of bacterial cells with virulent viruses is followed by a restriction of protein and nucleic acid synthesis to a comparatively small class of macromolecules relevant to the formation of viral components. Se-

cond, and more important, were the experiments of VOLKIN and ASTRACHAN (1956) who demonstrated the incorporation of P^{32} into an RNA fraction which was distinguishable by its apparent base ratios from the bulk of the host's RNA. These ratios were determined by estimation of the relative P^{32} content of the corresponding 2'-3' nucleotides isolated from an alkaline hydrolysate of RNA from infected cells. Of greatest interest was the fact that the numbers so obtained appeared to mimic the DNA base ratio of the infecting particle. Here, we have the first suggestion of a complementary RNA which could play an informational role in the synthesis of a particular class of proteins. The intriguing implications of these observations made it mandatory that they be extended in several directions.

The existence of an RNA peculiar to virus infected cells is inferred from the VOLKIN-ASTRACHAN experiments solely on the basis of the distribution of isotope amongst the four nucleotides. In the absence of further information, it was possible to interpret these results in terms of unequal pool labeling and other conceivable complications. It was clearly of importance to provide independent evidence in support of their inference. The unequivocal proof of the existence of a « T2-specific » RNA would immediately generate a host of new experimental possibilities. As a convenient guide to the following discussion, we may set down the questions which stimulated us to the performance of the experiments we will describe.

1) Is it possible to effect a physical separation of the RNA formed subsequent to T2-infection and thus provide proof that « T2-specific RNA » exists?
2) What relation, physical or other, does the T2-specific RNA have to the normal RNA containing components of the cell?
3) Is the similarity of base ratios between T2-RNA and DNA a reflection of a more detailed identity iu terms of base sequences? More specifically, are they complementary?

4) Can one find evidence of naturally occurring RNA-DNA complexes predicted by the assumption that RNA complementary copies are the normal intermediates in the flow in information from the genome?
5) Does a similar type of complementary RNA also exist in normal non-infected celles?

It will be noted that the posing of each question makes sense only if the preceding one has been answered in the affirmative. We now undertake to describe the nature of the experiments which led us to these conclusions.

II. Physical Properties and Location of T2-Specific RNA

Nomura, Hall and Spiegelman (1960) undertook to obtain information pertinent to the first two questions raised. In addition to radioactive labeling, two techniques were used in this investigation to examine the nature of the RNA synthesized following infection with T2. One, involved zone electrophoresis through starch columns (Rotman and Spiegelman, 1954). The other, employed centrifugation through linear sucrose gradients developed by Britten and Roberts (1960). These procedures were used to examine isolated ribosomes and purified RNA prepared from ribosomes by the phenol procedure of Gierer and Schramm (1956). The data so obtained should detect differences in size, electrophoretic mobility and degree of association with ribosomes.

An examination was first made of the distribution of RNA synthesized subsequent to T2-infection. Here, care was exercised to use cell breakage and extraction conditions which insured integrity of the ribosomes. It was found that with such methods most (63%) of the newly synthesized RNA was ribosome bound. However, the association of the new RNA involved a linkage which was much more labile to low Mg

concentrations than that which characterizes normal ribosomal RNA.

This is illustrated in Fig. 2 which describes the results of electrophoresis experiments of ribosome fractions obtained from T2-infected cells. In these experiments, free ribosomal RNA was included as a marker and is represented by the leading peak of the optical density profile. It will be noted that at 5×10^{-3} M Mg^{++} the bulk of the newly synthesized RNA (P^{32} labeled) travels with the ribosomes although there is some apparent dissociation evident even at this level of magnesium. When the magnesium concentration is lowered to 1×10^{-4} M most of the radioactive RNA leaves the ribosomes and travels as free RNA. Under similar conditions, no detectable separation of normal ribosomal RNA from the nucleoprotein particles can be detected.

Most significantly, it will be noted from the lower half of Fig. 2 that the liberated labeled RNA moves faster than the added carrier *E. coli* RNA. This difference in electrophoretic mobility provided the first evidence that the RNA synthesized in a T2-infected cell was indeed a distinct entity which could be physically separated from the bulk of the pre-existent ribosomal RNA. This property was examined in greater detail with purified RNA isolated by the phenol method. Fig. 3 compares the results of an electrophoretic analysis of ribosomal RNA prepared from T2-infected and control cells subjected to P^{32} pulses of the same duration. It is clear that virtually all of the RNA synthesized in the T2-infected cell possesses a higher electrophoretic mobility. In the case of the non-infected cell, the major proportion of the RNA synthesized possesses a mobility very similar to that of the non-labeled pre-existent ribosomal RNA.

Another difference between normal and infected cells emerged when ribosomal RNA synthesized during short (3 min.) P^{32} pulses was examined by centrifugation in linear suc-

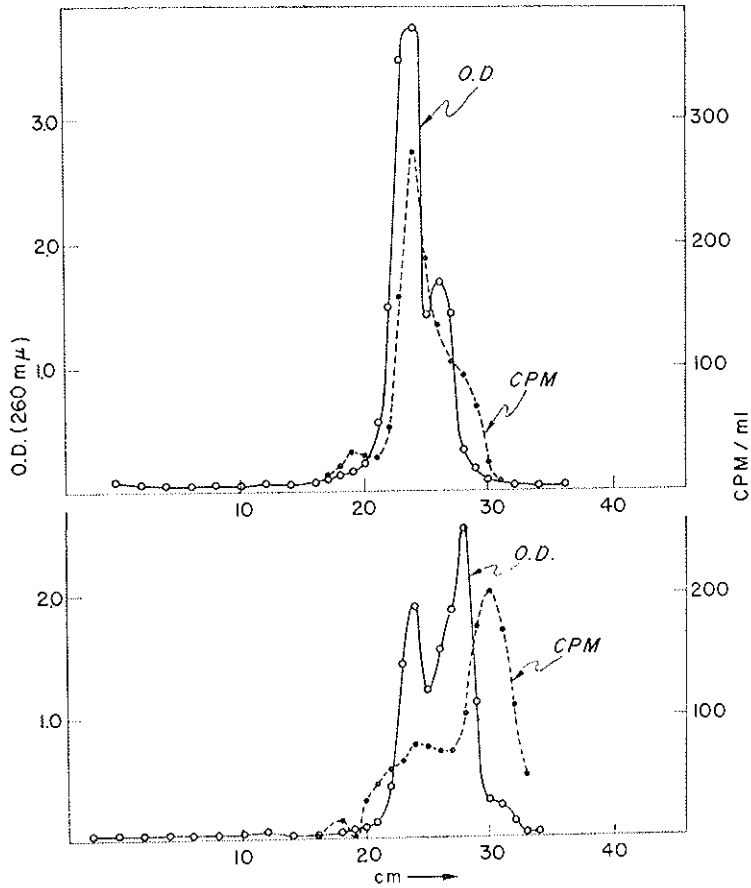


FIG. 2 — Starch column electrophoresis of the ribosome fraction of T2-infected cells. *a*) Electrophoresis in 0.03 M-tris, pH 7.8 with 5×10^{-3} M Mg^{++} after 3 hr. dialysis against this solvent. Duration of run - 13 hr. at 8.0 v/cm; 11 ma. *b*) Electrophoresis in 0.03 M-tris, pH 7.8, 1×10^{-4} M Mg^{++} after 24 hr. dialysis against this solvent. Duration of run - 13 hr. at 6.0 v/cm; 4 ma. Ribosomes used in this experiment came from T2-infected cells exposed to P^{32} between 5 and 7 min. after infection. *E. coli* particle RNA (prepared by the phenol method) was added as a marker just before electrophoresis. After elution of each fraction from starch OD at 260 mμ was measured; then the acid insoluble P^{32} which was solubilized by RNAase was counted (counts/min. curve).

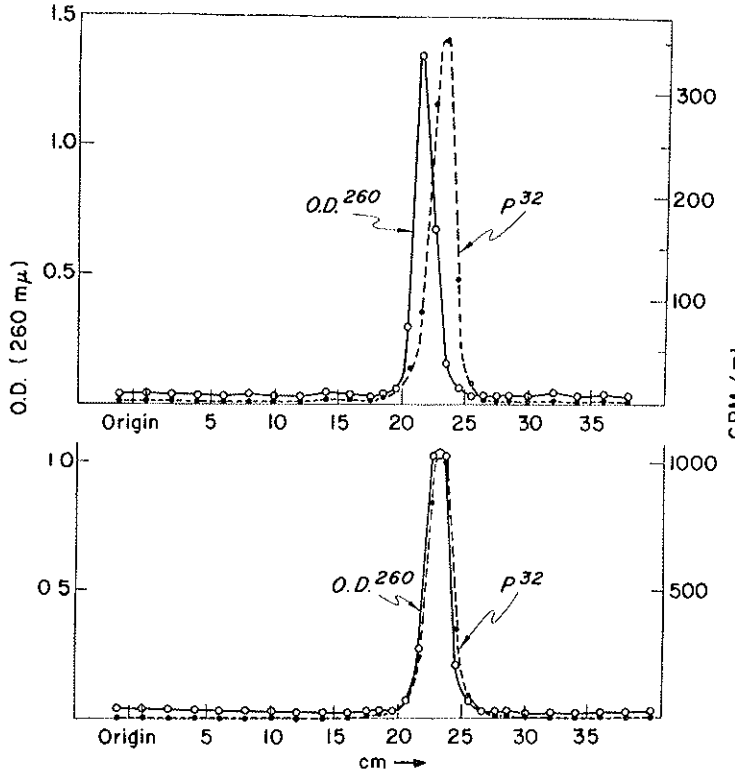


FIG. 3 — Starch column electrophoresis of P^{32} ribosome RNA. *a*) RNA from T2-infected cells, given P^{32} between 5 and 7 min. after infection. *b*) RNA from control cells given a 2 min. P^{32} pulse. In each case, electrophoresis was carried out for 10 hr. at 8.0 v/cm, 6 ma in 0.03 M-tris buffer, pH 7.8.

rose gradients. Fig. 4 compares the profiles obtained in the two situations. The first peak in the optical density profile corresponds to the 23S component and the second to the 16S. The striking difference between the infected and control pulses is that little, if any, of the RNA synthesized in the T2-infected cells corresponds to the two main ribosomal components. There is very little concordance between the O.D. and radioactive

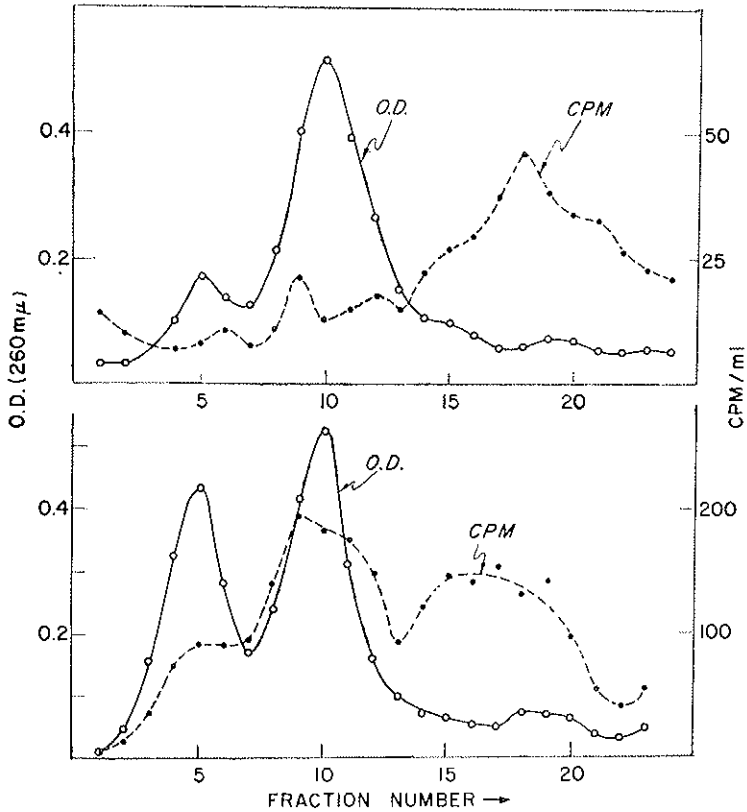


FIG. 4 — Sedimentation of P^{32} particle RNA. *a*) RNA from T2-infected cells, exposed to P^{32} between 5 and 7 min. after infection. *b*) RNA from the control cells given a 2-minute P^{32} pulse. 0.25 mg RNA in 0.03 M-tris (pH 7.8) were layered on 4.4 ml. of sucrose solution, with a concentration gradient from 3 to 20%. This was centrifuged for 8 hr. at 37,000 rev/min. Fractions were collected by piercing the lower end of the tube and collecting 10-drop fractions. To each fraction 1 ml. 0.03 M-tris buffer was added; optical density at 260 mμ and P^{32} content were then measured.

profiles in the 23S and 16S regions. Most of the RNA formed subsequent to T2-infection appears to possess, by these methods of isolation, a sedimentation lying between 8 and 12S. In the non-infected control we note a considerable formation of RNA characteristic of the 23S and 16S varieties. It must, however, again be emphasized what is clearly evident from the two sets of curves of Fig. 4. The synthesis of the smaller size of RNA is *not unique* to the virus infected cell. The short P^{32}-pulse (Fig. 4-*b*) readily reveals its existence in the uninfected controls. These experiments reveal that the essential effect of infection with virus is a preferential suppression of 23S and 16S RNA synthesis.

In summary then, the experiments just described establish the existence of a T2-specific RNA as a physically separable entity. The features which serve to distinguish it from normal ribosomal RNA may be listed as follows:

a) a base ratio homologous to the viral DNA,

b) metabolic instability,

c) a higher electrophoretic mobility,

d) a greater heterogeneity in size with an average sedimentation constant lower than 16S.

III. Sequence Complementarity of T2-DNA and T2-Specificity RNA

The procedures employed (zone electrophoresis and sedimentation) in the investigations just summarized led to the selective separation of T2-specific RNA. They, therefore, opened up the possibilities for further experiments relevant to an understanding of its nature. The fact that T2-RNA possesses a base ratio analogous to that of T2-DNA is of interest principally because it suggests that the similarity may go further and extend to a detailed correspondence of base sequence.

The central issue of the significance and meaning of T2-RNA is whether or not this is, in fact, the case.

A direct attack on this question by complete sequence determinations was and is technically not feasible. However, the findings of MARMUR (1960) and DOTY et al., (1960), suggested the possibility for an illuminating experiment. These authors demonstrated the specific reformation of double stranded DNA when heat denatured DNA is subjected to a slow cooling process. Such reconstitution of double stranded structures occurs only between DNA strands which originate from the same or closely related organisms. Presumably, the specificity requirement for a successful union of two strands reflects the need for a perfect or near perfect complementarity of their nucleotide sequences. We have here then a method for detecting the complementarity of nucleotide sequences in two strands of polynucleotide. The formation of a double stranded hybrid structure during the slow cooling of a mixture of two types of polynucleotide strands can be accepted as provisional evidence for complementarity of the bases of the input strands.

HALL and SPIEGELMAN (1961) undertook to use this procedure to examine for the complementarity of sequences between the T2-RNA and the T2-DNA. Purified T2-RNA was used in order to provide an optimal opportunity for the T2-RNA to combine with its DNA complement unhindered by non-specific interactions involving irrelevant RNA. Since the hybrid would have a lower density than uncombined RNA, a separation of the two should be attainable by equilibrium centrifugation in cesium chloride gradients (MESELSON, STAHL and VINOGRAD, 1957). To insure a sensitive and unambiguous detection of the hybrid, should it occur, double labeling was employed. The T2-RNA was marked with P^{32} and the T2-DNA with tritium. Two isotopes emitting beta particles differing in their energies are conveniently assayed in each other's presence in a scintillation spectrometer. This device, coupled with the use of the swinging bucket rotor for equilibrium centrifugation

permits the actual isolation of the pertinent fractions along with a ready and certain identification of any hybrids formed.

Mixtures of P^{32}-labeled to T2-RNA and H^3-labeled single stranded T2-DNA were subjected to slow cooling. They were then put in CsCl and centrifuged to density equilibration. Fig. 5 shows the optical density profiles and the distribution of tritium and P^{32} obtained from three preparations slow cooled from different starting temperatures. Comparison of the profiles of tritium and P^{32} show that in all three cases slow cooling of the DNA and RNA produced a new peak of P^{32} approximately centered on the band of tritium (denatured DNA). This new P^{32} containing band must contain a DNA-RNA hybrid having approximately the same density as denatured T2-DNA. The amount of complex formed on cooling from the three temperatures was the same within experimental error. The three differ, however, slightly in the density of the complex relative to DNA.

Exposure to the slow cooling process is necessary since an uncooled mixture of T2-RNA and single stranded T2-DNA exhibited no P^{32} peak in the DNA region when subjected to a centrifugal analysis. Furthermore, the presence of single stranded DNA is necessary during the cooling process. Double stranded DNA is unable to form hybrid with RNA under these conditions. Finally, the ability of T2-RNA to hybridize was found to be specific for T2-DNA. Thus, when the cooling process was carried with mixtures of P^{32}-labeled RNA of T2 and denatured DNA from heterologous sources (*Pseudomonas aeruginosa, E. coli* and bacteriophage T5) no evidence of significant hybrid formation was observed. It is of interest to note that although T5 has the same over-all base ratio as T2 no evidence of interaction of T2-RNA and T5-DNA was detected.

The data obtained in the course of this investigation showed that RNA molecules synthesized in bacteriophage infected cells have the ability to form a well defined complex with denatured DNA of the virus. Further, this interaction is unique to the

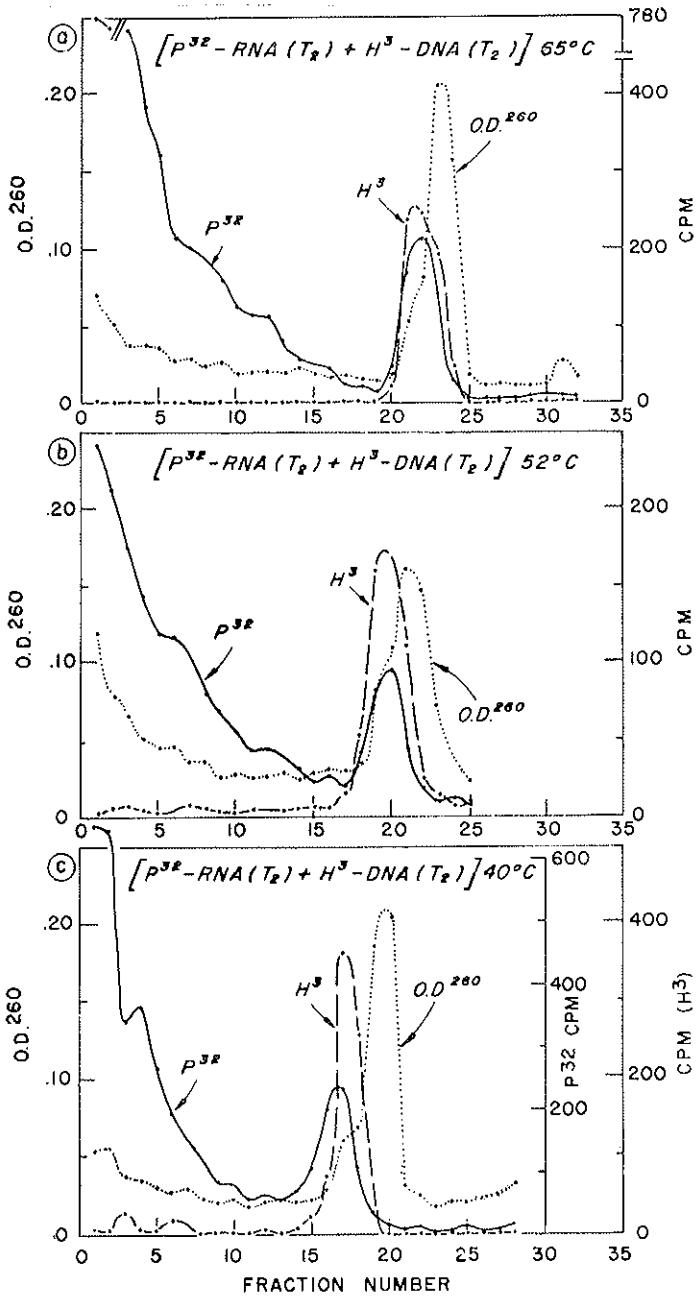


FIG 5 — Formation of DNA-RNA hybrid at various temperatures. CsCl-gradient centrifugation analysis. P^{32}-RNA(T_2) (14 μg) and H^3-DNA(T_2) (6.5 μg) were mixed in 0.6 ml 0.3 M NaCl and 0.03 M Na citrate (pH 7.8); then the solution was immediately placed in the slow-cooling bath. Three identical solutions were made; *a*) was placed in the bath at 65°, *b*) at 52°, and *c*) at 40°C. When the bath temperature reached 26°, CsCl and 25 μg T_2-DNA were added to each solution; then they were centrifuged for five days at 33,000 rpm.

homologous pair as shown by the virtual absence of such complexes when T2-specific RNA is slowly cooled with heterologous DNA. The fact that T2-RNA and DNA do satisfy the specificity requirement must reflect a correspondence of structure between the two. Structural specificity of this order in single polynucleotide strands can only reside in specified sequences of nucleotides. It was concluded, therefore, that the most likely interrelation of the nucleotide sequences of T2-DNA and RNA is one which is complementary in terms of the scheme of hydrogen bonding proposed by WATSON and CRICK (1953).

IV. THE OCCURRENCE OF NATURAL DNA-RNA COMPLEXES IN THE *E. coli* T2 COMPLEX

The fact that evidence could be provided for sequence complementarity between a specific RNA and its homologous DNA lends obvious support for the supposition that the normal process of transferring information from DNA to the protein synthesizing machine involves a mechanism whereby single stranded DNA serves as a template for the polymerization of a complementary ribopolynucleotide. If continued formation of complementary RNA is a necessary concomitant, it should be possible to find RNA-DNA hybrids in any cell actively engaged in protein synthesis.

SPIEGELMAN, HALL and STORCK (1961) undertook to find such complexes and the T2 *E. coli* system was selected as the most suitable in the initial search for native hybrid. The experimental devices employed were in essence similar to those nsed in the previous investigation on artificially formed complexes. These involved the use of double labeling and equilibrium centrifugation in cesium chloride gradients employing swinging bucket rotors. The T2-DNA was labeled with P^{32} by growth of the virus in a medium containing this isotope. T2-specific RNA was marked with tritium by introducing tri-

tiated uridine. It is known that some uridine ends up in the DNA once its formation begins in the T2-*E. coli* complex. To avoid the complications this would introduce in identifying radioactive peaks, the initial search for hybrids was confined to the period (2-5 minutes) when no DNA synthesis can be detected. In a double label experiment the presence of DNA-RNA hybrids would be signalled by the appearance of coincident peaks of the two isotopes. Further, these peaks should occupy a position in the cesium chloride gradient differing from those which characterize the densities of RNA and double stranded DNA.

A variety of procedures were surveyed for obtaining material suitable for the reliable detection of DNA-RNA hybrids by cesium chloride density equilibrium centrifugations. It was empirically established that removal of most of the protein was a necessary step prior to introducing the material into the cesium chloride. In the absence of this preliminary purification much of the nucleic acid, including hybridized material, was trapped in the protein layer found floating at the top of the gradient. The procedure ultimately adopted consisted essentially of the first few deproteinization stages normally (MARMUR, 1961) employed for preparing DNA from bacterial cells. Two modifications were made. None of the steps designed to remove RNA are included. Further, at the alcohol precipitation stage centrifugation, rather than winding around a glass rod, is used to collect the nucleic acid. Such DNA preparations are, of course, heavily contaminated with RNA but serve well the intended purpose.

In experiments which employ P^{32}-labeled T2-DNA and tritiated uridine as the RNA marker, existence of hybrid would be detected as a tritium peak in the DNA region somewhat heavier than the P^{32} peak corresponding to the T2-DNA. The P^{32} peak should exhibit signs of bi-modality as evidence of distribution of the T2-DNA between hybrid and unhybridized T2-DNA. Fig. 6 describes the results of a typical profile obtain-

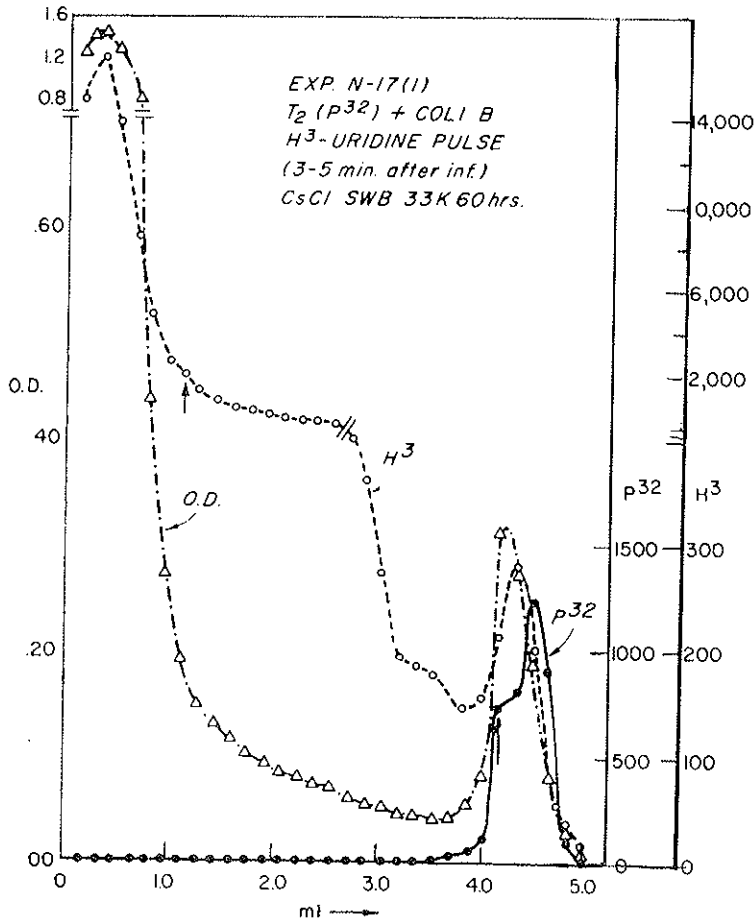


FIG. 6 — Equilibrium Density Centrifugation in CsCl of Nucleic Acid from T2-Infected Cells. Absorption of phage was carried out at a multiplicity of 15 of undiluted P^{32}-labeled virus. Pulsing was done between 3-5 minutes after zero time with H^3-uridine having a specific activity of 530 μc/μM at a level of 1 μg/ml. Note that the radioactivity scale is expanded in the hybrid region.

ed by subjecting a nucleic acid preparation from a pulse labeled T2-infected cell to equilibrium gradient centrifugation. Three peaks are readily discernable. The lightest one corresponds to the input P^{32}-labeled DNA of T2. The optical density peak identifies the position of double stranded *E. coli* DNA put in as a marker. The tritium peak locates the hybrid. Furthermore, a shoulder makes its appearance on the heavy side of the P^{32} profile corresponding in position to the tritium peak region. The relative positions of the peaks and the bi-modality of the P^{32} distribution are all consistent with what would be predicted from the existence of a DNA-RNA hybrid.

If it be assumed that a small amount of T2-DNA can be synthesized even in this early period of infection, another interpretation of the optical density and radioactive profiles can be entertained. This would presume the appearance of new T2-DNA, all in the form of single stranded material, along with the conversion of a portion of the input P^{32}-labeled DNA to single strands. To test for this possibility, the alkali lability of the tritium and P^{32} in relevant fractions of a number of experiments were examined. The same results were obtained in all cases and are exemplified by the data summarized in Table I. Here, the two fractions indicated by arrows in Fig. 6 were subjected to alkali digestion and the resultant effect on the acid precipitability of tritium and P^{32} were examined. Fraction 7 was included as a free RNA control. It is evident from the data in Table I that all of the tritium and none of the P^{32} counts are alkali labile. These data appear to eliminate the possibility that the displaced tritium peak can be ascribed to newly synthesized single stranded DNA. It is also evident that the displaced P^{32} cannot be ascribed to a conversion of the input viral DNA components to an RNA polynucleotide.

We may, therefore, conclude that a hybrid exists containing newly synthesized T2-specific RNA complexed with some of the input P^{32} labeled DNA. It should be noted that hybrids involving the input DNA were also observed in pulses covering

TABLE I — *Effect of Alkali Digestion on Acid Precipitability*

0.4 ml aliquots of Fractions 7 and 26, indicated by arrows in Fig. 6 were made 0.3 N with respect to NaOH and incubated for 24 hours at 30°C. Equivalent aliquots were held as controls under the same conditions. Following the incubation, the alkali was neutralized, carrier herring sperm DNA added and the contents precipitated with TCA. The precipitates were then washed and counted.

Fraction No.	Isotope	Acid Precipitable Counts/ml. Control	After Alkali Treatment
26	H^3	230	5
	P^{32}	735	720
7	H^3	1950	10
	P^{32}	—	—

in addition to the 2 to 5 minutes, 9 to 12 and 19 to 22 minutes after infection.

The existence of the natural hybrids constitutes another link in the chain of evidence supporting the simple complementary transcription mechanism mentioned earlier.

V. SELECTIVE SYNTHESIS OF INFORMATIONAL RNA IN NON-INFECTED CELLS

We now turn to the last question posed in the introductory paragraphs, namely the existence of complementary RNA in normal cells. The detection and study of the properties of complementary RNA formed in T2-infected cells was greatly facilitated by the fact that the larger ribosomal components are not synthesized. This advantage is not present in normal

uninfected cells which consequently complicates the search for normal informational RNA. That it is, nevertheless, feasible is suggested by the experiments of YČAS and VINCENT (1960) with yeast. These authors used P^{32} in a manner comparable to the procedures of VOLKIN and ASTRACHAN and, despite surprisingly long pulses, were able to detect formation of a fraction with a high metabolic turnover and possessing a base composition analogous to yeast DNA. In a preliminary report, ASTRACHAN and FISCHER (1961) suggest that very short P^{32} pulses leads to a distribution of labile in the RNA synthesized which indicates the synthesis of RNA mimicking the base composition of DNA.

It would clearly be of great advantage if a situation could be found or devised in normal cells which would be analogous to that which occurs on infection with T2. Essentially what we are demanding is a condition which suppresses ribosomal RNA synthesis and permits the formation of the informational variety. The possibility that a situation of this sort might, in fact, be realizable was suggested by studies on RNA and protein synthesis during passage from fast to slow growth.

Several features emerged which encouraged us to look more carefully into such transitions. It has been known for some time that the RNA content per cell is positively correlated with its growth rate. Since the bulk of the RNA is ribosomal, it would mean that cells with higher growth rates possess more ribosomes. Consider then the situation when one subjects a culture to a « step-down » transition by transferring cells from a rich (e.g. PENNASSAY) to a synthetic medium. The growth rate is decreased by a factor of two. But more important, the cells have more ribosomes than they can use. From the point of view of selective advantages it is not surprising, therefore, to find that such stepdown transitions cause a rather dramatic cessation of net RNA synthesis. Nevertheless, protein synthesis proceeds for a while at near normal rates. From the viewpoint of relative rates of net protein and RNA synthesis, such cul-

tures are analogous to T2-infected cells. It seemed not too unlikely that whatever RNA synthesis was occurring was restricted to the variety immediately necessary for the fabrication of new protein molecules. It was possible that this variety was the normal informational RNA for which we were searching.

HAYASHI and SPIEGELMAN (1961) undertook to see whether the above expectations were realizable. Base compositions of RNA synthesized in « step-down » cultures were examined. In all cases, such cultures were attained by transfer from a complete (PENASSAY) to a synthetic (Med. C. of ROBERTS *et al.*, 1957) medium. Base ratios were determined by procedures similar to those used in the study of the T2-*coli* system. P^{32} pulses were made and the ribonucleic acid isolated and purified, subjected to alkaline digestion, and the distribution of counts in the resultant 2′-3′ nucleotides determined. Table II summarizes experiments with three different organisms in which the P^{32} pulse was carried out at various periods subsequent to the transitional transfer. It will be noted that in each case the RNA synthesized during the transition period mimics the homologous DNA in its per cent GC and purine to pyrimidine ratio. What is even more remarkable is the length of time during which this selective synthesis of complementary RNA continues. Thus, in the case of *Ps. aeruginosa* even 60 minutes after the transfer a major fraction of the RNA formed is homologous to its DNA.

VI. THE SIZE DISTRIBUTION OF INFORMATIONAL RNA FORMED IN NON-INFECTED CELLS

The data summarized in Table II appeared to confirm our expectation that informational RNA is preferentially synthesized in step-down cultures. It was of obvious interest to continue this investigation and see whether the other properties

TABLE II — *Base Ratios of RNA in « Step-Down » Cultures*

All cultures were transferred in log phase from complete to synthetic medium at 30°C. At times indicated, they were subjected to a 3 min. pulse with P^{32}. The RNA was removed, purified and hydrolyzed with alkali in the presence of added carrier RNA. The nucleotides in the resulting hydrolysate were separated on Dowex columns and counted. The numbers given are derived from the distribution of the counts and isotope dilution. For purposes of comparison the total RNA base composition determined from UV absorption data are included for each organism along with the homologous DNA base composition.

Organism	Minutes after Transfer	*Moles Percent* C	A	U(T)	G	%GC	Pu/Pyr
E. coli	5	24.7	24.1	23.5	27.7	52.4	1.07
	60	25.2	24.1	22.1	28.6	53.8	1.12
	DNA	26	24	24	26	52.0	1.00
	Bulk-RNA	24.3	25.0	19.7	31.0	54.3	1.27
Ps. aeruginosa	5	29.0	21.3	20.2	29.5	58.5	1.03
	60	27.1	21.8	21.2	29.9	57.0	1.07
	DNA	32	18	18	32	64	1.00
	Bulk-RNA	22.3	23.1	23.6	31.0	53.3	1.21
B. megaterium	5	19.7	27.9	29.0	23.4	43.4	1.05
	DNA	19	31	31	19	38	1.00
	Bulk-RNA	21.9	22.4	23.6	32.0	53.9	1.19

of informational RNA revealed by the study of the T2-*E. coli* complex obtained here as well.

Size distributions of the RNA synthesized were examined by the usual procedure of labeling and subjecting the purified RNA to a swinging bucket analysis on a sucrose gradient. Fig. 7 gives the results obtained with *E. coli* exposed to a 30 min. « pulse » of H^3-uridine during a step-down transition.

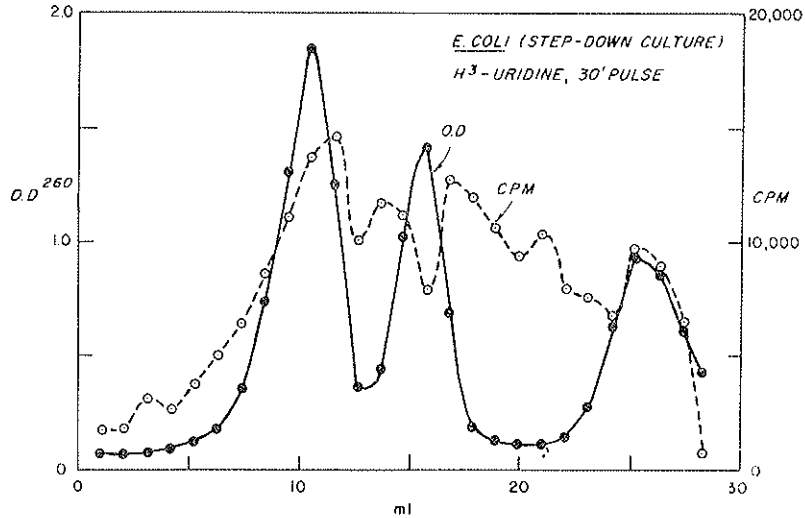


FIG. 7 — Swinging Bucket Analysis in 3 - 20% Sucrose Gradient of Phenol Purified RNA. Cells were exposed to H^3-uridine for 30 minutes during transition from complete to synthetic medium. Closed circles identify pre-existent and open circles newly synthesized RNA.

The optical density profile identifies readily the 23, 16 and 4S RNA components pre-existing in the cell. A similar period of labeling in a normal culture transferred from synthetic to synthetic medium would have resulted in virtually complete coincidence of the radioactivity and optical density profiles. Here we see little, if any, agreement between the two.

It was clearly desirable to extend our information on the RNA being synthesized, with particular reference to the base compositions of the various size ranges of RNA observed. For obvious numerical reasons and other technical considerations, *Ps. aeruginosa* was chosen for this more detailed analysis. A step-down culture was subjected to a P^{32} pulse and the ribonucleic acid examined centrifugally in a sucrose gradient. The

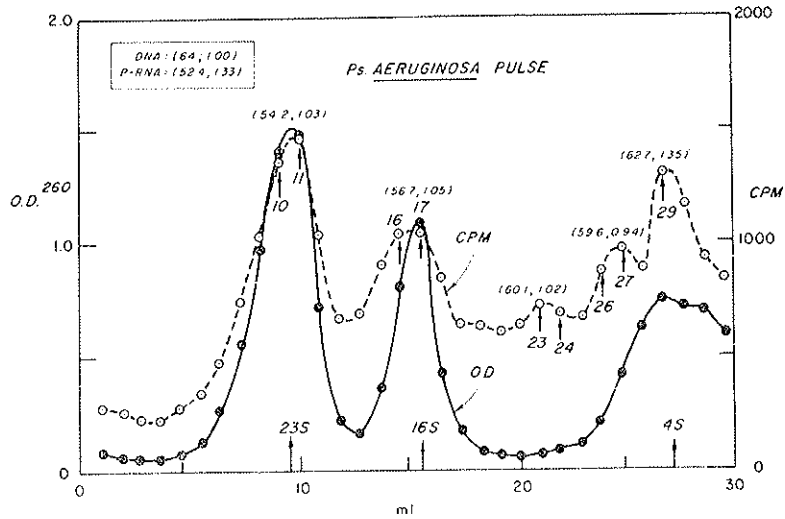


FIG. 8 — Swinging Bucket Analysis in 3 - 20% Sucrose Gradients of Phenol Purified RNA. Cells were exposed to a 3 minute P^{32}-pulse 5 minutes after they were transferred from complete to synthetic medium. Closed circles identify pre-existent and open circles newly synthesized RNA. The first number in parenthesis represents per cent GC and the second the ratio of purines to pyrimidines. Arrows indicate the fractions taken for base composition analysis.

optical density and radioactivity profiles are given in Fig. 8. The fractions indicated by arrows were analyzed for base compositions. For purposes of ready comparison, numbers corresponding to the per cent GC and purine to pyrimidine ratio which characterize each region are recorded in parentheses. Further details on the base ratio analyses are given in Table III.

TABLE III — *Ps. aeruginosa*: Base Composition of Various Sizes of RNA 3 min. P^{32}-Pulse of « Step-Down » Culture

Conditions of experiment and analyses similar to those described in Table II. The fractions taken are those indicated by arrows in Fig. 8. P-RNA means purified ribosomal RNA and the base composition was obtained from UV absorption data of the nucleotides.

		Moles Percent					
Fraction No. (Fig. 8)	Region	C	A	U(T)	G	%GC	Pu/Pyr
10,11	23S	25.9	22.4	23.4	28.3	54.2	1.03
16,17	16S	27.0	21.5	21.8	29.7	56.7	1.05
23,24	10-12S	30.3	20.9	19.0	29.8	60.1	1.02
26,27	6-8S	31.2	19.8	20.6	28.4	59.6	0.94
DNA		32	18	18	32	64	1.00
P-RNA		22.4	26.8	20.7	30.1	52.5	1.30

Comparison of the parameters reveals that DNA, like RNA of all size classes have been synthesized, confirming the findings with *E. coli* (Fig. 7). As one proceeds to the smaller size ranges (16S-6S) the homology between the RNA and DNA becomes excellent.

It was of obvious interest to examine the metabolic stability of the complementary RNA synthesized in the transition period. To accomplish this an aliquot of the culture used in the experiment of Fig. 8 was removed after the pulse, washed, reintroduced into the same medium containing P^{31} and allowed to chase for 0.7 generations. Fig. 12 describes the optical density and radioactivity profiles obtained when purified RNA from this preparation was centrifuged in the usual way. Here again individual fractions were taken for base ratio determinations. The numbers in parentheses give the results in terms of per cent GC and purine to pyrimidine ratios. Table IV provides further details on the base ratios determined by both the ultraviolet absorption derived from the added carrier and the distribution of radioactive counts amongst the 2′-3′ nucleotides as eluted from a DOWEX column. Comparison of Figs. 8 and 9 provides clear evidence of the metabolic instability of the heterogeneous RNA synthesized during the transition period. The interval of chasing eliminated almost completely the discordancies between the optical density and radioactivity profiles seen in the initial pulse (Fig. 8). Furthermore, as is evident from Table IV, the base ratios of the labeled RNA in the 23S and 16S region are now typically ribosomal. There is excellent agreement between the base ratios determined by the distribution of radioactive counts and ultraviolet absorption.

It is of interest to note that despite the fact that the chase extended for a period of 0.7 of a generation, there is still some discrepancy in both the profiles of Fig. 5 and the base compositions in the 4S regions (Table IV). This may be a reflection of the difficulty of completely removing informational RNA. It would be consistent with a mechanism which involves a com-

TABLE IV — Base Composition of RNA of Different Sizes Subsequent to Chase of the Culture of Table III

An aliquot of the culture used in the experiment of Table III and Fig. 8 was taken after the 3 min. P^{32}-pulse, washed and allowed to grow for 0.7 generations in an unlabeled medium. The RNA was prepared and analyzed in the usual way. The data from the UV absorption are included to permit a comparison of the degree of correspondence between the radioactive and UV calculation on the same samples. The fractions taken are indicated by the arrows in Fig. 9. The data obtained on P-RNA and S-RNA were from UV absorption data on separately purified material. P-RNA has the same meaning as in Table III. S-RNA is the RNA remaining in the supernatant after removal of ribosomes by means of a 38K spin for 6 hours.

Ps. aeruginosa

Moles Percent

		C		A		U		G		%GC		Pu/Pyr.	
Fr No.	Region	CPM	UV	CPM	UV	CPM	UV	CPM	UV	CPM	UV	CPM	UV
12	23S	22.2	22.1	26.0	26.9	22.1	20.8	29.7	30.2	51.9	52.3	1.26	1.33
19	16S	22.2	22.8	25.5	26.7	21.7	20.5	30.6	30.0	52.8	52.8	1.26	1.31
32,33	4S	31.9	29.2	19.4	24.4	18.8	20.5	29.9	25.9	61.8	56.1	0.97	1.03
DNA			32		18		18		32		64		1.00
P-RNA			22.4		26.8		20.7		30.1		52.5		1.30
S-RNA			29.2		24.4		20.5		25.9		56.1		0.97

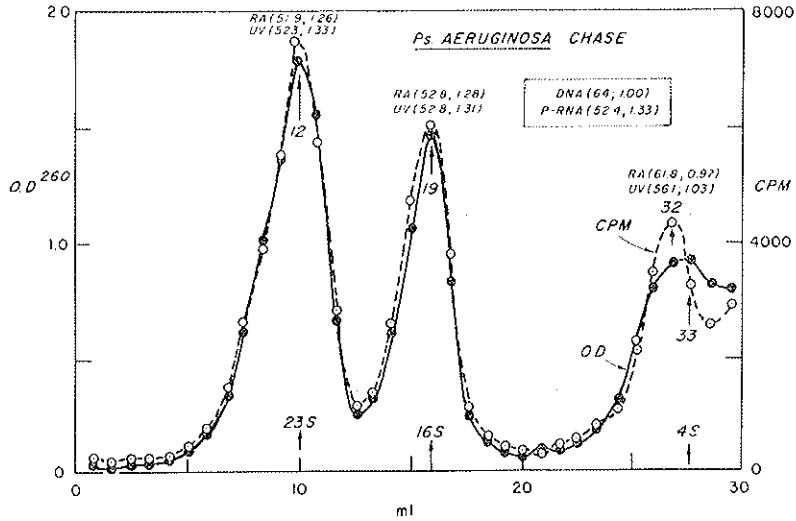


FIG. 9 — Swinging Bucket Analysis in a 3 - 20% Sucrose Gradient of Phenol Purified RNA. Cells represent an aliquot taken from the experiment of Fig. 8 and exposed to a « chase » for 0.7 generations in non-radioactive synthetic medium. The first number in parentheses represents per cent GC and the second, ratio of purines to pyrimidines. The values found by both UV absorption and radioactivity are given. Arrows indicate fractions subjected to base composition analysis.

paratively rapid breakdown of the larger informational RNA pieces to 4S size and a slower conversion of these to the level of nucleotide derivatives.

VII. Hybridizability of Informational RNA from Normal Cells with Homologous DNA

The experiments described thus far with non-infected cells established that « step-down » cultures preferentially synthesized a type of RNA which was heterogeneous in size, metabolically unstable, and possessed an over-all base ratio which was analogous to its homologous DNA. These are features

expected of informational RNA. To complete the identification, it was necessary to test for sequence complementarity by the hybridization procedure. An examination of the hybridizability of complementary RNA to homologous and heterologous DNA was undertaken by SPIEGELMAN, DOI and YANKOFSKY (1961). We cite here only a few representative experiments illustrating the principal features and findings. The general procedures employed may be outlined as follows:

1) Step-down cultures were pulsed with H^3-uridine to label the RNA synthesized during transition.

2) The RNA was isolated and purified by the phenol method.

3) The purified RNA was separated according to size on sucrose gradients.

4) Different regions of the radioactive profile were collected and concentrated.

5) Hybridizing tests were carried out by exposing mixtures of the labeled RNA and single stranded DNA to a slow cool from 55°C.

6) The resulting mixtures were then subjected to equilibrium centrifugation in CsCl gradients according to the methods described by HALL and SPIEGELMAN (1960).

Fig. 10 shows the outcome of a hybridization carried out between single stranded *E. coli* DNA and 8 - 12S H^3-RNA labeled during a step-down transition. It will be noted that excellent hybridization occurs. The shoulder in the optical density profile on the light side corresponds to marker double stranded *coli* DNA. That the interaction is specific is shown in Fig. 11 in which a similar hybridizing attempt was made between the same RNA fraction and single stranded DNA derived from *Pseudomonas aeruginosa*. There is no suggestion of any detectable mating.

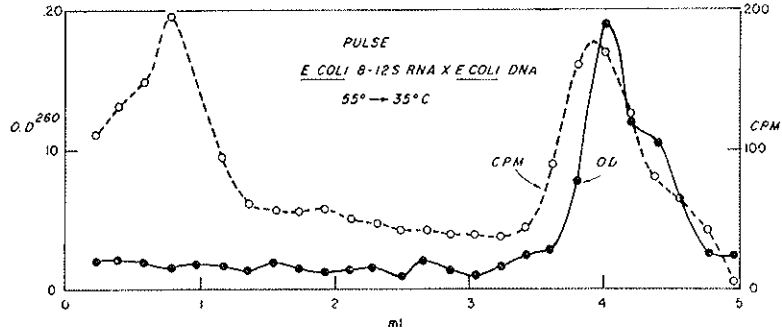


Fig. 10 — Equilibrium Density Centrifugation (33 K 60 hours) in CsCl. A mixture of H^3-RNA (8 - 12S) from an *E. coli* « step-down culture » slow cooled with single stranded *E. coli* DNA. Double stranded *E. coli* DNA was added as a marker and is represented by the shoulder on the light (right) side of the main O.D. peak.

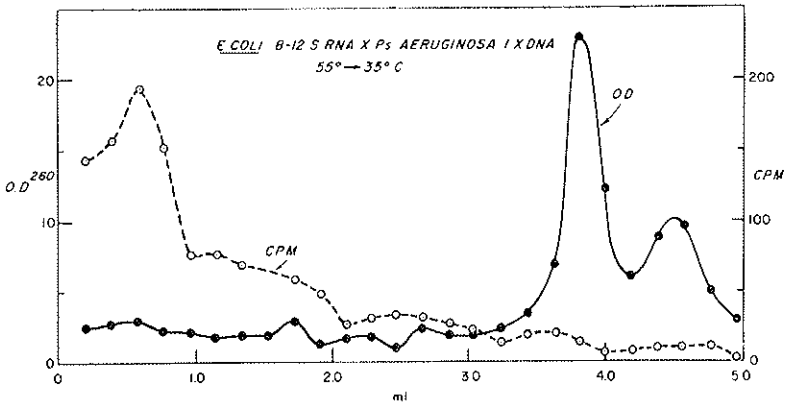


FIG. 11 — Equilibrium Density Centrifugation (33K 60 hours) in CsCl. A mixture of H^3-RNA (8 - 12S) from an *E. coli* « step-down » culture slow cooled with single stranded DNA from *Ps. aeruginosa*. Double stranded *E. coli* DNA was added as a marker and is represented by the second peak to the right.

Similar experiments were carried out with *Ps. aeruginosa*. The RNA was labeled with H^3-uridine during a step-down transition. Fig. 12 shows the outcome of a hybridization carried out with homologous single stranded DNA and H^3-RNA removed from the 16S region. Here again we note excellent hybrid formation as demonstrated by the peak of tritium in the DNA region. This same figure illustrates a feature which is extremely useful in attempts at detecting hybrid and distinguishing it from non-specific aggregation. Aliquots from each of the tubes were taken and treated with 10 gamma of RNAase for 15 mins. at room temperature and then carrier DNA was added and the material reprecipitated, washed and recounted. It will be noted that most of the counts corresponding to free RNA are almost

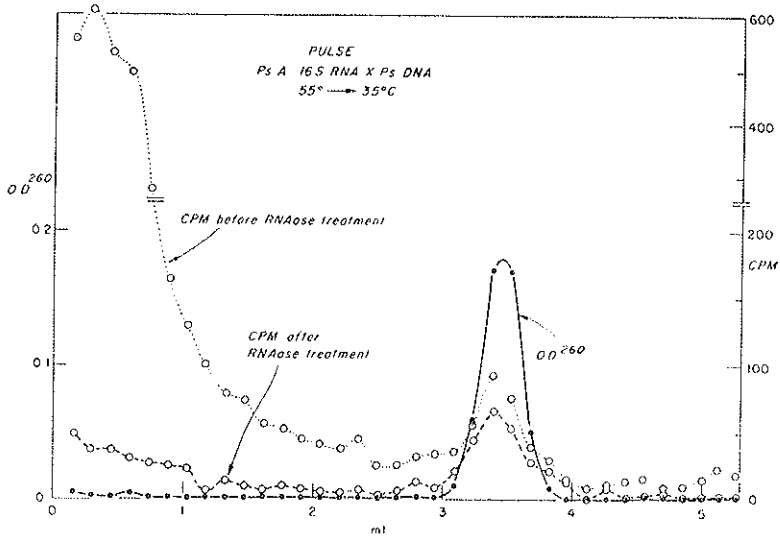


FIG. 12 — Equilibrium Density Centrifugation (33K 60 hours) in CsCl. A mixture of H^3-RNA (16S) from a *Ps. aeruginosa* « step-down » culture slow cooled with single stranded DNA from *Ps. aeruginosa*. No marker was added. Open circles - dashed line gives the effects on the cpm of treatment of the indicated fractions with RNAase prior to precipitation and counting.

completely removed by the RNAase treatment. However, the counts in the region of the hybrid are obviously much more resistant to RNAase.

Specificity tests with informational RNA from *Ps. aeruginosa* yielded results similar to those described for *E. coli*. No interaction with heterologous single stranded DNA was observed.

The experiments summarized in the present section establish that RNA molecules preferentially synthesized in « step-down » cultures possess base sequences complementary to their homologous DNA.

VIII. Discussion

We may perhaps begin with a few words on terminology. The terms « complementary » and « informational » have been used to describe the RNA molecules with which we are concerned. It is evident from the experiments discussed that these terms have well defined operational definitions. A given RNA molecule is defined as falling within the informational class if its base ratio is homologous and its sequence is complementary to a specific DNA molecule. At present the most sensitive test for complementarity of sequence is the hybridization experiment of Hall and Spiegelman (1960). Every « complementary » RNA is « informational » in at least one sense. Even if it is a complementary copy of a nonsense DNA sequence, it still contains the information necessary to specify the order of the bases.

It is important to emphasize that the word « informational » is not proposed as a substitute for the term « messenger » introduced in the elegant experimentations and theorizations of Jacob and Monod (1961). It seems likely that both terms will be useful. Thus, a given messenger RNA is presumed to constitute the structural program for the synthesis of a par-

ticular protein. It obviously must, therefore, be informational. However, not all informational RNA need serve a messenger function. It is conceivable, as is indeed implicit in the operon theory of JACOB and MONOD (1961), that informational RNA molecules will be found which serve regulatory rather than programming functions.

Before concluding it may be useful to discuss some problems and implications which arose during the investigations described and which have not received explicit mention.

A. *Size and Stability of Informational RNA.*

As our experience with informational RNA accumulated, it became more and more evident that this type of RNA is uniquely susceptible to degradation in extracts. Even highly purified preparations, obtained from either T2-infected cells or « stepped-down » cultures, were much more fragile than ribosomal RNA carried through the same procedures. In a sense, the behavior of informational RNA reminds one of the apparent inherent instability of the synthetic polypyrimidines (see discussion of RICH, 1958). The underlying reason may well be the same and there is a pressing need for an understanding of its chemical basis.

In any event, we became increasingly careful in the methods used to prepare this type of RNA. By the time the experiments involving informational RNA from normal cells (section VI) were performed, procedures had been devised which minimized contact of the newly synthesized RNA with enzymatically active extract. It will be noted from the results described in this section that there is clear evidence in these cases of informational RNA, 16S and larger, possessing hybridizing ability.

These observations led us to consider the possibility that the smaller, and more homogeneous size, deduced from the earlier experiments with the T2-infected cells (section II), might have been a consequence of the fact that we had not at that

time understood fully the instability problem. Experiments were, therefore, undertaken by SAGIK, GREEN and SPIEGELMAN (1961) to re-examine the size distribution of T2-specific RNA using the newer procedures. It was indeed found that T2-complementary RNA showed a much more heterogeneous distribution of sedimentation constants than that which characterized the earlier experiments. RNA hybridizable with T2-DNA was found in sizes ranging from 23S - 8S.

Our experience to date leads to the conclusion that informational RNA is extremely heterogeneous in size. Left unresolved is the question of whether the small size ranges observed are real or artifacts of unavoidable breakdown.

These considerations raise another issue of practical significance. It is evident that « soluble » RNA is very likely to contain small informational components as contaminants of the presumed transfer RNA molecules. In this same connection it should be noted that the chasing experiment of Fig. 9 did not suffice to completely remove informational RNA from the 4S region even after 0.7 generations. Stimulatory effects on amino acid incorporation with so-called « soluble » RNA are obviously open to more than one interpretation with respect to the responsible agent.

B. *Genesis of the Major RNA Components.*

The question naturally arises of the nature of the major RNA constituents, the base ratios of which bear no obvious relation to homologous DNA's. Two hypotheses can be entertained as to their origin. One is they are complementary in the sense of being synthesized by a DNA mediated mechanism. The second assumes that DNA is not involved. In this case, the mechanism of their formation might involve an RNA template as a guide or no template, the polymer being put together much as polysaccharides are formed.

It might perhaps seem a simple matter to settle this by examining for hybridizability of the relevant RNA to DNA. In the course of the present studies, many experiments have been performed in an attempt to detect hybrid formation with 23 and 16S ribosomal RNA. None has been found. On the basis of the data available we can state that on a per μg basis, the hybridizing ability of ribosomal RNA is 100 - 1,000 times poorer than informational RNA obtained by the methods described. However, this does not serve to eliminate the possibility of hybrid formation with ribosomal RNA. The numerology of the situation makes it technically difficult to obtain a definitively negative answer. Consider the 23S molecule, which is the most favorable. Its molecular weight of 1.6×10^6 represents approximately 0.1% of the bacterial genome. If there is only one section of the DNA concerned with its synthesis, then this is the only portion which will hybridize. To settle questions of this nature one will have to push the sensitivity of hybrid detection two orders of magnitude beyond that used to date. This is technically difficult but by no means impossible.

C. *Enzymological Implications.*

The transcription mechanism supported by the data described in the preceding sections would require a DNA dependent enzymatic mechanism for polyribonucleotide synthesis. Early evidence suggesting the existence of such a pathway emerged from a study (SPIEGELMAN, 1958) of a cell free system derived from *E. coli*. These preparations possessed considerable capacity to synthesize polyribonucleotide. It was routinely possible to obtain between 10 and 20-fold increases of polyribonucleotide. The observed synthetic activity exhibited a requirement for riboside-triphosphates, and Mn and was severely inhibited by treatment with DNAase. The last year has witnessed a notable advance in our understanding of the details of this

reaction. Recently a number of laboratories (WEISS, 1960; HURWITZ, BRESLER and DINGER, 1960; STEVENS, 1960; and OCHOA, BURMA, KROGER and WEILL, 1961) have independently achieved a considerable purification of an enzyme which synthesizes polyribonucleotide and requires riboside-triphosphates and DNA. Furthermore, the base ratio of the polyribonucleotide synthesized bears a striking homology to that of the DNA used as a primer (FURTH, HURWITZ, GOLDMANN, 1961; WEISS and NAKAMOTO, 1961).

IV. SUMMARY AND CONCLUSION

The present paper records our attempts over the past three years to gain an understanding of the nature of the agent which allows the genome to exert its control over macromolecular synthesis. The direction taken by the series of investigations described was greatly influenced by two factors. One was the assumption that the intermediary between the DNA and the protein synthesizing mechanism must be a polyribonucleotide. The second was the observation of VOLKIN and ASTRACHAN which suggested that an RNA homologous to the viral DNA was preferentially synthesized in an infected cell.

Before the VOLKIN-ASTRACHAN deduction could be accepted as a departure point for a more extensive search for this type of RNA, its existence had to be proved. This was accomplished by selectively separating the newly synthesized RNA from the bulk of the RNA of the cell. Separation was achieved by both zone electrophoresis in starch columns and centrifugations in sucrose gradients. In addition to possessing a base composition homologous to T2-DNA, T2-RNA was found to have a lower average size and higher electrophoretic mobility. It was further shown that T2-specific RNA was ribosome-bound but with a linkage much more sensitive to disruption than normal ribosomal-RNA.

Having established T2-specific RNA as a physical entity and provided methods for its selective enrichment, it was possible to go to the next stage of the investigation and inquire into the meaning and significance of the similarity of base ratios between it and T2-DNA. This was done in terms of attempts at hybrid formation. It was possible to show that RNA-DNA complexes were indeed formed when mixtures of single stranded T2-DNA and purified T2-specific RNA were subjected to slow cooling. The success of the hybridizing experiments suggested immediately that the original observation of a similarity in base composition between T2-RNA and DNA was, indeed, a reflection of a more profound homology. The fact that hybrid formation was found to be unique to the homologous pair led to the conclusion that the nucleotide sequences of T2-RNA and DNA are complementary.

The next obvious step was to look for the presence of native complexes in systems actively synthesizing enzyme. Using detection procedures very similar to the ones employed in artificially producing hybrids it was possible to exhibit evidence for the existence of natural DNA-RNA hybrids in the *E. coli* T2 system.

Finally came the question of universality. The principle of biological unity would suggest that the mechanism of information transfer in normal uninfected cells should also involve a complementary RNA. In an attempt to provide evidence for this belief attention was turned to cultures subjected to a transition from a rich medium, which supported fast growth, to a synthetic one in which the growth rate was lower. Evidence had accumulated which led us to believe that cells undergoing such a shift were analogous in their relative RNA and protein synthesizing capacities to a T2-infected cell.

The expectation that such cells were preferentially synthesizing complementary RNA was fully realized. It was possible to establish with three different organisms the existence of an RNA in normal cells having all of the properties which

had been established for the T2-specific RNA. This normal informational RNA exhibited a base ratio analogous to its homologous DNA, was metabolically unstable, very heterogeneous in size, and possessed the ability to hybridize specifically with its homologous DNA.

It would appear from the results summarized here and those reported from other laboratories that the search for the informational intermediary has been successful. All the data are happily consistent with the simplest transcription-mechanism mentioned in the introductory paragraphs, i.e., the informed intermediary is a complementary polyribonucleotide of its « parental » DNA.

BIBLIOGRAPHY

ASTRACHAN L. and FISCHER T.M., *Fed. Proc., 20,* 359 (1961).

BELOZERSKY A.N., in 11th Solway Conference; « Nucleoproteins » Interscience publishers, New York City, pp. 99-129 (1959).

BRITTEN R.J. and ROBERTS R.B., *Science, 131,* 32 (1960).

DOTY P., MARMUR J., EIGNER J. and SCHILDKRAUT C., *Proc. Nat. Acad. Sci., 46,* 461 (1960).

FURTH J.J., HURWITZ J. and GOLDMANN M., *Biochem. and Biophys. Res. Comm., 4,* 362 (1961).

GIERER A. and SCHRAMM G., *Nature, 177,* 702 (1956).

HALL B.D. and SPIEGELMAN S., *Proc. Nat. Acad. Sci., 47,* 137 (1961).

HAYASHI M. and SPIEGELMAN S., *Proc. Nat. Acad. Sci.* (in press) (1961).

HURWITZ J., BRESLER A. and DINGER R., *Biochem. and Biophys. Res. Comm., 3,* 15 (1960).

JACOB F. and MONOD J., *J. Mol. Biol.* (in press) (1961).

KURLAND C.G., *J. Mol. Biol., 2,* 83-91 (1960).

MARMUR J., *J. Mol. Biol.* (in press) (1961).

MARMUR J. and LANE D., *Proc. Nat. Acad. Sci., 46,* 453 (1960).

MESELSON M., STAHL F.W. and VINOGRAD J., *Proc. Nat. Acad. Sci., 43,* 581 (1957).

NOMURA M., HALL B.D. and SPIEGELMAN S., *J. Mol. Biol., 2,* 306-326 (1960).

OCHOA S., BURMA D.P., KROGER H. and WEILL J.D., *Proc. Nat. Acad. Sci., 47,* 670 (1961).

RICH A., *Brookhaven Symposia in Biol., 12,* 17 (1958).

ROBERTS R.B., ABELSON P.H., COWIE D.B., BOLTON E.T. and BRITTEN R.J., *Studies of Biosynthesis in « Escherichia coli »,* Carnegie Institution of Washington, Publication 607, p. 5 (1957).

ROTMAN B. and SPIEGELMAN S., *J. Bacteriol., 68,* 419 (1954).

SAGIK B., GREEN M. and SPIEGELMAN S., in Manuscript (1961).

SPIEGELMAN S., *Protein and Nucleic Acid Synthesis in Subcellular Fractions of Bacterial Cells* in *Recent Progress in Microbiology,* Symposia held at VII Internat. Congr. for Microbiology, pp. 81-103 (1958).

SPIEGELMAN S., HALL B.D. and STORCK R., *Proc. Nat. Acad. Sci., 47,* 1135 (1961).

SPIEGELMAN S., DOI R.H. and YANKOFSKY S., in Manuscript (1961).
STEVENS A., *Biochem. and Biophys. Res. Comm.*, *3*, 92 (1960).
VOLKIN E. and ASTRACHAN L., *Virology*, *2*, 149 (1956).
WATSON J.D. and CRICK F.H.C., *Nature*, *171*, 964 (1953).
WEISS S.B., *Proc. Nat. Acad. Sci.*, *46*, 1020 (1960).
WEISS S.B. and NAKAMOTO T., *Proc. Nat. Acad. Sci.*, *47*, 694 (1961).
WOESE C.R., *Nature*, *189*, 918 (1961).
YČAS M. and VINCENT W.S., *Proc. Nat. Acad. Sci.*, *46*, 804 (1960).

DISCUSSION

RICH

It may be possible to use the positions of the bands as an indication of the proportion of the DNA which is hybridized with the RNA. Perhaps you have something to say about this if you have done that calculation.

SPIEGELMAN

The experiments we have performed so far have not permitted an accurate calculation of the sort desired. We are, however, engaged in experiments at the present time which should provide us with the data permitting an estimation of the proportion of the DNA which is hybridized. A very rough estimation from existing data would suggest a number between one and five per cent, but it is already evident, from our current efforts, that this number can be increased by a factor of ten.

PERUTZ

There was another point. In the RNA polymerase system it has been proved that both strands of the DNA helix act as primers. But is it really sure that this is so *in vivo*?

SPIEGELMAN

At the present time data from *in vivo* studies do not eliminate the possibility that both strands of the DNA helix serve as primers

for the RNA polymerase. The ideal experiment would be to isolate complementary RNA synthesized *in vivo* and perform a saturation experiment with homologous DNA. If one can show that only 50% of the DNA can be hybridized and, further, that the residue cannot form hybrids with RNA, the use of one strand will have been established. Such experiments, while technically difficult, are feasible and will undoubtedly be successfully performed soon. They should answer the question definitively in both a positive and negative sense.

[22] *Spiegelman* - pag. 42

THE INTERACTION BETWEEN NUCLEOTIDE COENZYMES AND ENZYME PROTEINS

HUGO THEORELL

Biokemiska Avdelningen, Medicinska Nobelinstitutet
Stockholm (Sverige)

In the true nucleic acid-protein-compounds the bonding between the two kinds of macromolecules is, of course, very complicated. No doubt all sorts of ordinary chemical bonds play a role; for instance, in the nucleoproteins evidently the phosphoric acid residues on the surface of nucleic acid helices must form salt-like linkages with the basic amino acid residues of the proteins. But we should notice the profound difference between single nucleotides and the highly polymerized nucleic acids. Nucleotides, and nucleotide coenzymes, are in a way much more versatile in their capacity to form different compounds with proteins, because their three-dimensional shape is not rigidly fixed through hydrogen bonding. Therefore nucleotides can presumably use practically any part of their molecules to form bonds with proteins.

The nucleic acids, on the other hand, according to the CRICK-WATSON-WILKINS model, have their purine and pyrimidine moieties hidden and hydrogen-bonded pairwise to one another in the interior of the macromolecules and therefore probably inaccessible to ordinary chemical bonding with protein groups.

I think it is correct to say that the relative importance of ordinary chemical bonding decreases parallel with increasing molecular size and rigidity, so that complex formation between macromolecules will depend to a large extent on mutually fitting molecular shape. I hope Dr. PERUTZ will agree this could be the case with the four subunits of the hemoglobin molecule. And in the double helices, f.ex. in viruses, I think the fitting shapes are of preponderant importance compared with what we generally define as chemical bonds between the two helices.

The nucleic acids and the high molecular proteins are highly specialized for a symbiotic function, as a matter of fact to reproduce one another. This function is in a way the most complicated in Nature — but the type of bonding is monotonous from a chemical point of view.

Interaction reactions and complex formation between the low molecular mono- or di-nucleotide coenzymes form the basis for innumerable and very different enzymatic reactions. Generally spoken, the binding sites for coenzymes and substrates on the enzyme surface serve the purpose of bringing these into an optimal spacial arrangement for reacting. Obviously this requires a great many different types of binding for different substrates, which are nearly endless in number, whereas Nature has confined itself to use only a few types of coenzymes, but these are bound to the enzyme in various manners.

We are only in the very beginning of finding out how a few of these complicated reactions work, but I thought it might be of interest to bring some of our results to your attention at this conference.

In some cases protein and coenzyme can form an active complex without the assistance of any metal, as for example flavinmononucleotide in the old yellow enzyme. Some years ago, using fluorimetric determinations of the association and dissociation rate constants of FMN and the enzyme protein, we could demonstrate that there are at least two bonds between FMN and protein. The phosphoric acid residue of FMN is

joint to a basic amino acid side chain, probably ε-lysine, by a saltlike bond. This bond causes a millionfold stabilization of the weak bond operating between the imino group [3] in the urea ring of the flavin, and a tyrosine residue in the protein. This work was made with Dr. AGNAR NYGAARD. The substrate is in this case reduced triphosphopyridine-nucleotide, which is reversibly bound to the same enzyme molecule, as recently found by my collaborators EHRENBERG and LUDWIG. Which chemical groupings are involved here, we do not know.

In other cases metals like magnesium or manganese or many others are necessary, for example in the transaminases. In these cases the metal ions are reversibly bound in chelates.

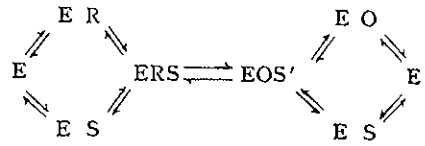
In a third type of enzyme systems metal is irreversibly bound to the enzyme, and serves as an active center. The haemoproteins are a wellknown example. However, to-day I would like to describe some very recent work we have done on a different system, the liver alcohol dehydrogenase, which uses diphosphopyridine-nucleotide as coenzyme. The enzyme molecule contains one atom of zinc in each of its two binding sites, and it seems from our experiments that the zinc functions by binding both coenzyme and substrate to ternary, reactive complexes. The zinc is very strongly bound to the protein.

Many collaborators have taken part in this work at different times (R. BONNICHSEN, A. WASSÉn, B. CHANCE, A. NYGAARD, A. WINER, J.S. MCKINLEY MCKEE, CH. WORONICK, R. PLANE and others). A great deal of effort has been spent on determining kinetics and equilibria, with or without inhibitors. Some of the « inhibitors » have shown ambivalent properties, decreasing the reaction rate at certain concentrations of the reaction partners, increasing it at others.

The horse liver alcohol dehydrogenase was crystallized by BONNICHSEN and WASSÉN in the Nobel Institute in 1948, and has ever since proved to be remarkably suitable for experimen-

tation. It is comparatively easy to obtain in a pure state and is under suitable conditions stable for months or years. In addition the combination of the reduced coenzymes, « DPNH » (disphosphopyridine nucleotide) leads to light absorption (THEORELL and BONNICHSEN, 1952) and fluorescence (BOYER and THEORELL, 1956) changes which have been of great importance for studying equilibria and kinetics.

Its reaction mechanism may be represented by this scheme:



Here E represents free enzymes, ER its compound with DPNH and ES with aldehyde; EO *vice versa* is E · DPN, ES' = E · alcohol. ERS and EOS' represent the ternary complexes with both substrates and coenzymes which are in rapid equilibrium. It has recently been finally proved that this somewhat complicated reaction scheme can be simplified down to a scheme with 6 instead of 16 rate constants. This implies the advantage that 6, or even 8 but not 16, constants can be determined in kinetic experiments, and the results checked against independent equilibrium constant measurements.

Already in 1951 (THEORELL and CHANCE) we tentatively proposed a mechanism:

$$E + R \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ER \qquad (1)$$

$$ER + S \underset{k_3'}{\overset{k_3}{\rightleftharpoons}} EO + S^1 \qquad (2)$$

$$E\ O \underset{k_1'}{\overset{k_2'}{\rightleftharpoons}} E + O \qquad (3)$$

As seen this formula neglects the formation of substrate (alcohol, aldebyde) compounds with the enzyme or enzyme-coenzyme complexes.

It has taken years to collect accurate enough experimental evidence to prove whether this simplification could be kinetically allowed; our recent results, however, seem to be convincingly in favor of the "THEORELL-CHANCE" mechanism [1-3]. We shall not now go into details on this point. It may be enough to emphasize that the knowledge of the mechanism has been the basis of explaining the inhibition and stimulation effects caused by different substances.

Fig. 1 shows the influence on the initial reaction velocities of adding different concentrations of imidazole to enzyme - DPN - alcohol mixtures.

In this "LINEWEAVER-BURK" presentation the inverse of the varied alcohol concentrations at constant concentration of DPN (abscissa) is plotted against the inverse of the reaction rate (v) per enzyme concentration, e, (ordinate), (fig. 1). In fig. 2 we see the *vice versa* case with constant alcohol concentration and varied (DPN). It is evident that imidazole at low concentrations of the varied reaction partner (to the right in the pictures) causes inhibition, but at high concentrations of the varied partner stimulation. The plots thus intersect with the plots without imidazole (x) at points where the reaction rates are the same, and imidazole therefore apparently has no effect at all. The explanation of this effect was given by some experiments which I shall mention briefly.

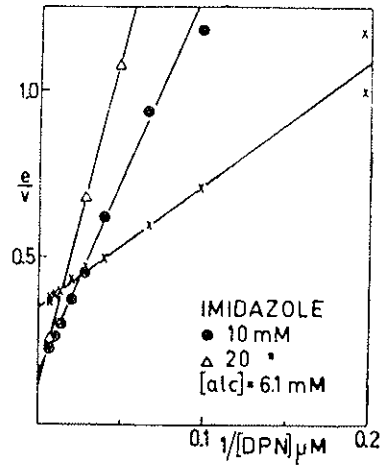


FIG. 1 — Lineweaver-Burk plots [DNP] varied, [alcohol] = 6.1 mM, pH = 7, 23.°5 C, μ = 0.1, phosphate
× without imidazole
● 10 mM »
△ 20 mM »

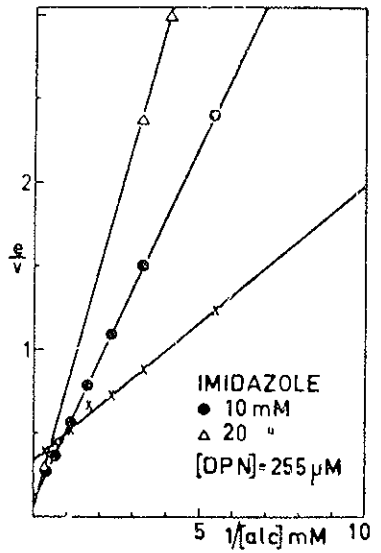


FIG. 2 — Some as Fig. 1, but [alcohol] varied at [DPN] = 255 μM. Intercepts with ordinate and slopes corrected to [DPN] = ∞

The enzyme forms binary and ternary complexes with the coenzymes and different inhibitors. Those with the reduced coenzyme, DPNH, are fluorescent to various degrees. As seen in fig. 3, curve 4, free DPNH when excited in the 340 mμ absorption band gives a moderately intense emission band. When coupled to liver ADH the DPNH fluorescence increases and shifts its maximum to a shorter wavelength (curve 6). When imidazole is added a ternary E - DPNH - imidazole compound is formed with still more intense fluorescence (curve 7).

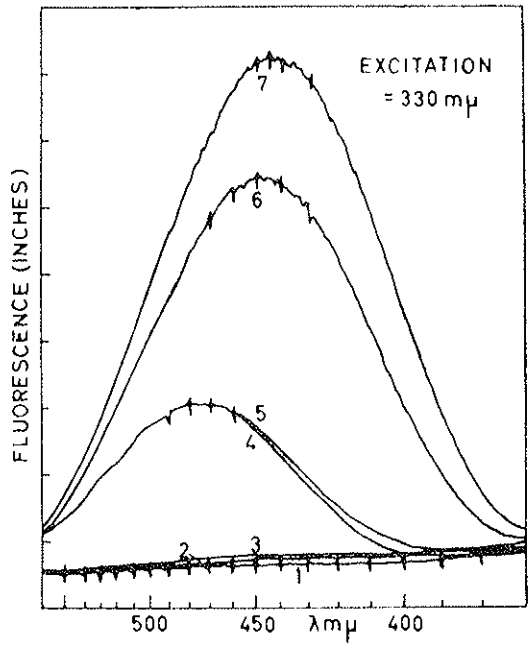


FIG. 4 — Fluorescence emission spectra of
1: Buffer
2: » + imidazole
3: » + » + ADH
4: » + DPNH
5: » + » + imidazole
6: » + » + ADH
7: » + » + » + imidazole

These effects, first found for E+DPNH by BOYER and myself in 1956, give excellent possibilities of determining the equilibrium constants of these complexes by the aid of spectrophotofluorimetry. Due to the shifts in wavelength of the maximum intensity the differences are especially large on the short wavelength side of the maxima. The intensities were therefore recorded at 410 mμ.

The compounds with DPN do not fluoresce, but can nevertheless be studied fluorimetrically, because DPN and DPNH strictly compete for a binding site at the enzyme surface. Fig. 4 shows an experiment for determining the dissociation constant operating between ADH - imidazole, DPN and DPNH.

Successive additions of DPNH to a mixture of ADH, imidazole and DPN give increments in fluorescence which decrease with the saturation with DPN. Comparing these results with titrations without DPN added, and without imidazole allows calculations of all dissociation constants involved, K_{ER}, K_{EO}, K_{EI}, $K_{ER,I}$, $K_{EO,I}$, $K_{EI,R}$ and $K_{EI,O}$, where E = ADH, R = DPNH, O = DPN, I = imidazole.

The results of hundreds of such experiments are summarized in Table I.

TABLE I

Dissociation Constants for Complexes of LADH with Coenzymes and Imidazole (uncharged, alkaline form). μ-0.1 23° .5C.

pH	K_{ER} μM	$K_{E,I}$ mM	$K_{ER,I}$ mM	$K_{EI,R}$ μM	$K_{E,O}$ μM	$K_{EO,I}$ mM	$K_{EI,O}$ μM
6	0.23	0.47	4.2	2.0	266	0.8	450
7	0.31	0.55	3.5	2.0	160	1.5	430
8	0.41	0.68	3.6	2.2	51	5.9	440
9	0.65	0.67	3.0	2.9	12	24.2	435

These experiments gave some valuable information. Liver ADH contains 2 atoms of Zn per molecule, and the same num-

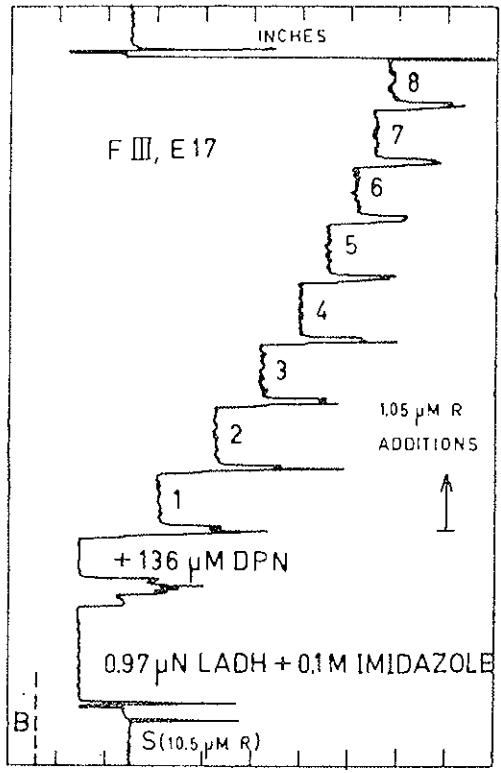
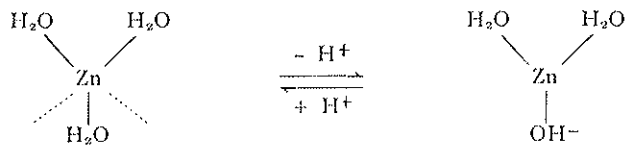


FIG. 4 — Determination of dissociation constants by spectrophotofluorimetric titration

ber of binding sites for DPNH. It is very hard to imagine that imidazole in its uncharged form would combine strongly with any ordinary group in a protein, but Zn-imidazole complexes are well known from the literature, and the values found for K_{EI} are of the same order of magnitude as for Zn ions and imidazole. In these complexes neutral imidazole uses the free electron pair of its ternary nitrogen to coordinate with Zn, and a maximum of 4 imidazoles per 1 Zn can be attached in this way. The results given in table I indicate that Zn in LADH

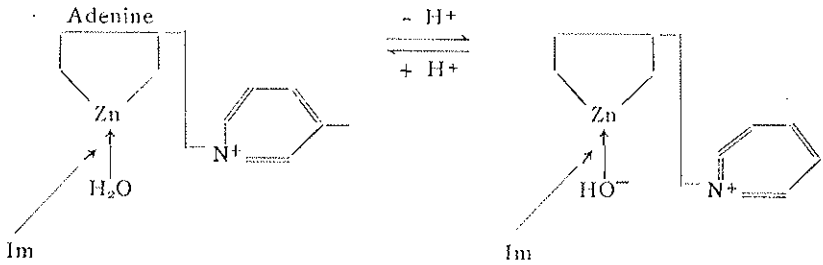
can bind only one imidazole. Zn might therefore be bound by three out of a total of six octahedral bonds to ADH, leaving the three other ones free for reacting with coenzymes and substrates. This would explain the very firm bonding of Zn to the enzyme protein. Experiments (with R. PLANE) on the effect of inhibitors like o-phenantroline, ethylene diamine and others were entirely in line with this theory. I shall not go into details on these items here. However, our conclusion is that Zn in LADH takes part in the binding of both coenzymes, using two of its three free bonds for attaching perhaps the adenine moiety — as earlier suggested by WALLENFELS and SUND — and of substrates, using the third bond. As seen from table I most of the dissociation constants operating between ADH, coenzymes and imidazole are scarcely influenced by pH-changes between 6 and 9. The two exceptions are K_{EO} and $K_{EO,I}$ which show an inverse proportionality. We explain this by assuming that at increasing pH one of the water molecules at lower pH-values occupying the three free Zn bonds loses a proton (pK = 8.6), leaving a negative hydroxo group on the Zn:



pK = 8.6

In the complex with DPN ("EO") the positively charged pyridine ring interacts, we think, with this hydroxo group so that a stabilization of the binding of DPN to ADH occurs, in agreement with the changes in K_{EO} with pH (see table I). Since two of the free Zn bonds are already occupied by the

adenine part of the DPN, imidazole has to compete with the pyridine ring for the third bond. This explains why $K_{EO,I}$ increases with pH, whereas $K_{EI,O}$ is insensitive to such changes.



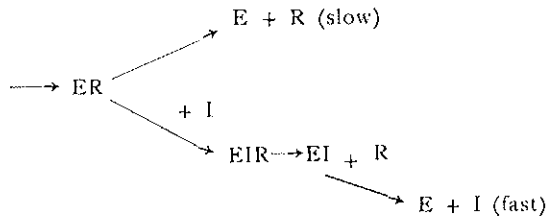
$pK = 7.07$

It seems very plausible to believe that alcohol in the form of an alcoholate ion — the dissociation favored by the proximity to the pyridine $\overset{+}{N}$ — substitutes the hydroxo group.

In work with A. WINER fatty acids were found already a couple of years ago to form ternary complexes ADH - DPN - fatty acids, but no ternary complexes with ADH - DPNH. This is obviously to be expected from the suggested arrangement, since the negatively charged carboxyl group must be stabilized by the positive pyridine $\overset{+}{N}$. Fatty acid amides form ternary complexes with ADH-DPNH ("ERI"), but not with ADH - DPN (EOI). This indicates that the binding sites for DPN and DPNH, and for fatty acids and their amides, must be different. It was found in kinetic experiments that fatty acids compete with alcohol and with DPNH, whereas the amides compete with aldehyde and with DPN. Now it must be noticed that the three free bounds of the Zn are not equivalent, being fixed in their positions relative to the enzyme surface. The oxidoreduction occurring between DPN-alcohol and DPNH-aldehyde therefore seems to involve an « edge

shift » at the Zn bonds. Suppose the adenine of DPN uses Zn-bonds A and B, alcohol C, the adenine could flick over to A and C when DPN is reduced, and the aldehyde formed to B. This mechanism would explain one mystery in Nature: how the formation of the non-reactive, useless complexes ADH - DPNH - alcohol and ADH - DPN - aldehyde is prevented. This prevention seems to be total at physiological concentrations of the reactants; at very high alcohol concentrations, however, we found indications of an ADH - DPNH - alcohol complex to be formed. We think the alcohol in such cases is forced to form a weak bond with the aldehyde site.

Let me finally turn back to the question how imidazole can either stimulate or inhibit the reaction velocity. The answer is that these effects occur in different phases of the reactions. In the « backwards » reaction (see the « THEORELL-CHANCE » mechanism, formulae 1, 2 and 3), starting with E + O, at high concentrations of both O and S', the dissociation rate constant, k_2, of ER will be rate limiting, and most of the enzyme is in the form of ER. When imidazole is present ERI will be formed to an extent depending on the imidazole concentration and on $K_{ER,I}$. Now according to table I, ERI is less stable than ER, with higher dissociation constants than ER or EI. We found this to be dependent on the dissociation velocity constants being increased, whereas the association velocities remain practically unchanged. This means that the decomposition of ER to liberate free E for next reaction cycle has now in the presence of imidazole found a faster pathway.



[23] *Theorell* - pag. 12

Inhibition occurs when either DPN or alcohol are present in low concentrations. The dissociation of ER is then only partly rate limiting. E and EO will be considerably present, leading to the formation of EI and EOI, that is part of the enzyme is inhibited. At some intermediate concentrations stimulation by ERI formation will cancel inhibition by EI and EOI formation; this is where the LINEWEAVER-BURK plots intersect.

I can very well imagine that such effects may be of considerable biochemical and pharmacological interest. Substances acting like imidazole — and we have found many others — will have a homeostatic function, helping to maintain a constant substrate concentration. An occasional excess will be rapidly metabolized, and when the concentration goes down increasing inhibition occurs. Perhaps some hormones act in this way — that we do not know yet; but it should certainly be possible to find drugs with this kind of homeostatic function.

DISCUSSION

LINDQVIST

The complex formation of zinc with alcohols and fatty acids was studied by MEERWEIN, who showed that the acidity increased for both these types of compounds. Now, can you be sure about the octahedral coordination around zinc? Tetrahedral coordination is very common with nitrogen or sulfur as ligands, while the octahedral is generally obtained by oxygen ligand. The existence of both types in aqueous solution has been discussed by ORGEL.

THEORELL

We have considered the possibility that zinc in the free enzyme could be held by tetrahedral bonds, thus probably by nitrogen (f.ex. imidazole) and sulfur groups. In the ternary complexes the bonding could then change to octahedral, perhaps by the aid of the oxygen atoms in the substrates, or inhibitors like fatty acids or their amides. Some difficulty arises with the ternary imidazole complexes; in case two adenine nitrogen in the coenzymes are attached to zinc, there would be no oxygen added by the introduction of imidazole. This would speak against the adenine part being attached to zinc. On the other hand we think there is no doubt about the zinc having three bonds to the protein, and that both coenzymes and substrates are bound to zinc. This requires more than four bonds; that is to say, octahedral configuration in the ternary complexes.

PUTZEYS

Ovalbumine crystallizes with 3.6 zinc ions for one ovalbumine molecule (1). If there are seven histidine molecules in one ovalbumine, this means two imidazols for one zinc. Obviously one may not infer from this, may be fortuitous, stoichiometry, that the complex formation is really with two imidazols.

ANFINSEN

Professor THEORELL, what sorts of methods have you used to investigate the possibility that alcohol dehydrogenase is a dimer? With myosin for example, HARRINGTON and KIELLEY found that 4M guanidine was not sufficient but that at 6 M, monomers were produced from the normal myosin trimer.

THEORELL

We have not made systematic investigations into this, but we have never observed any tendency to splitting into monomers. However, I thank it is very well possible that guanidine might do the trick.

(1) P. PUTZEYS - E. SWAELEN, *Med. Kon.* VI. Acad. Wetensch. België (1960), *22*, n. 3.

SOME RECENT ADVANCES IN BIOCHEMICAL SEPARATION METHODS AND THEIR SIGNIFICANCE AS AN ANALYTICAL APPROACH TO THE STUDY OF BIOLOGICAL STRUCTURES

ARNE TISELIUS

Biokemiska Institutionen, Uppsala Universitet, Uppsala (Sverige)

Biochemistry, being the branch of chemistry which aims at an investigation of the mechanism of the processes of life, has had and will always have as one of its most important goals to isolate substances out of biological materials which somehow play an important part in the complex phenomena which we call Life. This problem has come in to the foreground not least because many of the substances involved owing to their molecular size, their complicated structure and their fragility to common chemical operations, present special problems. Furthermore many of the most important substances occur in extremely small quantities. Consequently it is natural that analytical biochemistry has come to occupy a key position and that special methods, which take the difficulties mentioned into consideration, have had to be elaborated. The purpose has aften been to isolate an enzyme or some other biologically active substance in pure form and to characterize it chemically as far as possible. Often one has had to stop here — first lately, as you know, one has succeeded in realizing a real chemical structure analysis of such complicated proteins as

insulin and ribonuclease, so that the exact sequence of amino acids in these molecules is now known. Likewise X-ray crystallographical methods, in some few cases (myoglobin, hemoglobin), have gradually made an almost complete localization of the atoms in the molecule (M = 67,000) possible. Other proteins are now subject to similar structure investigations, but we still have to admit that this important branch of the biochemistry on the whole still is on the pre-structural level. And also when dealing with the best known structures we are still very far from an « explanation » of the important biochemical functions these substances show in the living organism or as isolated substances (e.g. enzymatic activity).

However, the aim of the biochemist can not be restricted to reducing his problems to merely chemical terms by purification and structure determination of important substances from biological material. When discussing the problem of the *function* of these substances we must never forget that this function in nature is performed in a very special environment — the living cell — and often is an integral part of a system of chemical reactions connected to each other, which also have a structural basis, though of another character. We deal here with biological-morphological structures which generally have been studied with microscopic or electron microscopic methods, sometimes including also microspectrophotometry of different kinds or X-ray crystallographical analysis. It is true that even these structures must have a chemical background, but chemical methods used till now are generally not sufficient here. We know far too little of the laws which decide how macromolecules and other material associate to the very specific and highly organized structures, which we can observe in the cell by the methods mentioned. We have every reason to assume that we have here new and promising fields for chemical research and that these « biological » structures are also of fundamental interest to the chemist.

It is not my intention to enter upon a further general discussion of these in themselves so attractive problems. Instead I thought of illustrating the character of these problems and the methods for their study with some examples — especially from my own laboratory —. As to the methods I intend to, as far as time admits, tell you about some new advances at my own laboratory which may be of some interest also for other problems of biochemistry and chemistry than those involved in the main theme of this lecture.

In the blood of the well-known snail *Helix pomatia* there is a blue protein, hemocyanin, which serves as a carrier of the oxygen in the blood in the same way as the hemoglobins in higher animals. Through examinations by SVEDBERG and his collaborators [1] in 1930 it appeared that the hemocyanin is a real giant molecule (M = 6.7 millions), but that this molecule can be splitted *reversibly* to fragments of 1/6 of the original size through a shifts of the pH-value from neutral reaction (pH = 6.8) in the alcalic direction (pH = 8.5). Interestingly enough, on neutralization, the primary molecules are reformed. We come across similar phenomenon with other hemocyanins and on the whole with a number of proteins, especially those with high molecular weight. Here we have an example of an unusually simple and clear-cut case of reactions between biological macromolecules in solution. Today we have the possibility of observing directly in the electron microscope not only the hemocyanin molecule but also something of its inner structure. Fig. 1 shows such a picture (Dr. Ö. LEVIN, Uppsala) where we can see that the molecules are short cylinders (1). Some of these lie on their side (the rectangles), some are standing upright (the circles), and they appear to have a six-fold structure which would explain their dissociation into sixths.

In 1937 Dr. FRANK HORSFALL and I [2] made some experi-

(1) See also: VAN BRUGGEN, WIEBENGA and GRUBER, *Biochim. Biophys. Acta*, *42*, 171 (1960).

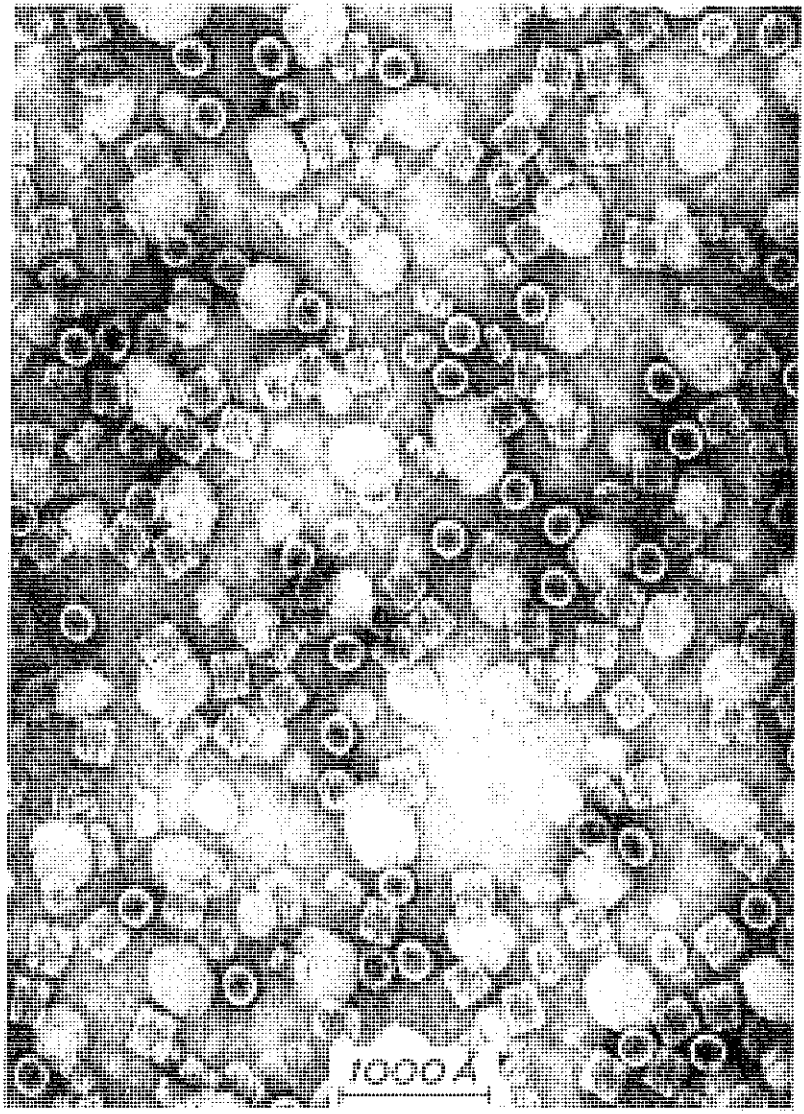


FIG. 1 — Electron micrograph of molecules of *Helix pomatia* hemocyanin (Ö. LEVIN).

ments in order to see what happens if one performs such a reversible recombination operation with a mixture of two *different* hemocyanins. We chosed *Helix pomatia* and *Helix nemoralis* hemocyanins, which easily can be separated and identified by electrophoretic analysis. The experiments showed that in this case the recombination does *not* direct to a reformation of the two primary kinds of molecules, but we obtain a series of « hybrid » molecules, almost corresponding to what one might expect the following combinations: 1 Hp+5 Hn, 2 Hp+4 Hn.....5 Hp+1 Hn, distributed according to a statistical Gauss curve. The same experiment was repeated but with two hemocyanins from more distant species, namely Helix pomatia and Litorina litorea. Here the primary molecules were reformed on recombination, at least to a considerable extent.

These experiments allow some conclusions of interest. Obviously, those affinities, which keep the fragments together, will consist of bonds which each are rather weak, as the dissociation occurs through a relatively moderate pH-shift (2 units). The aggregates (the whole molecules) may still be rather stable because of the large number of such bonds which keep them together. The pronounced biological specificity which is reflected in the results of the recombination experiments might also be explained as a consequence of a cooperation of many affinities, arranged according to a specific pattern, to some extent perhaps analogous to a kind of crystallization.

Lately similar recombination experiments have been performed with hemoglobins especially by ITANO and his collaborators at the National Institutes of Health in U.S.A. These experiments go back to the discovery (1949) by LINUS PAULING, ITANO and their associates of the so called sickle-cell hemoglobin (HbS) which appears at certain hereditary pathological states (anaemias) in human beings. This hemoglobin has the same molecular weight (67,000) as normal hemoglobin (HbA) and is also in other properties very similar to HbA. As

a matter of fact we know now that the difference between the molecules are restricted to one single residue of amino acid per half of the molecule ((glutamic acid in HbA is replaced by valine in HbS), which means a difference of only 1:300 amino acids. This, however, is sufficient for demonstrating the difference by electrophoretic analysis, and it was through studies of electrophoretic diagrams that HbS was first discovered. Today we know at least 20 pathological human hemoglobins, which might be differentiated electrophoretically or chromatographically. The hemoglobins are dissociated reversibly at moderate pH-shifts in the same way as the hemocyanins, but the process is more simple here as hemoglobins dissociate into halves, and thus the alternative recombination possibilities are fewer and the experiments easier to interpret. As an example I would like to show a figure (ITANO and ROBERTSON [3], *183*, 1799, 1959) with HbA recombined with HbS *or* HbI. From the figure (*c*) you can see that A increases, while S and I decrease, and that a new component appears between A and I — a double abnormal hemoglobin. Similar results were obtained with HbC instead of HbS, though no new component could be observed here. The increase in A (relatively to the two other components) is clear, however. Consequently, it is possible to produce « hybrides » of related proteins through such simple « in vitro » recombination experiments. It does not concern just « curiosities » as the formed hybrides have biological (pathological) properties different from those of their « parents ». It seems probable that such phenomena also appear in other important relations. Thus it is known that the microsomes and other submicroscopic cell particles may undergo a reversible dissociation-association-reaction, where the magnesium concentration is the determining factor. I do not know if recombination experiments with microsomes from different species have been performed yet, but such experiments would no doubt be of great interest. Even viruses, in any case the simpliest types like tobacco mosaic virus consist as you know of sub-

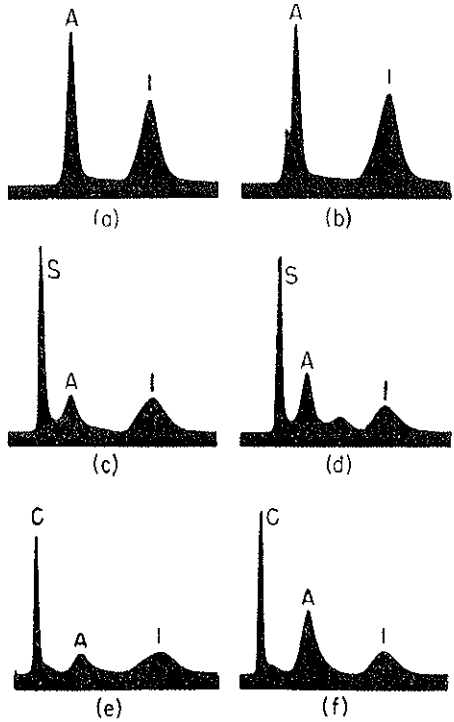


FIG. 2 — Recombination experiments with hemoglobin A, S, I and C, as studied by electrophoretic analysis. *a*) A, I control - *b*) A, I recombined - *c*) S, A, I control - *d*) S, A, I recombined - *e*) C, A, I control - *f*) C, A, I recombined. (ITANO and ROBINSON [3]).

microscopical particles containing protein and ribonucleic acid. These, by appropriate treatment, can be dissociated from each other and be recombined, which first was proved by FRAENKEL-CONRAT in Berkeley and GIERER and SCHRAMM in Tübingen. It is also well known to the participants in these discussions that highly significant recombination experiments recently have been performed with DNA molecules by a heating-annealing cycle (see for example [4]). Such experiments open very wide perspectives indeed as one here is able to « manipulate »

with some of the very key substances in reproduction processes. But they have their significance also as a way to that aim I mentioned as especially important in the beginning of this lecture: an elucidation with chemical-analytical methods of those affinities and those structure relationships which are important to an understanding of the character and function of the biological structure.

The methods which have been used in these experiments in order to study the structure of certain biological particles are to some extent analogous to those which the chemist uses when he tries to establish the structure of an organic compound. It is true that the analytical methods are different — they have to be adapted to the complicated and delicate structure of the biological material, but the principle is the same: a break down, an identification of the fragments and, if possible, a resynthesis. In the « biological » case the fragments are not atoms or smaller atom groups but complicated macromolecules the structure of which we generally do not know in detail, but which we anyhow can identify with some confidence. But otherwise the method is analogous in principle at least.

We have every reason for asking ourselves if not a similar procedure could be applied when dealing with biological structures in general, that is not only with more a less well defined cell particles, viruses and macromolecules, but also e.g. for cell membranes, cell nuclei, chromosomes — perhaps even for biological tissues. I pointed out in the beginning of this lecture how the biochemist not only has to try to isolate important substances out of biological material in as pure form as possible. Furthermore, he has to take interest in how these substances fit into the living cell, where they exercise their functions together with others, to which their action is linked through some kind of structural relation. We could in a somewhat pointed formulation say that the biochemist not only ought to be interested in the pure substances but also in the impure ones. I want to illustrate this with a figure which is supposed to

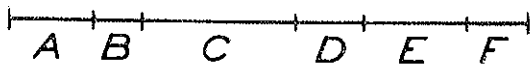


FIG. 3 — A linear structure consisting of structure elements A, B, C, D, E and F.

represent a « linear » structure of the substances A, B, C, D, E and F (fig. 3). If you isolate e.g. C in pure form, then we actually just know that C is a part of the structure, but if certain quantities of the adjoining components B and D still remain on C then we know in addition that C in the original structure is bound to B and D. We can generalize this approach (fig. 4, *a*, *b*, *c*). Here we imagine that a certain biological structure (e.g. a cell) through some mechanical method is decomposed into fragments of gradually diminishing size (indicated by a system of squares). Suppose that we perform

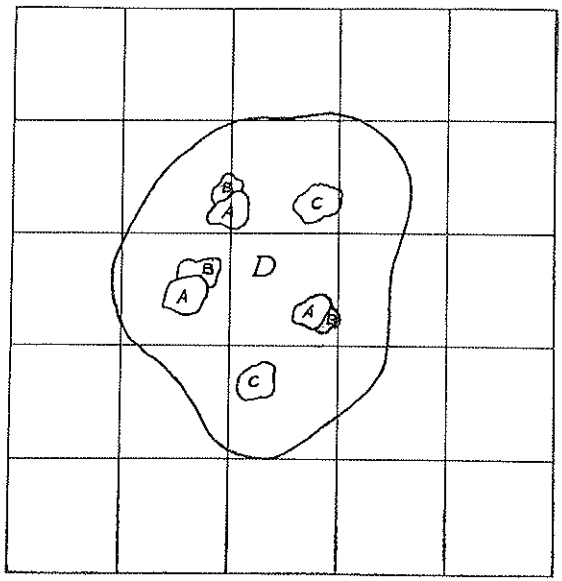


FIG. 4-*a*

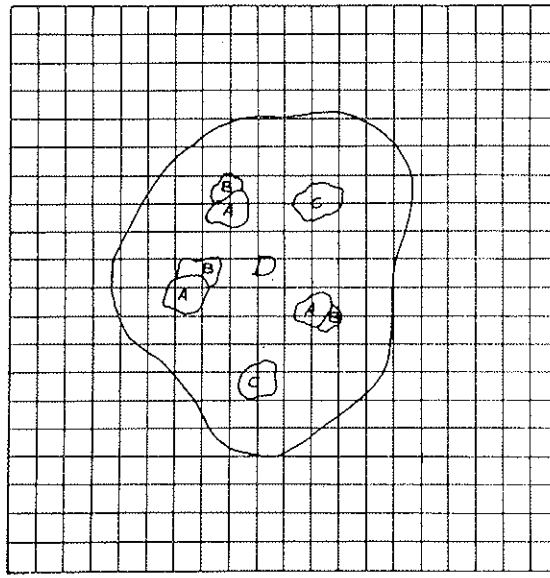


FIG. 4-*b*

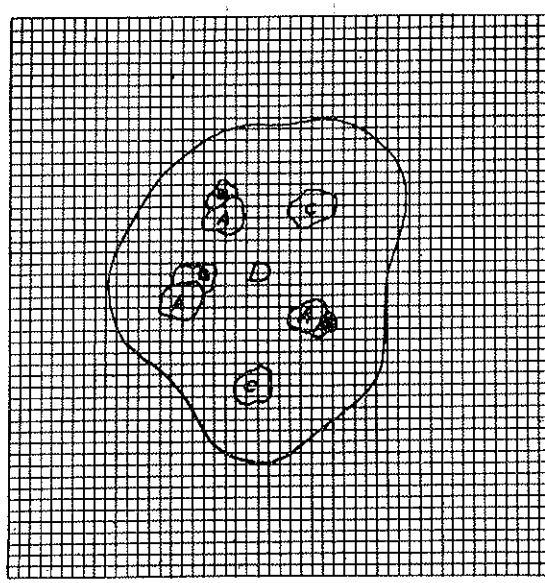


FIG. 4-*c*

Successive disintegration of a structure for analysis of the fragments.

such an operation with a very great number of cells and that we study the combination of the fragments of different size classes. We can see, that relatively big particles (fig. 1) still are able to contain both A, B and C, the medium sized only A and B or B and C, while the smallest fragments only contain one component A or B or C. That particle size at which a « segregation » occur is a measure of the « volume distance » between the components in the structure (TISELIUS [5]).

Such an analysis would thus in principle make possible a structure determination but would also facilitate a correlation of structure and function. A certain biological function, e.g. respiration or assimilation, will survive down to a certain particle size. This in relation to chemical analysis of the fragments might give a clue to such features of the structure which are essential to function.

I have been tempted to indicate this somewhat sketchy principle for an analytical method for the study of biological structures not because I think that this — still — is a practically feasible procedure but because out of this we can see what methodical advances would be especially desired for such an approach. It is evident, already from those examples I have given you on studies of certain aspects of the structure of biological particles that separation methods of the highest degree of resolution and at the same time of a very gentle character are essential, and the same is of course applicable to the more general problem. But in addition methods must be worked out where fragments of biological structures can be separated most effectively, and where fragment — or particle *sizes* in some way will enter into the picture. Actually such lines of thought have to some extent been lying behind some of the latest improvements in biochemical analysis methods which has been realized at my laboratory in Uppsala, and I shall finish this lecture by indicating something about this, briefly.

I shall not discuss details in new and old methods for the disintegration of biological material down to submicroscopical

dimensions even if these procedures are of importance in this connection. The most important recent advances have been made with methods of separating and isolating particles and macromolecules, that is of fragments of submicroscopical dimensions.

In principle the simplest method of separation in chemistry is the partition method but so far this method has been chiefly restricted to molecules of small or medium size. If one tries to use such methods with large molecules, for example with proteins, one will find that they will remain in one phase only and thus no separation is obtained. During the last few years, Dr. ALBERTSSON at the Institute of Biochemistry in Uppsala has, however, worked out some special phase systems which make possible a very specific and well defined partition even of very large molecules, particles or whole cells [6]. Also it is an advantage that these systems do not seem to have any aggresive action. ALBERTSSON's two-phase-systems consist of aqueous solutions of two water soluble polymer substances in comparatively low concentration (a few percent) and some salt. If one dissolves in water for example polyethylene glycol and dextran and some phosphate buffer one obtains a system of two phases chiefly consisting of water and separated by a very sharp boundary. As the phase-forming polymers have a low osmotic pressure the phases end to become very similar and ordinary substances will distribute themselves with a partition coefficient close to 1. Larger molecules and, above all, particles will, however, give finite partition coefficients (see table 1) which depend among other things of the salt concentration. This provides excellent possibilities of an efficient and gentle fractionation of systems which earlier have been difficult to work with. The method has already with success been used for example for virus purification, separation of bacteria, fractionation of proteins and nucleic acids etc.

Another new method which is of particular interest in this connection is the so called gel-filtration which has been worked

TABLE I

Partition coefficient, K, and the surface area of a number of protein and virus particles in the dextran-methylcellulose system D 68 - MC 4000:

Particle	Surface area ($m\mu^2 \times 10^{-3}$)	K
Phycoerythrin o	0.3	0.95
Hemocyanin, « eighth »	0.86	0.65
Hemocyanin, « whole »	3.5	0.25
ECHO virus	2.3	0.2
Polio virus	1.3	0.3
Southern bean mosaic virus		0.41
Phage ∅X 174		0.34
Phage T3	8.7	$2.1 \cdot 10^{-2}$
Tobacco mosaic virus	14.4	$(1\text{-}2) \cdot 10^{-2}$
Phage T2	25.5	$(6\text{-}10) \cdot 10^{-4}$
Phage T4	25.5	$(3\text{-}5) \cdot 10^{-4}$
Vaccinia	220	$(4\text{-}12) \cdot 10^{-5}$

out by Dr. J. PORATH at the Institute of Biochemistry and Dr. PER FLODIN at the Research Laboratories of Pharmacia in Uppsala [7]. The principle is shown in fig. 5. A cylindrical tube is packed with a suspension of gel particles (Sephadex) prepared by « cross-linking » of dextran molecules. These particles swell strongly by uptake of water but they remain insoluble and resistant even towards moderately alcaline and acid media. If a mixture of small and big molecules is allowed to pass such a column the smaller molecules will move more slowly than the larger ones as the former to some extent will penetrate into the gel and thus become retarded. The result is, as shown in the figure, a separation, so that the substances leave the bottom of the column in more or less separated zones, the larger molecules emerging first and the smaller last. Thus

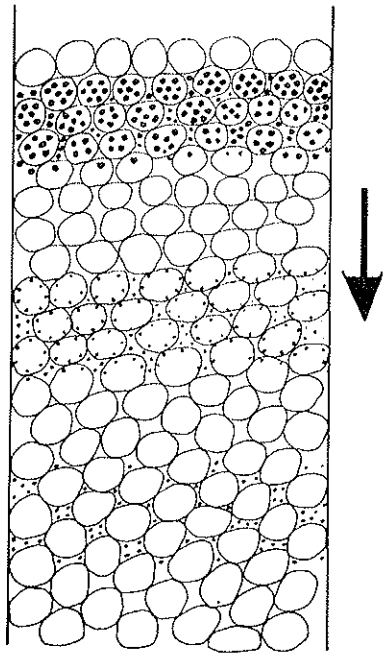


FIG. 5 — *The principle of separation by gel filtration (see text).*

this is a kind of chromatography largely based upon differences in molecular size, even though special affinities sometimes may interfere. The method has rapidly come into general use in biochemistry.

So far the application has been chiefly in the field of molecules of medium size but the range of the method is gradually being extended even to the largest molecules. Thus some recent experiments with « Sephadex 200 » have shown that serum proteins may be separated by this type of gel filtration. The method is simple, rapid and gives very good yields. Also the fact that differences in molecular size play a decisive role makes

the method particularly interesting in connection with the problems just discussed. Methods of the kind I have discussed here are useful tools in attempts to study problems of biological structures by analytical procedures. It is difficult to say yet how far we can proceed along this road. The methods referred to here will still be valuable as tools in biochemical research in general, where specific and gentle separation methods of high resolution occupy a key position.

REFERENCES

[1] SVEDBERG T. and PEDERSEN K.O.P., *The Ultracentrifuge*. (Clarendon Press. Oxford, 1940), p. 363 ff.

[2] TISELIUS A. and HORSFALL F.L., *Journ. Exp. Med.*, *68*, 83 (1939).

[3] ITANO H.A. and ROBINSON E., *Nature*, *183*, 1799 (1959).

[4] SCHILDKRAUT C.L., MARMUR J. and DOTY P., *Journal Molec. Biol.*, *3*, 595 (1961).

[5] TISELIUS A., *Arkiv f. Kemi*, *15*, 171 (1960).

[6] ALBERTSSON P.A., *Partition of Cell Particles and Macromolecules*. (Th. Wiley and Sons, New York, 1961).

[7] PORATH J. and FLODIN P., *Nature*, *183*, 1657 (1959).

DISCUSSION

CHARGAFF

I wonder whether there has been an improvement in ALBERTSSON's techniques? If one is dealing with solutes of a size not suitable for recovery by sedimentation, we have found it difficult to separate the substances we were interested in from the dextran sulfate and methyl cellulose in the two phases.

TISELIUS

ALBERTSSON's method was originally introduced for the separation of particles, and in this case — as Dr. CHARGAFF points out — it is easy to remove the phaseforming polymers by centrifuging the separated particles out of the suspension samples. When the method was gradually extended to material of lower particle size, e.g. proteins and nucleic acids, the problem of removing the polymers becomes more difficult and is now being studied intensively. In addition to using polymers of the type of dextran sulphate, which can be removed by salt precipitation there are some alternative procedures, which have proved useful. Most uncharged polymers will pass through chromatographic columnes with little or no retention, leaving proteins etc. behind for subsequent elution.

In some cases, the solute may be transferred to a polymer-free phase by partition. Thus DNA or RNA preparations will move almost completely into the bottom phase (potassium phosphate) if

the top phase is a very concentrated polyelethylenglycol solution. This type of system has been described by ALBERTSSON (1) and its application to nucleic acid fractionation by LIF (2).

CHARGAFF

What actually enters into the separation? Density or surface area?

TISELIUS

Density does not come into the picture. The decisive factor determining the partition coefficient K is the product of the surface area of a particle ($M^{2/3}$) and a factor λ which is characteristic of the surface chemical properties of the particle and constant in a family of « isochemical » substances, that is f.ex. for proteins (if the influence of charge is eliminated by sufficient salt concentration).

Thus $K \sim e^{-\frac{M^{2/3}\lambda}{RT}}$

For nucleic acid preparations a definite correlation seems to exist between K and the sedimentation constant (3).

FRAENKEL-CONRAT

The question is raised whether the extent of incorporation of tritium (T_2O) upon « reconstitution » of hemocyanine could or was used as a measure of the number of groups involved in intersubunit bonding. Such studies with TMV protein indicate a finite number of protons to become immobilized both, upon reversal of denaturation of the protein and upon the aggregation of the native protein to rod particles.

(1) ALBERTSSON P.Å., *Partition of Cell Particles and Macromolecules*, p. 69 (John Wiley and Sons, New York, 1961).

(2) LIF T. to be published shortly.

(3) LIF T., FRICK G. and ALBERTSSON P.Å., *J. Mol. Biol.*, 3, 727, 1961.

TISELIUS

So far no reconstitution experiments on hemocyanins with incorporated isotopes have been performed. This would no doubt be very interesting.

ZAMECNIK

As Professor TISELIUS has suggested, it is possible to combine the molecular sieving properties of the crosslinked dextran with ion exchange separation by using, for instance, a DEAE-dextran. Thus, one may obtain a synergistic effect under favourable circumstances. In this way, we have obtained a rather good separation of a species of aminoacyl-s-RNA from the family of closely related s-RNA molecules, and at a later time in this meeting I shall perhaps present the details of these experiments.

TISELIUS

I have seen the report of Prof. ZAMECNIK's experiments on the separation of a species of aminoacyl-s-RNA, and to me this is one of the most striking explase of the efficiency of the dextran and DAEA-dextran gel columns in separation work.

JACOB

The possibility of distinguishing different types of cells or of bacteria is of course of high interest. It may allow one for instance to select for mutants in a population. I am thinking also of another system where the method might be extremely valuable. We have tried in the passed years to determine what material, besides the chromosome, can migrate from male to female during crosses of E. coli. For this we need to label the males with isotopes, mate them with non-labelled females and then separate males from females.

I wonder how far one can separate the two types, i.e. starting with a 1:1 mixture, can one recover a population of females with less than 1 male in 10^3 or 10^4?

TISELIUS

It is difficult to make a general statement, as some systems are more favourable than others. A thorough study of the separation of bacteria in two-phase systems will appear shortly (4). Selective action on mutants has been observed in several cases.

No systematic study of the distribution behaviour of the K12 strains of E. coli has yet been made, but so far, differences in distribution pattern among these strains do not seem to depend only on the mating type. It appears that only some male and female strains have sufficiently different surface properties to allow an effective fractionation of a mixture of two types of cells. Some Hfr and F-strains have been fractionated with an efficiency of at least 99% after 50-100 transfers in counter current experiments. Particularly we have been able to fractionate mixtures of the F-strain W1177 and the Hfr strain 58, and these strains might therefore be useful for the very interesting type of experiment that Dr. JACOB has suggested.

RANZI

With regard to the reaggregation in the molecular field, may I remind you that in 1948 my collaborators and I were able to prepare myosin and actin from different species of vertebrates and invertebrates. These actins and myosins of different origin (for example teleostean fish and pigeon) can combine. In such a way it is possible to obtain hybrid actomyosins which can be found in treads which are contracted by ATP action (5).

TISELIUS

With regard to Prof. RANZI's interesting observation I wish to say that I believe that molecular recombination may be a rather general phenomenon among certain proteins in living tissue.

(4) ALBERTSSON P.Å. and BAIRD D., *Exp. Cell Research* (to be published) 1962.

(5) CIGADA, CITTERIO, RANZI and TOSI, *Experientia*, 9 (1948).

PERUTZ

Can Professor TISELIUS tell us anything about the symmetry of the haemocyanin particles?

TISELIUS

We do not at present know more than can be seen from the electron micrographs (fig. 1) which show that the molecules are cylindric, being built up of 6 sections transversal to the cylinder axis. The end view also shows some structural details, but we are not yet sure about the interpretation.

RICH

The hemocyanin molecule recombination should give rise to several isomers. For example, in the subfamily of one *Helix pomatia* + 5 *Helix nemoralis*. There could be a number of positions for the one helix pomatia. That it could be at an end, or somewhere buried inside the cylinder. Further, it would be very interesting to know why the native cylinder consists of only 6 units. It would be interesting to know if it could be reaggregated into longer units, with more than 6 parts.

TISELIUS

No doubt, such isomers may occur, but the electrophoretic resolution does not yet permit us to differentiate them out of the rather continuous-looking mixture of different recombinants. We have never observed cylinders longer than 6 units, but frequently fragments with only 3 units. Evidently the rule valid for well defined proteins — that they are essentially monodisperse systems — applies also to these giant molecules. As a matter of fact, this is the only justification for speaking of « molecules » and not just « particles » in this case.

THE STRUCTURE OF HORSE CYTOCHROME *c* AND COMPARATIVE STUDIES OF CYTOCHROME *c* OCCURING IN DIFFERENT ORGANISMS

H. TUPPY

Institut für Biochemie der Universität, Wien (Osterreich)

The cytochromes, « haemoproteins whose principal biological function is electron and/or hydrogen transport by virtue of a reversible valency change of their haem iron » [1], were discovered by MACMUNN [2] in the eighties of the last century. Their almost ubiquitous occurence in living organisms, however, and their paramount importance for cell respiration did not become apparent until KEILIN [3], forty years later, reinvestigated them and greatly extended the early findings. His work, and that of many others since, firmly established the fact that the electron transferase function of the cytochromes is one of the most fundamental and most universal activities of living cells. The association of oxidative phosphorylation with electron transfer mediated by cytochromes appears to be a common denominator throughout aerobic organisms.

Among the electron transport haemoproteins, the cytochromes of type *C* are characterized — and they can thus be differentiated from other cytochrome types — by having covalent linkages between the prosthetic group and the protein moiety. Acid acetone which readily causes catalase, haemoglobin, myo-

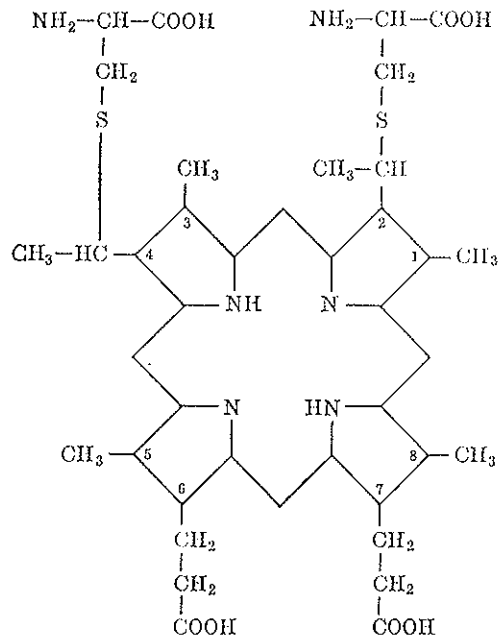
globin, and the cytochromes *a* and *b*, to dissociate into haem and apoprotein, does not affect the covalent bonds which conjugate the prosthetic group and the apoprotein in cytochromes of type *C*. Another criterion used for assigning a cytochrome to group *C* is the display by its reduced form of a haemochrome spectrum with a sharp alpha band maximum around 550 mμ. Cytochrome *c* proper will exhibit this absorption peak upon simple reduction, whereas other respiratory pigments, which require an additional treatment with pyridine and alkali in order to show the characteristic maximum at 550 mμ, have been included in the wider category of cytochromes of *type C* [1].

Cytochrome *c* has been identified in most animal and plant cells investigated [4] and in many microorganisms [5]. It is frequently associated with cytochrome c_1, another respiratory pigment belonging to type *C*; c_1 differs from *c* by showing light absorption bands shifted slightly toward the red; it has been isolated and purified from mammalian heart [6-10]. Other members of type *C* are the plant chloroplast haemoprotein, cytochrome *f* [11]; cytochrome *h*, a pigment found in the hepatopancreas of molluscs [12, 13]; and numerous haemoproteins which occur in various microorganisms and also conform to the chemical and spectroscopic criteria defined above for this group of pigments.

Cytochrome *c* proper has been extracted from a variety of organisms, the most popular sources being the heart muscle tissue of vertebrates [18-21] and bakers' yeast [14-17]. Owing to its relatively great stability, it could be subjected to rigorous operations of fractionation and purification. Crystallized samples were obtained by BODO, by PALÉUS & THEORELL, and by HAGIHARA, HORIO & OKUNUKI, from sources such as the king penguin [22], from beef, pig and horse heart [23-25], from the whale [26], from fish and bird [17, 27-29], from wheat germ [30], yeast [16, 17], and a Pseudomonas species [31].

Cytochrome *c* as derived from horse or beef heart is a strongly basic protein [32] with a molecular weight of around 12,000 [19]. It contains 0.45% iron, the metal being part of a single prosthetic group of the haem type. The low molecular size, the ease of purification, the haemoprotein nature which facilitates observation and physicochemical investigation, the biological importance, and the wide distribution of cytochrome *c*, have all contributed to make this protein especially attractive for structural studies and for investigations of the relationship between structure and catalytic activity.

Great attention has for a long time been paid to the two features most characteristic of cytochrome *c*, namely the covalent bonds which link apoprotein and prosthetic group, and the special architecture of the molecule which is responsible for its haemochrome spectrum. THEORELL has been the first to show convincingly that sulphur bridges are accomplishing the stable conjugation of haem and apoprotein [33]. Each cytochrome *c* molecule contains four sulphur atoms [34]. Two of them are accounted for by methionine residues of the apoprotein [35], whereas the other two are peculiar in remaining attached to the porphyrin moiety, upon acid hydrolysis of the haemoprotein. From acid digests of cytochrome *c* ZEILE & MEYER [36], as well as THEORELL [37], isolated and purified a compound called porphyrin *c*, which may be considered as an adduct of one protoporphyrin and two cysteine molecules (fig. 1). In cytochrome *c* an iron atom occupies the centre of the tetrapyrrole structure of porphyrin *c* and the cysteine residues of the iron-containing « Haem *c* » are incorporated by peptide bonds into the protein moiety. The two thioether bridges of cytochrome *c* which are stable to acid can, however, be broken by treating the haemoprotein with silver or mercuric salts in weakly acidic solution; PAUL has found that the apoprotein devoid of pigment and the free pigment in the form of optically active haematoporphyrin can thereby be obtained separately [38].



Porphyrin c

FIG. 1

In addition to the thioether bridges, there are other bonds linking haem and protein in cytochrome *c*. Functional groups of the protein moiety are coordinated with the iron of the prosthetic group, a fact clearly revealed by the haemochrome spectrum of cytochrome *c*. The nature of these haem-linked groups has raised special interest since they strongly influence the properties of the prosthetic group and doubtless play a crucial role in building up the catalytically active site of the cytochrome *c* molecule. THEORELL & ÅKESON [32] who admirably investigated this point, as early as twenty years ago, using spectrophotometric, titrimetric and magnetic methods, concluded that at least one, and most likely two haem-linked

groups were imidazole rings of histidine residues. It should be mentioned in this context that, on the whole, three histidine residues have been found to occur in horse and beef cytochrome *c* [32, 39, 40].

The facts so far reviewed were the starting point and the basis for structural investigations with which the author of this paper has been closely associated. The first aim, several years ago, was to determine the nature of the amino acid residues adjacent to the cysteines which are engaged in the thioether linkages, with an eye to the possibility that among them there might be histidine residues complexed with the haem iron. Partial hydrolysis of horse and beef cytochromes *c* with diluted sulphuric acid gave a series of iron-free porphyrin-*c*-peptides, with the porphyrin-linked cysteine residues still bound to other amino-acids. A study of these coloured compounds established the occurrence of the two cysteines in the following peptide sequences, Lys-*Cys*-Ala-Glu and Glu-*Cys*-*His*. Thus, one histidine residue was indeed found to be closely attached to porphyrin *c* [41].

When cytochrome *c* was subjected to digestion by either pepsin or trypsin, iron-containing haem-*c*-peptides were obtained. Such red coloured compounds had first been prepared by Tsou [42], but they had not yet been pure enough for a study of amino-acid sequences. In order to arrive at homogeneous preparations, a series of purification steps involving precipitations with ammonium sulphate and trichloroacetic acid, isoelectric agglutination at pH 5, and partition chromatography on columns of Hyflo Supercel had to be carried out.

The structure of the *peptic* degradation product of beef cytochrome *c* is represented in fig. 2 [43]. Its peptide moiety is composed of eleven amino acid residues which are arranged in one single polypeptide chain and contain the two porphyrin-bound cysteine residues. In order to establish this amino acid sequence the haemopeptide was first cleaved into its haematin and peptide moieties with the help of silver salt in acetic acid,

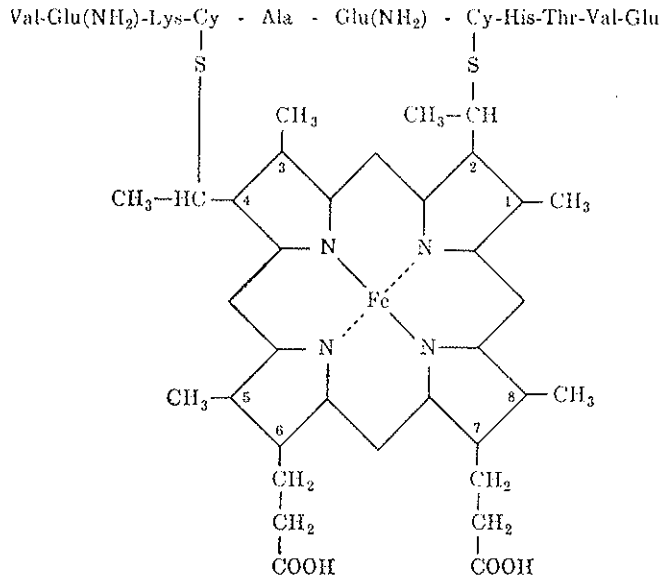


FIG. 2 — Haem-*c*-peptide from beef heart cytochrome *c*.

following the procedure which PAUL had successfully used for the breakage of the thioether bonds in cytochrome *c*. The peptide moiety set free by this treatment was then oxidized with performic acid in order to convert the cysteine residues into cysteic acid residues. The sequence of amino acids in the porphyrin-free undecapeptide thus obtained was elucidated with the help of end-group determinations, and by fragmenting the chain with the enzyme subtilisin (which produced five small peptides, Val-Glu(NH_2), Lys-$CySO_3H$, Ala-Glu(NH_2), $CySO_3H$-His, and Thr-Val-Glu) and with mineral acid.

The haemopeptide prepared by *tryptic* hydrolysis of horse cytochrome *c* [44] contained only nine amino acid residues aligned in one polypeptide chain: Cys-Ala-Glu(NH_2)-Cys-His-Thr-Val-Glu-Lys. N-terminally this nonapeptide sequence was

shorter by three amino acid residues than the corresponding sequence of the peptic haemopeptide, whereas the C-terminus was longer by one (lysine) residue.

These results provided evidence that the two cysteine residues of cytochrome *c*, which are involved in the stable sulphur bridges between protein and haem, occur in one and the same polypeptide chain and are separated from each other by only two other amino acid residues. The detection of an histidine residue in the position adjacent to one of the cysteine residues made it appear likely that its imidazole group was involved in rendering cytochrome *c* a haemochrome.

In several laboratories the properties of the haem-*c*-peptides were compared with those of cytochrome *c*. Only a few of the results obtained can be referred to in this paper. The haemopeptides lack the typical electron-transfer capacity of cytochrome *c* in the succinic oxidase system of the respiratory chain. They are auto-oxidizable, in contrast to cytochrome *c*, and they exhibit a strong oxydatic activity towards substrates such as ascorbic acid [42, 44]. In the presence of hydrogen peroxide they exert a peroxidatic action far superior to that of cytochrome [43]. Their visible absorption spectrum, in the reduced state and at neutral pH, is very similar to the haemochrome spectrum of cytochrome *c*, including the presence of the characteristic alpha-band at 550 mμ, though the extinction coefficients are significantly lower. Addition of imidazole compounds to neutral solutions of the haemopeptides was shown to increase the height of the absorption maxima and to cause them to approach those of cytochrome *c* [39, 44]. Concomitantly, the oxidatic activity of the haemopeptides was suppressed [44]. The oxidation-reduction potential of the haem undecapeptide has been determined by HARBURY & LOACH in the presence of an excess of imidazole, and has been found to be around -200 mV, in the neutral pH region [45]; it thus differs greatly from the corresponding one reported for cyto-

chrome *c* (+250 mV) [46] and resembles more nearly that found for horseradish peroxidase (−280 mV) (47).

EHRENBERG & THEORELL [48], who constructed and discussed steric models of haemopeptides, were able to show that the sulphydryl groups of two cysteine residues in a polypeptide could be readily brought into contact with the two vinyl side chains of a protoporphyrin molecule provided that the cysteine residues were separated from one another by two other amino-acid residues and that the polypeptide chain was turned into an alpha-helix. When the polypeptide chain in the model of the haem undecapeptide had the conformation of a rigid alpha-helix of the left-hand type with its axis parallel to the plane of the haem disk, the imidazole group of the histidine residue adjoining one of the cysteine residues took a steric position utterly favourable for coordination with the haem iron. Since X-ray diffraction studies by ARNDT & RILEX [49] have indeed revealed that the polypeptide chains in the closely packed molecule of cytochrome *c* form left-hand alpha-helices, at least over rather long stretches, it is likely that the imidazole ring of the histidine residue adjacent to cysteine is one of the haem-linked nitrogenous groups responsible for the haemochrome properties of cytochrome *c* as well as of the haemopeptides.

Whereas the attachment of one imidazole nitrogen to the iron does not appear to be controversial for haem peptide solutions at pH values above 4,6, the functional group which occupies the sixth coordination position around the iron has not yet been identified unequivocally. For neutral or slightly alkaline solutions, it might be either the α-amino group of the N-terminal valine [50] or the ε-amino group of the lysine residue [51]. A simultaneous *intra*molecular coordination of the imidazole group of the histidine and of the ε-amino group of the lysine residue, however, is possible only if the polypeptide chain is in an extended configuration [51], but not if it conforms to an α-helix [48]. It has been shown by EHRENBERG & THEORELL and by HARBURY & LOACH that, when the sixth

position around the iron is filled, polymerization of the haem peptide to form pentamers or hexamers is taking place [45, 48]. Therefore, in the peptides investigated haemochrome formation appears to involve an *inter*molecular rather than an *intra*molecular coordination of a sixth ligand with the iron.

In native cytochrome *c* the imidazole ring of a histidine residue belonging to another portion of the polypeptide chain of the haemoprotein is considered to be the most fitting group for coordinating with the iron in position 6. The discussion as to the nature and location of this second iron-complexed group of the protein was greatly stimulated by MARGOLIASH who, in 1955, reported experiments which seemed to indicate that two histidine residues might be N-terminal in the cytochrome *c* molecule and that one of them might be bound to iron [52]. MATSUBARA and others, on the other hand, claimed to have found an arginine residue in N-terminal position [53]. In our own experiments [54] and in those of HOLLEMAN & BISERTE [56] neither of these findings could be confirmed. The reason why these attempts to identify an amino acid residue carrying a free α-amino group in cytochrome *c* gave negative results, became apparent when the peptides present in a tryptic hydrolysate of horse cytochrome *c* were thoroughly investigated [55], in recent studies which aimed at working out the entire structure of the haemoprotein.

According to recent analyses, horse cytochrome *c* contains 2 arginine [56-58] and 18 [57, 58] or 19 [56, 59] lysine residues per molecule. Digestion of the haemoprotein with trypsin, which is known to cleave peptide bonds involving the carboxyl groups of arginine and lysine residues, was therefore expected to yield 21 or 22 fragments. However, since the commercial trypsin used exhibited a slight chymotroptic activity, a greater number of peptides was obtained. In order to minimize chymotryptic activity, the enzyme was pretreated with dilute hydrochloric acid [60] and the duration of the enzymic digestion of cytochrome *c* was reduced to two hours. The peptide mixture

obtained was resolved on sheets of filter paper with the help of a two-dimensional « finger-printing » technique very similar to that introduced by INGRAM [61]. The peptides were located with the help of ninhydrin or by observing the fluorescence which appeared on heating the sheets at 105° [62]. The resulting peptide pattern is given in fig. 3. For the qualitative de-

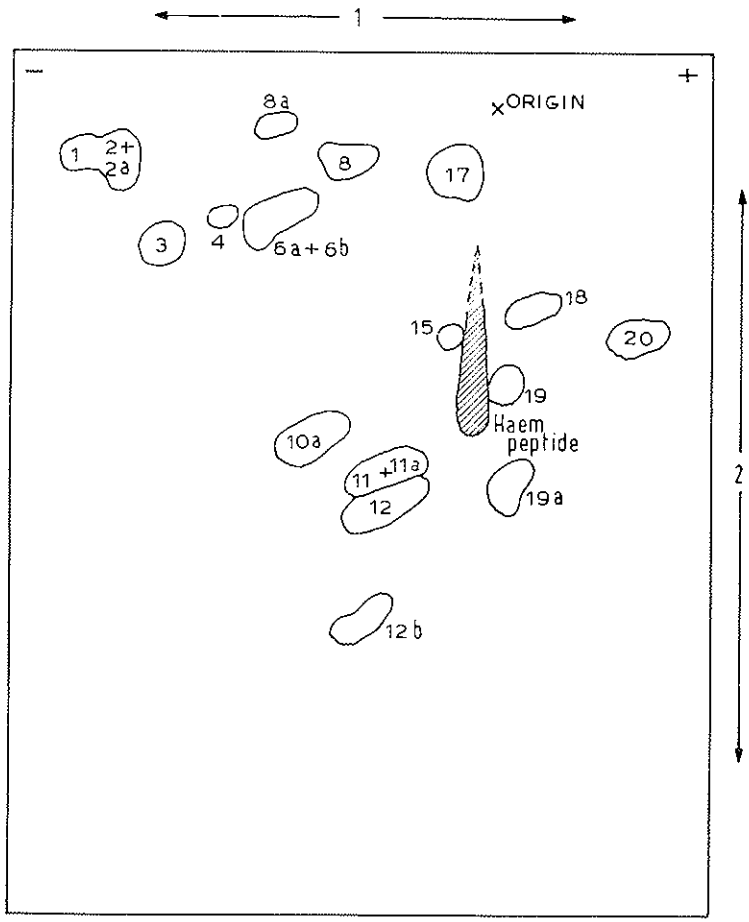


FIG. 3 — 'Finger-print' of horse heart cytochrome *c*.
A tryptic digest of cytochrome *c* resolved (1) by paper electrophoresis in pyridine acetate buffer, pH 4.7, and (2) chromatography with n-butanol, acetic acid, water, 4:1:5, v/v/v.

termination of their amino-acid composition the peptides from one sheet were eluted, hydrolyzed, and the hydrolysates were analysed by paper chromatography. When larger amounts of peptide were required for sequential analysis, the material eluted from several « fingerprints » was combined. Alternatively, a larger-scale resolution of the peptide mixture obtained on tryptic digestion was carried out in two consecutive steps: larger portions of the digest were first subjected to one-dimensional paper electrophoresis so as to give a series of fractions of different mobility; corresponding fractions eluted from several electrophoretic runs were then combined and purified by descending paper chromatography. The methods used to elucidate the amino acid sequence in most of the peptides occuring in the digest were, largely, the classical ones introduced by SANGER, such as end-group determinations and partial hydrolysis by acid and by various enzymes, followed by characterization of the resulting smaller fragments.

In the tryptic digest of horse cytochrome *c* an acidic peptide (*T-20*) was observed which gave a very weak ninhydrin colour and which was devoid of a free α-amino group. Cleavage by pepsin resulted in a neutral tripeptide, Val-Glu-Lys, and an acidic fragment which failed to react with ninhydrin. The latter, on total acid hydrolysis, gave glycine and aspartic acid, whereas partial hydrolysis with 2N HCl (40 min., 100°) yielded the dipeptide Gly-Asp; heating at pH 2 (90 min., 105°) produced free aspartic acid and a substance chromatographically indistinguishable from acetylglycine; hydrazinolysis gave aspartic acid and two hydrazides behaving like glycylhydrazide and acetyl-hydrazide. It has been concluded, therefore, that T-20 is *acetyl-Gly-Asp-Val-Glu-Lys* and represents the N-terminal portion of cytochrome *c*. Thus cytochrome *c* can be added to the growing list of proteins and polypeptides characterized by an acetylated N-terminal amino acid residue, viz. bovine fibrinogen [63], ovalbumin [64], the proteins of tobacco mosaic [65], cucumber 4 [66], and turnip yellow mosaic vi-

rus [67], and porcine α-melanocyte-stimulating hormone [68]. The same conclusion, i.e. that the amino-end of the polypeptide chain of horse cytochrome *c* was N-acetylated, was also arrived at by MARGOLIASH & SMITH [59] and, independently, by TITANI, YAOI, NARITA & OKUNUKI [69]; the latter authors isolated acetyl-Gly-Asp-Val-Glu from a mixture of peptides which was obtained from horse and beef cytochrome *c* upon digestion with a proteinase produced by Streptomyces griseus.

When trypsin virtually free of chymotryptic activity was employed for the digestion of horse cytochrome *c*, all the fragments arising in good yield contained C-terminal lysine or arginine, except for one (*T-18*) whose C-terminal residue, as determined by hydrazinolysis, proved to be glutamic acid. Glutamic acid had previously been reported by TITANI et al. [70] to be in the C-terminal position in horse cytochrome *c*, a finding that we could readily confirm. This obviously suggested that T-18, the structure of which was determined to be *Ala-Thr-AspN-Glu*, was derived from the C-terminus of the polypeptide chain of horse cytochrome *c*.

The results of sequence studies on all the peptides, terminal and internal, released from cytochrome *c* by digestion with trypsin, are summarized in Table 1.

While the tryptic cleavage products of cytochrome *c* were studied in Vienna (KREIL & TUPPY [55], the peptides arising from the haemoprotein upon digestion with chymotrypsin were investigated at Salt Lake City, by E. MARGOLIASH & E. L. SMITH (59). The peptides were separated by ion exchange chromatography, quantitatively analyzed and their amino acid sequence determined. The sequences of chymotryptic peptides listed in Table 2 are taken from the paper of MARGOLIASH & SMITH [59].

From the results obtained by the two groups of investigators a unique sequence for horse heart cytochrome *c* could be derived [71] (fig. 4). This sequence comprises 104 amino-acid residues which, including the prosthetic group, add up to a

TABLE I — Peptides obtained from horse heart cytochrome *c* by digestion with trypsin.

T-20	Acetyl-Gly-Asp-Val-Glu-Lys
T-2	Gly-Lys-Lys
T-12	Ileu-Phe-Val-GluN-Lys
Haem peptide	Cys-Ala-GluN-Cys-His-Thr-Val-Glu-Lys
T-6a	Gly-Gly-Lys
T-2a	His-Lys
T-10a	Thr-Gly-Pro-AspN-Leu-His-Gly-Leu-Phe-Gly-Arg
T-17	Thr-(Ala,GluN,Gly,Pro)-Phe-Thr-Tyr-Thr-(Ala,Asp)-Lys
T-6b	AspN-Lys
T-11a	Gly-Ileu-Thr-Try-Lys
T-19	Glu-(Glu,Leu,Thr)-(Glu,Met)-Tyr-Leu-Glu-AspN-Pro-Lys
T-11	Tyr-Ileu-(Gly,Pro)-Thr-Lys
T-12b	Met-Ileu-Phe-Ala-Gly-Ileu-Lys
T-8a	Lys-Lys-Thr-Glu-Arg
T-19a	Glu-Asp-Leu-Ileu-Ala-Tyr-Leu-Lys
T-18	Ala-Thr-AspN-Glu

TABLE 2 — Peptides obtained from horse heart cytochrome *c* by digestion with chymotrypsin (MARGOLIASH & SMITH) (59).

C-XIXa	Acetyl-(Asp,Gly)-Val-Glu-Lys-Gly-Lys-Lys-Ileu-Phe
C-XXXIIIa (haem peptide)	Val-GluN-Lys-Cys-Ala-GluN-Cys-His-Thr-Val-Glu-Lys-Gly-Gly-Lys-His
C-XXVa	Lys-Thr-Gly-Pro-AspN-(Leu,His)-Gly-Leu-Phe
C-XXVIa	Gly-Arg-Lys-Thr-Gly-GluN-Ala-Pro-Gly-Phe
C-XIIIa	Thr-Tyr
C-XXa	Thr-Asp-Ala-AspN-Lys-AspN-Lys-Gly-Ileu-Thr-Try
C-XVIIIb	Lys-Glu-Glu-Thr-Leu-Met-Glu-Tyr
C-XXIVa	Leu-Glu-AspN-Pro-Lys-Lys-Tyr
C-XVIIIa	Ileu-Pro-Gly-Thr-Lys-Met
C-XIVa	Ileu-Phe
C-XXVIIIa	Ala-Gly-Ileu-Lys-Lys-Lys-Thr-Glu-Arg-Glu-Asp-Leu-(Ileu,Ala)-Tyr
C-IVa	Leu
C-XXIIIa	Lys-Lys-Ala-Thr-AspN-Glu

The arrows indicate the points of hydrolysis by trypsin (T) and chymotrypsin (C).

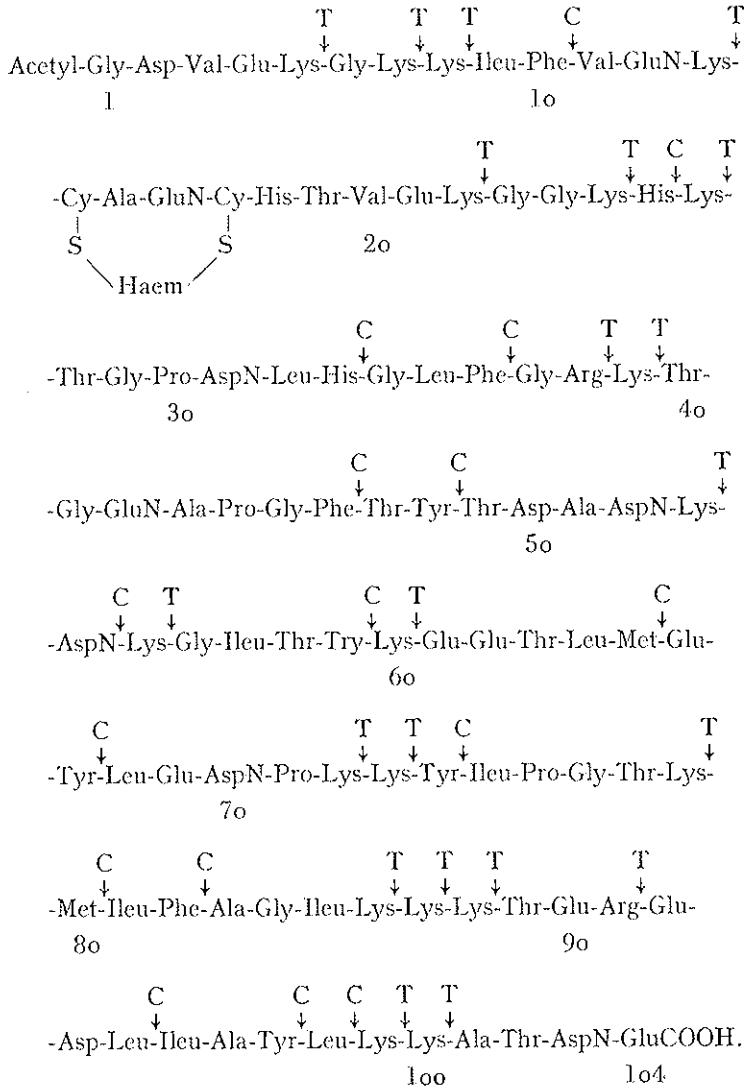


FIG. 4 — The amino acid sequence of horse heart cytochrome *c*.

molecular weight of 12,398. As already pointed out, the N-terminal α-amino group is acetylated. The haem is linked by thioether bonds to the cysteine residues 14 and 17 and is thus located near the N-terminus of the chain. Interestingly, all the three histidine residues occuring in horse cytochrome *c*, and not only the one adjacent to cysteine that was dealt with before, are situated in the N-terminal third of the sequence (residues 18, 26 and 34). It cannot be stated yet whether the imidazole group either of residue 26 or of residue 34 is coordinated with the iron of the prosthetic group, in addition to that of residue 18, but the elucidation of the entire amino acid sequence of horse cytochrome *c* will greatly stimulate and facilitate physical and chemical attempts to settle this question.

A few other points are worth noting. The basic amino acid residues in the polypeptide chain show a tendency to occur in groups (Lys-Gly-Lys-Lys, residues 5-8; Lys-His-Lys, 25-27; Arg-Lys, 38-39; Lys-Lys, 72-73; Lys-Lys-Lys, 86-88; and Lys-Lys, 98-99). A less marked grouping is observed for the non-amidated acidic residues (Glu-Glu 61-62; Glu-Asp, 92-93) and for large hydrophobic residues (Leu-Phe, 35-36; Tyr-Ileu, 74-75; Ileu-Phe, 81-82; Leu-Ileu, 94-95).

Recently the amino acid sequences in two haemoproteins other than cytochrome *c* were established, viz. in the α- and β-chains of human haemoglobin A, by BRAUNITZER et al. [72] and by KONIGSBERG et al. [73], and — partially as yet — in the single chain of sperm-whale myoglobin, by EDMUNDSON & HIRS [74] and by KENDREW et al. [75]. Marked similarities were found to exist between the two haemoglobin chains [76] as well as between haemoglobin and myoglobin [77] which were interpreted as homologies pointing to a common extraction of the genes responsible for their synthesis [78]. Horse cytochrome *c*, however, does not show any obvious relationship in its amino-acid sequence to either haemoglobin or myoglobin. This lack of similarity is not surprising, however, in

view of the different functions of the electron-transferring cytochrome and the oxygen-carrying haemoproteins.

Cytochromes of type *C* appear to be an extraordinarily good object for an extensive study of species differences of protein structure since they have been found and can be investigated in organisms belonging to all different levels of the phylogenetic scale. Provided that they are homologous in the genetic sense, the study of their amino acid sequences should indicate the nature and scope of the variations which the structure of a protein may incur in the course of the (hypothetical) evolutionary process. Certain features of the structure, naturally, should always remain constant, features which are so intimately linked with the specific structure and function of the cytochrome *c* molecule that they cannot be changed without loss of its biological activity.

So far, the complete sequence of amino acid residues has not yet been established for any cytochrome other than that of horse heart. A comparison, however, has been carried out of the arrangement of amino acids in haemopeptides prepared by peptic and/or tryptic digestion of *C*-type cytochromes of three mammals [79], a bird [80], a fish [80], an insect [81] and three microorganisms [15, 82, 83].

It is obvious from the results summarized in Table 3 that a certain pattern of amino acid residues has indeed been found to occur in all *C*-type cytochromes studied. This pattern includes two porphyrin-bound cysteine residues separated from each other by two other residues, and a histidine residue adjacent to one of the cysteines. Their invariable presence in the same sort of arrangement confirms the contention put forward before, that they are involved in building up the catalytically acitve site of the cytochrome *c* molecule. A threonine residue following upon histidine and a basic residue, either lysine or arginine, preceding one of the cysteine residues, are also appearing regularly; we do not know yet whether or not this constant occurrence is required by cytochrome *c* function or by the

TABLE 3 — Amino-acid sequences in haem-*c* peptides obtained from cytochromes *c* of different origin.

Beef Horse Pig	...-Val-Glu(NH$_2$)-Lys-Cys-Ala-Glu(NH$_2$)-Cys-His-Thr-Val-Glu-Lys-...
Salmon	...-Val-Glu(NH$_2$)-Lys-Cys-Ala-Glu(NH$_2$)-Cys-His-Thr-Val-Glu-...
Chicken	...-Val-Glu(NH$_2$)-Lys-Cys-Ser-Glu(NH$_2$)-Cys-His-Thr-Val-Glu-...
Silkworm	...-Val-Glu(NH$_2$)-Arg-Cys-Ala-Glu(NH$_2$)-Cys-His-Thr-Val-Glu-...
Yeast	...-Phe-Lys-Thr-Arg-Cys-Glu-Leu-Cys-His-Thr-Val-Glu-...
Rh. Rubrum	...- Cys-Leu-Ala-Cys-His-Thr-Phe-Asp.Glu-Gly-Ala-Asp(NH$_2$)-Lys-...

stability of molecular architecture or whether it is due to the fact that these residues simply happened not to be changed in the course of evolution.

In horse, beef and pig and even in salmon cytochrome *c* the sequence of amino acids in the neighbourhood of the prosthetic group proved to be exactly the same. (This does not mean, however, that the structures of mammalian and fish cytochromes are wholly identical. An electrophoretic difference between horse and beef cytochrome *c* has been noted by several authors [84-86] and a recent comparison by the fingerprinting technique [87] has confirmed the existence of minor structural variations.) In chicken cytochrome *c* a serine residue was found to be present where alanine was occurring in mammalian and fish cytochromes. Silkworm cytochrome *c* differs from the others in possessing an arginine residue instead of a lysine residue. More marked deviations were observed in a yeast cytochrome *c* and even more pronouncedly in cytochrome c_2 of *Rhodospirillum rubrum*. All the cytochromes studied here, with the exception of the last-mentioned Rhodospirillum cytochrome c_2, have virtually the same catalytic activities and are interchangeable in their respective systems. Thus, e.g., yeast cytochrome *c* is active in the mammalian succinic oxidase and cytochrome oxidase systems, and mammalian cytochromes *c* are active in the yeast systems. C_2 of *Rh. rubrum*, however, is somewhat different. It possesses many but not all the properties usually attributed to mammalian cytochrome *c*. As shown by KAMEN, ELSDEN and VERNON (88-90) its absorption spectrum closely resembles that of mammalian and yeast cytochromes *c*; it is not auto-oxidizable, and it can be reduced with DPNH *via* DPNH-cytochrome *c* reductase of pig heart, like mammalian cytochrome *c*. If differs from the latter in that it is not oxidized by oxygen in the presence of the mammalian cytochrome oxidase system; it has an oxidation-reduction potential higher than that of mammalian cytochrome *c*. The amino acid sequence in the tryptic haemopeptide obtained from

Rhodospirillum rubrum cytochrome c_2 still contains the characteristic pattern involving two cysteines separated from each other by two other residues, and an adjacent histidine residue, whereas most of the other residues differ from those found in the cytochromes of higher organisms.

Quite recently, DUS, BARTSCH & KAMEN [83] reported the amino-acid sequence of an unusual peptide containing 2 haem groups and 27 amino acid residues (including 3 cysteines and two histidines) isolated from an haemoprotein of the obligate photoanaerobe, Chromatium: Phe-Ala-Gly-Lys-Cys-Ser-Glu-Cys-His-Thr-Leu-Val-Ala-Asp-Glu-Gly-Ser-Ala-Lys-Cys-His-Thr-Phe-Asp-Glu-Gly-Ser. One haem group appeared to be attached to the peptide chain in the usual manner through thioether linkages involving two cysteine residues (adequately spaced by two intercalated amino acid residues), one of the cysteines being followed in the sequence by histidine in full agreement with the pattern found in other *C*-type cytochromes. For the covalent attachment of the second haem, only one cysteine was available which, however, likewise had histidine for its neighbour. Moreover, the heptapeptide sequence, -Cys-His-Thr-Phe-Asp-Glu-Gly-, has proved common to this haem peptide and to that derived from Rh. rubrum cytochrome *c*, a factor pointing to a genetical relationship between the parent haemoproteins.

If we disregard the cytochromes of microorganisms and only compare the amino acid sequences in haemopeptides derived from animal cytochromes, we notice that they are distinguished by the substitution of single amino acid residues for others (Ala/Ser, Lys/Arg). Analogous species differences have been observed for insulin [91-95], hypertensins [96, 97], melanophore-stimulating hormones [98-101], etc. These small variations are similar in nature and in scope to those which have been found by INGRAM and others [102] in « pathological » types of human haemoglobin and by WITTMANN [103] and FRAENKEL-CONRAT [104, 105] in the protein moiety of tobacco

mosaic virus strains that had been obtained from the wild type virus by treatment with mutagen. It would thus appear that changes in protein structure appearing *inter*specifically and *intra*specifically are of the same kind and are similarly due to point mutations in the genetic determinants of protein structure . If that be so, the interspecific variation of protein structure might be taken as an indication of gene evolution and, more generally, of phylogenetic evolution. The consideration of the molecular morphology of homologous proteins should prove to be a valuable supplement to the comparative studies of macromorphology upon which phylogenetic theories are mostly based. In such a project, naturally, the lack of fossil protein molecules will be a serious handicap. This drawback will be counterbalanced, however, by the wealth of information residing in the amino acid sequences of the proteins in the existing species of organisms.

The comparative study of *C*-type cytochromes has as yet involved too short a portion of their amino acid sequence and too few different species as to warrant any phylogenetic conclusions. The approach, however, appears to be promising and should eventually prove fruitful.

REFERENCES

[1] Report of the Commission on Enzymes of the International Union of Biochemistry. Pergamon Press, 1961 .

[2] C.A. Mac Munn, *Phil. Trans. Roy. Soc. London*, *176*, 661 (1885); *177*, 267 (1886); *J. Physiol.*, *8*, 57 (1887); *Z. physiol. Chem.*, *13*, 497 (1889).

[3] D. Keilin, *Proc. Roy. Soc. London*, *98B*, 312 (1925); *100B*, 129 (1926).

[4] E.F. Hartree, *Advances in Enzymol.*, *18* (1957).

[5] M.D. Kamen, *Bacteriol. Revs.*, *19*, 250 (1955).

[6] D. Keilin & E.F. Hartree, *Nature*, *164*, 254 (1949); ibid., *176*, 200 (1955).

[7] E. Yakushiji & K. Okunuki, *Proc. Imp. Acad.* Tokyo, *16*, 299 (1940).

[8] D.E. Green, J. Järnefelt & H.D. Tisdale, *Biochim. et Biophys. Acta*, *31*, 34 (1959).

[9] I. Sekuzu, Y. Orii, K. Okunuki, *J. Biochem.* Japan, *48*, 214 (1960).

[10] R. Bomstein, R. Goldberger & H. Tisdale, *Biochim. et Biophys. Acta*, *50*, 527 (1961).

[11] H.E. Davenport, R. Hill, *Proc. Roy. Soc.*, *139B*, 327 (1952).

[12] J. Keilin, *Nature*, *180*, 427 (1957); *Biochem. J.*, *64*, 663 (1956).

[13] L. Tosi & F. Ghiretti, *Arch. Biochem. Biophys.*, *93*, 399 (1961).

[14] D. Keilin, *Proc. Roy. Soc.* London, *106B*, 418 (1930).

[15] H. Tuppy & K. Dus, *Monatsh. Chem.*, *89*, 407 (1958).

[16] B. Hagihara, T. Horio, J. Yamashita, M. Nozaki & K. Okunuki, *Nature*, *178*, 629 (1956).

[17] B. Hagihara, M. Yoneda, K. Tagawa, I. Morikawa & K. Okunuki, *J. Biochem. Japan*, *45*, 565 (1958).

[18] D. Keilin & E.F. Hartree, *Proc. Roy. Soc. London*, *122B*, 298 (1937).

[19] S. Paléus & J.B. Neilands, *Acta Chem. Scand.*, 4, 1024 (1950).

[20] E. Margoliash, *Biochem. Preparations*, 5, 33, 37 (1955).

[21] B. Hagihara, K. Tagawa, I. Morikawa, M. Shin & K. Okunuki, *Proc. Japan. Acad.*, *34*, 169 (1958).

[22] G. Bodo, *Nature*, *176*, 829 (1955).

[25] *Tuppy* - pag. 22

[23] S.A. KUBY, S. PALÉUS, K.-G. PAUL, H. THEORELL, *Acta Chem. Scand.*, *10*, 48 (1956).

[24] B. HAGIHARA, I. MORIKAWA, I. SEKUZU & K. OKUNUKI, *J. Biochem. Japan*, *45*, 551 (1958).

[25] B. HAGIHARA, I. MORIKAWA, I. SEKUZU, T. HORIO & K. OKUNUKI, *Nature*, *178*, 630 (1956).

[26] S. MINAKAMI, *J. Biochem. Japan*, *45*, 547 (1958).

[27] B. HAGIHARA, K. TAGAWA, M. NOZAKI, I. MORIKAWA, J. YAMASHITA & K. OKUNUKI, *Nature*, *179*, 249 (1957).

[28] B. HAGIHARA, I. SEKUZU, K. TAGAWA, M. YONEDA & K. OKUNUKI, *Nature*, *181*, 1588 (1958).

[29] B. HAGIHARA, K. TAGAWA, I. MORIKAWA, M. SHIN & K. OKUNUKI, *J. Biochem. Japan*, *45*, 725 (1958).

[30] B. HAGIHARA, K. TAGAWA, I. MORIKAWA, M. SHIN & K. OKUNUKI, *J. Biochem. Japan*, *46*, 321 (1959); *Nature*, *181*, 1589 (1958).

[31] T. HORIO, *Biochem. J.*, *77*, 194 (1960).

[32] H. THEORELL & A. AKESON, *J. Amer. Chem. Soc.*, *63*, 1804, 1812, 1818, 1820 (1941).

[33] H. THEORELL, *Enzymologia*, *4*, 192 (1937).

[34] S. PALÉUS, *Acta Chem. Scand.*, *9*, 335 (1955).

[35] Å. ÅKESON, *Acta Physiol. Scand.*, *4*, 362 (1942).

[36] K. ZEILE & H. MEYER, *Z. physiol. Chem.*, *262*, 178 (1939).

[37] H. THEORELL, *Biochem. Z.*, *298*, 242 (1938); *Enzymologia*, *6*, 88 (1939).

[38] K.-G. PAUL, *Acta Chem. Scand.*, *4*, 239 (1950); *5*, 389 (1951).

[39] A. EHRENBERG & H. THEORELL, *Acta Chem. Scand.*, *9*, 1193 (1955).

[40] R. NUNNIKHOVEN, *Biochim. et Biophys. Acta*, *28*, 108 (1958).

[41] H. TUPPY & G. BODO, *Monatsh. Chem.*, *85*, 807 (1954).

[42] C.L. TSOU, *Biochem. J.*, *49*, 362, 367 (1951).

[43] H. TUPPY & S. PALÉUS, *Acta Chem. Scand.*, *9*, 353 (1955).

[44] H. TUPPY & G. BODO, *Monatsh. Chem.*, *85*, 1024 (1954).

[45] H.A. HARBURY & P.A. LOACH, *Proc. Nat. Acad. Sci. U.S.*, *45*, 1344 (1959).

[46] F.L. RODKEY, E.G. BALL, *J. Biol. Chem.*, *182*, 17 (1950).

[47] H.A. HARBURY, *J. Biol. Chem.*, *225*, 1009 (1957).

[48] A. EHRENBERG, & H. THEORELL, *Acta Chem. Scand.*, *9*, 1193 (1955).

[49] U.W. ARNDT & D.P. RILEY, *Proc. Roy. Soc. London*, *114B*, 93 (1953); *Phil. Trans. Roy. Soc.* (London), *A247*, 409 (1955).

[50] S. PALÉUS, A. EHRENBERG & H. TUPPY, *Acta Chem. Scand.*, *9*, 365 (1955).

[51] E. MARGOLIASH, N. FROHWIRT & E. WIENER, *Biochem. J.*, *71*, 559 (1959).

[52] E. MARGOLIASH, *Nature*, *175*, 293 (1955).

[25] *Tuppy* - pag. 23

[53] H. MATSUBARA, B. HAGIHARA, T. HORIO & K. OKUNUKI, *Nature, 179,* 251 (1957).
[54] S. PALÉUS & H. TUPPY, unpublished.
[55] G. KREIL & H. TUPPY, *Nature, 192,* 1123 (1961).
[56] J.W. HOLLEMAN & G. BISERTE, *Bull. Soc. chim. biol., 41,* 975 (1959).
[57] R. NUNNIKHOVEN, *Biochim. et Biophys. Acta, 28,* 108 (1958).
[58] G. LEAF, N.E. GILLIES & R. PIRRIE, *Biochem. J., 69,* 605 (1958).
[59] E. MARGOLIASH & E.L. SMITH, *Nature, 192,* 1121 (1961).
[60] R.P. REDFIELD & C.B. ANFINSEN, *J. Biol. Chem., 221,* 385 (1956).
[61] V.M. INGRAM, *Biochim. et Biophys. Acta, 28,* 539 (1958).
[62] D.M.P. PHILLIPS, *Nature, 161,* 53 (1948).
[63] J.E. FOLK & J.A. GLADNER, *Biochim. et Biophys. Acta, 44,* 383 (1960).
[64] R.D. MARSHALL & A. NEUBERGER, *Biochem. J., 78,* 31P (1961).
[65] K. NARITA, *Biochim. et Biophys. Acta, 28,* 185 (1958).
[66] K. NARITA, *Biochim. et Biophys. Acta, 31,* 372 (1959).
[67] J.I. HARRIS & A. HINDLEY, *J. Mol. Biol., 3,* 117 (1961).
[68] J.I. HARRIS & P. ROSS, *Biochem. J., 17,* 434 (1959).
[69] K. TITANI, Y. YAOI, K. NARITA & K. OKUNUKI, in press.
[70] K. TITANI, H. ISHIKURA & S. MINAKAMI, *J. Biochem.* (Tokyo), *44,* 499 (1957).
[71] E. MARGOLIASH, E.L. SMITH, G. KREIL & H. TUPPY, *Nature, 192,* 1125 (1961).
[72] G. BRAUNITZER, R. GEHRING-MÜLLER, N. HILSCHMANN, K. HILSE, G. HOBOM, V. RUDLOFF & B. WITTMANN-LIEBOLD, *Z. physiol. Chem., 325,* 283 (1961).
[73] W. KONIGSBERG, G. GUIDOTTI & R.J. HILL, *J. Biol. Chem.*, 1961, *236,* PC 55.
[74] A.B. EDMUNDSON & C.H.W. HIRS, *Nature, 190,* 663 (1961).
[75] J.C. KENDREW, H.C. WATSON, B.E. STRANDBERG, R.E. DICKERSON, D.C. PHILLIPS & V.C. SHORE, *Nature, 190,* 666 (1961).
[76] G. BRAUNITZER, N. HILSCHMANN & B. WITTMANN-LIEBOLD, *Z. physiol. Chem., 325,* 94 (1961).
[77] H.C. WATSON & J.C. KENDREW, *Nature, 190,* 670 (1961).
[78] V.M. INGRAM, *Nature, 189,* 704 (1961).
[79] H. TUPPY & G. BODO, *Monatsh. Chem., 85,* 1182 (1954).
[80] H. TUPPY & S. PALÉUS, *Acta Chem. Scand., 9,* 353 (1955).
[81] H. TUPPY, *Z. Naturforsch., 12b,* 784 (1957).
[82] S. PALÉUS & H. TUPPY, *Acta Chem. Scand., 13,* 641 (1959).
[83] K. DUS, R.G. BARTSCH & M.D. KAMEN, *J. Biol. Chem., 236,* PC 47 (1961).
[84] H. TINT & W. REISS, *J. Biol. Chem., 182,* 385 (1950).
[85] R.W. HENDERSON & W.A. RAWLINSON, *Biochem. J., 62,* 21 (1956).

[25] *Tuppy* - pag. 24

[86] R.P. FARRAS et al., *Chem. Abstr., 52,* 15634e (1958).
[87] G. KREIL, unpublished.
[88] S.R. ELSDEN, M.D. KAMEN & L.P. VERNON, *J. Amer. Chem. Soc., 75,* 6347 (1953).
[89] L.P. VERNON & M.D. KAMEN, *J. Biol. Chem., 211,* 643 (1954).
[90] M.D. KAMEN & L.P. VERNON, *Biochim. et Biophys. Acta, 17,* 10 (1955).
[91] J.I. HARRIS, F. SANGER & M.A. NAUGHTON, *Arch. Biochem., 65,* 427 (1956).
[92] Y. ISHIHARA, T. SAITO, Y. ITO & M. FUJINO, *Nature, 181,* 1468 (1958).
[93] D.S.H.W. NICOL & L.F. SMITH, *Nature, 187,* 483 (1960).
[94] L.F. SMITH, cf. F. SANGER, *Brit. Med. Bull., 16,* 183 (1960).
[95] A. KOTAKI, *J. Biochem.* (Japan), *50,* 256 (1961).
[96] L.T. SKEGGS, K.E. LENTZ, J.R. KAHN, N.P. SHUMWAY & K.R. WOODS, *J. Exp. Med., 104,* 193 (1956).
[97] D.F. ELLIOTT & W.S. PEART, *Biochem. J., 65,* 246 (1957).
[98] I.I. GESCHWIND, C.H. LI & L. BARNAFI, *J. Amer. Chem. Soc., 78,* 4494 (1956); ibid., *79,* 1003 (1957).
[99] J.I. HARRIS & P. ROOS, *Nature, 178,* 90 (1956).
[100] C.H. LI, *Adv. Protein Chem., 12,* 269 (1957).
[101] T.H. LEE, A.B. LERNER & V. BUETTNER-JANUSCH, *J. Biol. Chem., 236,* 1390 (1961).
[102] V.M. INGRAM, « Hemoglobin and its Abnormalities », C.C. Thomas, Springfield, Ill., 1961.
[103] H.G. WITTMANN, *Virology, 12,* 609 (1960); Proc. V. Internat. Congr. Biochem. Moscow. 1961. Symp. I.
[104] A. TSUGITA & H. FRAENKEL-CONRAT, Proc. Nat. Acad. Sci. (Wash.), *46,* 1463 (1960).
[105] H. FRAENKEL-CONRAT & A. TSUGITA, Proc. V. Internat. Congr. Biochem. Moscow. 1961. Symp. I.

DISCUSSION

THEORELL

A similar thing to what we discussed concerning the hemoglobin could be discussed in the case of cytochrome *c*, too. It has a peculiar property of being especially adapted not to react with oxygen at all.

It would be high time to produce hig crystals of cytochrome *c* in order to study its tertiary structure. We speculated years ago about what would be the reason for the inability of cytochrome *c* to react with oxygen at neutral reaction. If you go down to low pH's below 3 or 4 or above 12, then the structure is opened in such a way that iron becomes autoxidizable. An obvious reason could be that the heme iron is so well hidden in the interior of the molecule that it is not accessible for oxygen molecules. Then one may wonder how the electrons come to the iron atom. There are different possibilities. The iron is linked octahedrically to porphyrin and two histidine imidazoles and it appears plausible in this case that the heme-linked imidazols could render some service in forming a conjugated double bond system with the iron and thus lead electrons from the surface of the cytochrome molecule to the iron hidden in its interior.

FRAENKEL-CONRAT

It is not very clear to me whether the sequence of amino acids in each of the peptides has been worked out by you, Dr. TUPPY, for the tryptic peptides and whether it is in agreement with the work reported for the chymotryptic peptides, wherever these overlap.

TUPPY

The sequences of amino acids in all of the tryptic peptides of horse cytochrome *c* have been established in Vienna, with the exception of a few amino acid residues the order of which has not yet been worked out unambiguously. Those sites have been indicated in Table 1 by setting in parentheses the residues concerned. Similarly, MARGOLIASH and SMITH established the structure of the chymotryptic peptides, with the exception of three dipeptide arrangements, as shown in Table 2. The results obtained in the two laboratories are in full agreement whereever the amino acid sequences of tryptic and chymotryptic peptides overlap. Where uncertainties (indicated by parentheses) had been left in one laboratory, they were resolved by the results of the other laboratory, and *vice versa*. Thus, the entire amino acid sequence of horse cytochrome *c* could be derived from the two sets of data.

CHARGAFF

Is anything known about the aminoacid sequences in the rest of the molecule of the different cytochromes *c* for which you showed a comparison of what one might call the active centers that seem to be quite similar? Is the rest of the sequence very different in different species or what are the similarities?

TUPPY

According to what Dr. KREIL in my laboratory has so far been able to find out, the differences between the amino acid sequences between horse and beef cytochromes *c* are limited to the replacement of single amino acid residues at not more than two or three points along the chain. However, the residue exchanges by which horse and tunny fish cytochromes *c* appear to differ are as many as twenty, approximately.

ANFINSEN

I would like to mention that with some proteins, for example the lysozymes of egg whites, there is a very significant change over

a very small evolutionary « distance ». For example, we have examined the egg white lysozymes of chickens, turkeys and pheasants. One cannot make viable crosses between chickens and pheasants, but one can make such crosses between turkeys and pheasants. If one examines the sequential structures of lysozymes from these three species, one finds a considerable number of differences: 5 or 6 amino acid replacements for the chicken-pheasant case and 2 or 3 in the pheasant-turkey example.

TUPPY

As far as cytochrome *c* is concerned, the « evolutionary » change of protein structure appears to have occurred relatively slowly.

CHANTRENNE

You showed that there was always a threonine close to your histidine. Is there anything similar in myoglobin or in hemoglobin?

TUPPY

The histidine residues in cytochrome *c*, on the one hand, and in haemoglobin and myoglobin, on the other hand, have been found to occur in quite different amino-acid sequences.

PERUTZ

In connection with some of Professor THEORELL's remark, I emphasize that we have some nice big crystals of tuna heart cytochrome *c* which give good X-ray diffraction pictures, but we are stuck over the heavy atom problem. We cannot get isomorphous heavy compounds of cytochrome *c*, and if any of those present can help us, we should be delighted.

AMINOACYL RIBONUCLEIC ACID AND INTERMEDIARY REACTIONS IN PROTEIN SYNTHESIS

PAUL C. ZAMECNIK

J.C. Warren Laboratories of Harward University, Mass. General Hospital, Boston, Mass. (U.S.A.)

A growing body of evidence indicates that the genetic message inherent in the sequence of purine and pyrimidine bases in deoxyribonucleic acid (DNA) is transformed into a sequence of bases in ribonucleic acid (RNA), and that the latter serve to order the sequence of amino acids which polymerize to form a particular protein molecule. If this reasoning is correct, at some point in the translation of the sequence of bases in DNA to the sequence of amino acids in the protein molecule, there must be a key molecule in which an amino acid becomes coded to an oligonucleotide or polynucleotide. Independently of CRICK's [1] prediction that such would be the situation, we came upon such a molecular type several years ago, and it became known as soluble RNA (s-RNA) because of its relatively small molecular weight [2], and more recently, as transfer RNA [3] in recognition of its function. Following activation of an amino acid by interaction with an enzyme and adenosinetriphosphate (ATP) to form a bound aminoacyl-AMP [4], a transfer of this compound to ester linkage on the terminal adenosyl unit of an s-RNA molecule occurs [5-8], a reaction catalyzed by the same enzyme as that which activates the amino acid.

Activating enzymes, of which there is very likely one for each amino acid, are therefore unusual catalysts, deserving of more attention than they currently receive. One must consider the presence of three separate recognition sites on a single activating enzyme, as follows: one for the amino acid side chain (since there is great specificity for a particular amino acid [cf. 9]); one site for the ATP (since GTP, CTP, and UTP will not serve as substitutes); and one site for the particular s-RNA molecule to which the activated amino acid is transferred. With regard to the last-mentioned point, one would postulate either a large scale fit of a portion of the enzyme surface and the s-RNA surface; or alternatively that the activating enzyme has, built into its structure, an oligonucleotide coding piece which is complementary to a section of the s-RNA chain. Either one of these possibilities makes the structure of the amino acid activating enzymes an inviting puzzle.

In any event, once the amino acid becomes esterified to the hydroxyl group (whether 3′- or 2′-hydroxyl is unsettled) of the ribosyl moiety of the terminal adenosyl unit of the s-RNA, it would appear that the task of placing the amino acid in proper position on the ribosomal template prior to the polymerization step has been transferred to the s-RNA. This positioning could be accomplished by appropriate selective base pairing of a sequence of 3 or more mononucleotide residues on the s-RNA molecule with complementary bases in the ribonucleic acid chain located on the surface of the ribosomal particle [10].

From the foregoing considerations, it has become of great interest to try to isolate a single species of s-RNA, coded to a particular amino acid. Difficult though it was to accept the point of view, in addition to a separate activating enzyme for each amino acid, there would be a separate s-RNA molecule as well for each amino acid, the evidence has become too compelling to consider otherwise [7, 8, cf. also 11], and we have become engaged in separation techniques.

As mentioned in detail elsewhere [11], both physical and chemical methods have been employed by a number of laboratories to achieve this goal. Our present separation procedure is a combination of two methods based on different principles. By mild alkaline hydrolysis we first remove all aminoacyl groups from their esterification sites on the terminal adenosyl units of the s-RNA [12]. A single, labeled amino acid (i.e. C^{14}-l-valine) is then reesterified on the s-RNA by incubation with activating enzyme and ATP. Sodium periodate is now added, and the cis-hydroxyl groups of the ribosyl moieties of the terminal adenosyl residues of the s-RNA molecules are oxidized to the dialdehydes [12]. The species of s-RNA which has the aminoacyl ester present on either the 2′ or 3′ ribosyl carbon is, however, resistant to periodate oxidation, and this is a key feature of the separation method. A hydrazone is formed by addition of 2-hydroxy, 3-naphthoic acid hydrazide to the above-mentioned aldehyde groups, and a coupling agent is then added (tetra-azotized diorthoanisidine). This reagent couples specifically with the naphthoic acid hydrazono-RNA, forming a bulky, deep blue dye, attached covalently to the terminal adenosyl residue of the RNA. This dye-bearing RNA has a decreased solubility in alcohols and water, and may be separated from the aminoacyl-RNA by precipitation methods [12]. Unfortunately, this procedure endeavors to remove the 97-98%, of RNA not bearing a valine ester from the 2-3% which does. To the extent that the dye attachment falls short of 100% completion, the method will fail to achieve separation of the single amino-acyl-RNA species. This has proven to be the case, even with later improvements in the general technique [13, 14].

We therefore looked for a different fractionation principle to employ subsequent to the dye addition step, and explored molecular sieving [15] and column chromatography. Although cross-linked dextran [16] (Sephadex, Pharmacia, Uppsala, Sweden), and diethylaminoethyl (DEAE)-cellulose and starch

columns were tried without notable success, we found that cross-linked DEAE-dextrans gave promising results, with approximately 25 fold enrichment of specific activity of valyl-RNA. Crucial to the success of this adsorption-elution procedure were 1) the length of the column, 2) the proper sodium chloride elution gradient, and 3) the elution rate, and details of these conditions have recently been published [11]. A purification of valyl-RNA of approximately 65% was thereby obtained. More recently, in the last five columns, we have obtained pooled peak tube purifications of the order of 95% for valyl-RNA, by means of the following changes in conditions: 1) increasing the height of the column from 55 to 95 cms. (1.2 cm. diameter in each case), 2) changing the salt gradient from the 0.35-0.40 M rise in 2000 ml. of elution fluid to 0.32-0.38 M NaCl gradient change in the same volume, and by decreasing the rate of flow through the columns from 100 ml. per hour to 40 ml. per hour [17]. This purification figure is based on the assumption of a molecular weight of 25,500, one molecule of valine esterified to one molecule of RNA, and an A_{260} absorbancy index for s-RNA of 21.4 cm_2mg_{-1} (as determined by Dr. JESSE F. SCOTT). Thus the valyl-RNA species approaches a degree of purity which makes sequential degradation studies worthy of consideration.

It is interesting to speculate on how the DEAE-dextran column operates in separating valyl-RNA from the other two general types of s-RNA present in the s-RNA placed on the column. The dye-RNA, accounting for some 80± % of the bulk of this RNA, becomes firmly bound to the dextran at the top of the column, imparting to it a deep blue color. The dye-RNA migrates very little down the column, even with high (1.0 M) concentrations of NaCl. The valyl-RNA is the fastest moving major RNA component, and emerges first when the NaCl concentration reaches a level sufficient to displace the weak linkages which bind to the column the valyl-RNA and the RNA fraction which has neither valine nor dye attached.

This latter RNA fraction amounts to around 17-18% of the total, with the valyl-RNA accounting for 2-3% of the original s-RNA.

It is well established that s-RNA has a secondary structure, and is largely intramolecularly hydrogen bonded [18]. It is therefore reasonable to regard exposed hydroxyl groups, such as those on the ribosyl moiety of the terminal adenosyl unit, as being capable of contributing to a greater extent to hydrogen bonding with the nitrogen or oxygen atoms of the DEAE-dextran column than do internally situated groups in the RNA molecule. When either the 2′ or 3′-hydroxyl group of this ribosyl unit is converted to an aliphatic aminoacyl ester, there may thus be a decrease in the hydrogen bonding, initially inherent in the presence of these two hydroxyl groups, of sufficient magnitude to account for the slightly more rapid passage of the valyl-RNA through the column. This suggestion is in line with the more rapid flow of methylated oligosaccharides through chromatographic columns and paper in aqueous solvents than unsubstituted oligosaccharides [19, 20]. This type of reasoning would imply that aminoacyl-RNA's bearing good hydrogen bonding constituents on the side chain of the amionacyl residue would also be retarded in their migration through the column. This is the case with tyrosyl-RNA, which does not emerge first, but rather in the middle of that 17-18% of the total s-RNA which has neither aminoacyl ester nor dye attached to it.

The use of this DEAE-dextran column in s-RNA fractionation may thus be pointed out as an example of the use to which the techniques mentioned by Professor TISELIUS may be put.

At this meeting, Professor LIPMANN has mentioned experiments which suggest that the addition of puromycin to a cell-free reticulocyte system results in the liberation of peptides from the ribosomal surface. Dr. DAVID ALLEN in our laboratories has carried out recent experiments which bear on this point [21]). He has synthesized C^{14}-puromycin, labeled in the methoxyl group of the p-methoxyphenylalanine portion of pu-

romycin, by treating the tyrosyl analogue of puromycin with C^{14}-diazomethane. When reticulocyte ribosomes are incubated at 37°C with the complete cell-free system of SCHWEET et al. [22] — i.e. with ATP, GTP, phosphoenol-pyruvate, pyruvate kinase, a soluble enzyme fraction, an amino acid mixture, and C^{14}-l-valine, — C^{14}-labeled protein is released linearly into the soluble protein fraction of the system over a 15 minute period of time. This protein closely resembles or may be identical with a hemoglobin peptide chain. It has an N-terminal valine residue, approximately 10 or 11 internal valine residues, and chromatographic behavior indistinguishable from that of authentic rabbit hemoglobin peptide.

When puromycin is added to this cell-free peptide-synthesizing system, peptide is released rapidly into the soluble protein fraction during the ensuing 1-2 minutes, by which time maximal release has occurred. The peptide release is independent of the addition of the energy generating system mentioned above. The peptides released are largely if not completely valyl N-terminal. They are not hemoglobin peptides, but appear to be partially synthesized chains, originally destined to become hemoglobin peptide, which have started at the N-terminal valine end and have been terminated short of completion by the attachment of a puromycin molecule. The puromycin can only be separated from this peptide by strong acid hydrolysis, with release of amino acids. Prior to acid hydrolysis, the amino group of the p-methoxyplenylalanyl residue of the puromycin is not available for reaction with dinitrofluorobenzene, but is so available following acid hydrolysis. It can be calculated that roughly one puromycin residue is bound to the released peptide per one N-terminal valine residue. Thus the evidence suggests that puromycin acts as a « chainender » for the hemoglobin peptide chain undergoing synthesis on the ribosomal surface.

As suggested by YARMOLINSKY and DE LA HABA [23], puromycin appears to behave as an analogue of an aminoacyl-RNA molecule. The present experiments point to the further

mechanistic detail that the puromycin actually becomes bound in peptide linkage as the last unit in the peptide chain, which then separates from the ribosomal surface, due to the absence of the hydrogen-bonding purine and pyrimidine bases which are herein presumed to bind the genuine aminoacyl-RNA molecule to the ribosomal surface.

Dr. ALLEN has recently found [24] in our laboratories that 2-3 M salt-free system hydroxylamine acta in this same cell-free system in a manner very similar to that of puromycin, inducing the release of peptides from the ribosomal surface. This action of hydroxylamine would be expected from its esterolytic effect on aminoacyl-RNA. There are now available, therefore, two agents — puromycin and hydroxylamine — which can be used in time sequence studies of nascent peptide synthesis on the ribosomal surface.

The time is also nearly at hand when gentle sequential degradative procedures [25] [26] may be applied to relatively pure species of transfer RNA molecules.

REFERENCES

[1] CRICK F.H.C., *Soc. Exper. Biol.*, Symposium XII, London, 1958.
[2] HOAGLAND M.B., STEPHENSON M.L., SCOTT J.F., HECHT L.I. and ZAMECNIK P.C., *J. Biol. Chem.*, *231*, 241 (1958).
[3] ALLEN E., GLASSMAN E., CORDES E. and SCHWEET R., *J. Biol. Chem.*, *235*, 1068 (1960).
[4] HOAGLAND M.B., KELLER E.B. and ZAMECNIK P.C., *J. Biol. Chem.*, *218*, 345 (1956).
[5] HECHT L.I., STEPHENSON M.L. and ZAMECNIK P.C., *Biochim. et Biophys. Acta*, *29*, 460 (1958).
[6] ZACHAU H.G., ACS G. and LIPMANN F., *Proc. Natl. Acad. Sci.*, *44*, 885 (1958).
[7] PREISS J., BERG P., OFENGAND E.J., BERGMANN F.H. and DIECKMANN M., *Proc. Natl. Acad. Sci.*, *45*, 319 (1959).
[8] HECHT L.I., STEPHENSON M.L. and ZAMECNIK P.C., *Proc. Natl. Acad. Sci.*, *45*, 505 (1959).
[9] DAVIE E.W., KONINGSBERGER V.V. and LIPMANN F., *Arch. Biochem. Biophys.*, *65*, 21 (1956).
[10] HOAGLAND M.B., ZAMECNIK P.C. and STEPHENSON M.L. in *A Symposium on Molecular Biology* (Chicago University Press, 1959), chapter 9, p. 105.
[11] STEPHENSON M.L. and ZAMECNIK P.C., *Proc. Natl. Acad. Sci*, *47*, 1627 (1961).
[12] ZAMECNIK P.C., STEPHENSON M.L. and SCOTT J.F., *Proc. Natl. Acad. Sci.*, *46*, 811 (1960).
[13] SAPONARA A. and BOCK R.M., *Federation Proc.*, *20*, 356 (1961).
[14] VON PORTATIUS H., DOTY P. and STEPHENSON M.L., *J. Am. Chem. Soc.*, *83*, 3351 (1961).
[15] PORATH J. and FLODIN P., *Nature*, *183*, 1657 (1959).
[16] MCKERNAN W.M. and RICKETTS C.R., *Biochem. J.*, *76*, 117 (1960).
[17] STEPHENSON M.L. and ZAMECNIK P.C., *Biochem. and Biophys. Res. Com.*, *7*, 91 (1962).
[18] BROWN G.L. and ZUBAY G., *J. Mol. Biol.*, *2*, 287 (1960).
[19] ISHERWOOD F.A. and JERMYN M.A., *Biochem. J.*, *48*, 515 (1951).
[20] LEDERER E. and LEDERER M., *Chromatography. A review of principles and applications*. Second edition. Elsevier, London, 1957, p. 246.
[21] ALLEN D.W. and ZAMECNIK P.C., *Biochim. et Biophys. Acta*, *55*, 865 (1962).
[22] SCHWEET R., LAMFROM H. and ALLEN E., *Proc. Natl. Acad. Sci.*, *44*, 1029 (1958).
[23] YARMOLINSKY M.B. and DE LA HABA G.L., *Proc. Natl. Acad. Sci.*, *45*, 1721 (1959).
[24] ALLEN D.W. (unpublished data).
[25] YU C.T. and ZAMECNIK P.C., *Biochim. et Biophys. Acta*, *45*, 148 (1960).
[26] KHYM J.X. and COHN W.E., *J. Biol. Chem.*, *236*, PC 9 (1961).

DISCUSSION

FRAENKEL-CONRAT

My question regards the periodate oxidation and stepwise degradation method which many people are now investigating. Yet nobody is really publishing any data on successful degradation of oligonucleotides. Do you have any such data, or can you tell us how well it works?

ZAMECNIK

On s-RNA we can say (ref. 25) that the periodate-cyclohexylamine-monoesterase procedure has been carried through four degradation steps. The recoveries are of the order of 90% or more for each step, but it is difficult to give a more quantitative answer at the present time. In the case of mononucleotides and nucleosides, the recoveries of bases are virtually 100%. Since the time of this conference, Dr. YU has degraded the trinucleotide ApApC, kindly furnished by Dr. A.M. MICHELSON, with recovery of 95-100% of the calculated values for the two bases (cytosine and adenine) and the 3'-adenylic acid residue, which result from the two degradation of steps.

CHARGAFF

I have a couple of questions I want to ask you. To begin with, when one prepares supernatant RNA, say by the phenol procedure, of course, not all of it is transfer RNA, and if one separates, for instance, such a supernatant RNA, into the NaCl soluble RNA preparation and into the NaCl insoluble portion, one finds a very different distribution of nucleotides. That is, this NaCl soluble

portion is really probably mostly the transfer RNA. There is quite an important fraction in the supernatant RNA which doesn't seem to have transfer functions, which may be a piece of the ribosomal RNA broken off in the course of the protein synthesis, in which it has — I think — been neglected. It is quite an interesting fraction.

And in this connection, I was wondering if you have an idea of what can be the reason for the difficulty of preparing an amino-acyl-s-RNA following periodate oxidation.

I wonder whether you have any explanation for both the difficulties in the dye procedure and also in your adsorption procedure, i.e. what really is not oxidized?

ZAMECNIK

I can't give a completely satisfactory answer to these questions. It is possible that a small fraction of the periodate oxidizable sites on the terminal ribosyl moieties of the transfer RNA are inaccessible, due to aggregation of molecules of transfer RNA. The hydrazone and dye coupling reactions may also fail quite to reach 100% of completion. It is also possible that some, that is a small percentage, of spontaneous degradation of the dye-RNA compound may occur.

TISELIUS

In connection with the remarkable separation on DEAE-Sephadex reported by Dr. ZAMECNIK, I should like to point out that too little is yet known about the exact mechanism of the separation on Sephadex material. With unsubstituted Sephadex we have essentially a « molecular sieve » effect, even though specific affinities clearly often play a role. With substituted Sephadex material this is clearly the case, and the many advantages of these synthetic polymers (great permeability, good stability, easy introduction of additional groups of desired affinity) would seem to offer many possibilities, some of which are now being explored in my laboratory and in the laboratories of Pharmacia Ltd. in Uppsala. It would be valuable to us to receive suggestions as to systems and separation problems for which special properties are desired.

CONCLUSIONS

CONCLUSIONS

I) FROM THE POINT OF VIEW OF MOLECULAR BIOLOGY (1)

Structural studies of biological macromolecules have passed through several distinct but overlapping stages. Without too much over-simplification one might distinguish successively a carbohydrate phase, a protein phase and a nucleic acid phase. The effort devoted to the investigation of a particular class of macromolecule at any given time is a complex function of its availability in pure, or at least well-characterized, form; of the state of development of techniques appropriate for its study; and of the biological significance of the material in question. The history of nucleic acid research offers good examples of the interplay between these factors. For many years the biological significance of nucleic acids as information carriers was not generally understood, or was at best dimly apprehended by a few; methods of obtaining « pure » native specimens were primitive or non-existent, simply because the incentive to develop such methods, or techniques for investigating the properties of the extracted material, was lacking. Since the central biological role of nucleic acids has become generally recognised, techniques have advanced at a prodigious rate, as will be realized by those attending the present Study Week if they cast their minds back, as the author found himself doing while he listened to this series of papers, to nucleic acid sym-

(1) Compiled by J.C. KENDREW.

posia of ten or fifteen years ago. In fact the dramatic developments in the nucleic acid field, discussed in the accompanying article by Prof. CHANTRENNE, have secured so much public notice as to divert attention and interest from the less spectacular, but equally important, developments in protein studies. It is true that nucleic acids have been shown to have a unique and indispensable role in being responsible for the carriage and storage of information; but it remains true also that the information they purvey cannot be used except through the mediation of proteins. Investigations of the structure and function of proteins remain, therefore, of central importance in our attempts to understand the workings of living cells. Progress in these investigations was exemplified in a wide variety of the paper presented during the Study Week.

Basic to our knowledge of protein behaviour must be a determination of the amino acid sequence. It is only about ten years since the first sequence determination was carried out by SANGER, and since then, in spite of intensive study in many laboratories, the technique has remained a slow and laborious one; even today the number of protein sequences completely and unambiguously determined can be counted on a hand. One example of such a complete determination is that of horse cytochrome *c* discussed by TUPPY. The chemical mode of attachment of the haem group to the protein moiety is known, but it is unfortunate that as yet we have no knowledge of the three-dimensional make-up of this ubiquitous and very important protein. TUPPY in his paper reports, however, on comparative studies of cytochrome from a number of different species, and from the homologies between these it is possible at least to see which amino acids are important for function, even if the reason for their importance is still obscure. In any event, studies like those of KREIL & TUPPY, and of MARGOLIASH & SMITH with whom they collaborated at long range, will always remain of basic importance to protein studies, for reasons which will be suggested below; perhaps the most urgent

requirement in the field today is that more rapid and certain methods be developed, so that amino acid sequences can be determined as a matter of routine, not (as now) requiring years of patient work.

Logically following a determination of the amino acid sequence of a protein would be the elucidation of its three-dimensional structure. The only direct methods capable in principle of doing this is that of X-ray crystallography — and even here we have problems concerning the relation between the structure of the protein in the crystalline phase (as determined by X-rays) and in solution under physiological conditions. Principle has been converted into practices only very recently, however, and so far in very few cases and then still partially (as decribed in the papers of PERUTZ and KENDREW). There has, therefore, been a strong incentive to develop other methods, less direct and capable of giving less complete information, but at least able to measure in a straightforward manner some of the parameters of the molecule in solution.

Prominent among these have been the hydrodynamical methods which give values for a thermodynamically related set of parameters. An example was provided in the Study Week in the paper of DEBYE who showed how measurements of light scattering could give information not only about the molecular weight and dimensions of a macromolecule, but also about interaction forces *between* molecules by measurements made in the neighbourhood of the critical mixing point. Such measurements can be compared with those derived by other related physico-chemical methods; the difficulty comes when one tries to interpret the « equivalent ellipsoids » of the hydrodynamician in terms of a real molecular structure.

An alternative approach has been to use techniques sensitive to particular elements of a structure, rather than to some parameter of the structure as a whole. The infrared, ultraviolet and RAMAN spectra studied by MIZUSHIMA are good examples of this approach; particular spectral lines can be used for

diagnostic purposes and their precise frequencies may be a measure of particular bond types. Such methods may be more fruitfully applied to model compounds (amino-acids and peptides) and the results used to define the limits within which the configuration of the elements of the polypeptide chain may vary, and hence, by means of suitable mathematical tools, to enumerate possible stable configurations of the chain. Such calculations do not in general lead to a unique solution for a particular structure and this must be sougbt by more direct experimental attack, e.g. by the analysis of the X-ray diffraction pattern of a protein fibre.

The application of physico-chemical methods has also led to the development of extremely sensitive methods for the separation of very similar molecules — for example the partition and the gel filtration methods applied to protein separation in the laboratory of TISELIUS; the latter in particular have very rapidly found a place in every laboratory of protein chemistry. The development of such methods of high sensitivity is essential to progress, since it has become clear that in very many cases proteins are not homogeneous, and normally occur as mixtures of slightly different species, which must be separated before they can be adequately characterized.

Another aspect of protein structure which has received much attention in the past has been the nature of the forces stabilizing the three-dimensional arrangement of the molecule, in other words the intramolecular interactions. A number of types of force may be enumerated — covalent bonds (e.g. disulphide bridges), electrostatic attractions between oppositely-charged residues, hydrogen bonds, and VAN DER WAALS attractions between non-polar groups. A discussion of the role of all these forces has been given in the paper by PUTZEYS, who emphasizes the role played by each and every type of force; it is most improbable that any one type has overwhelming significance or is not represented at all. Quantitative assessment may, however, be beyond the power of the classical techniques and

must await explicit solution of three-dimensional protein structures. Thus in myoglobin the high resolution analysis has shown that the hydrophobic bonds between internal non-polar residues are quantitatively the most important; virtually all the polar residues are on the surface of the molecule and have relatively few interactions among themselves, though a few of these may have special importance in determining the conformation of the polypeptide chain. It is still not clear, however, how far these features of myoglobin are characteristic of proteins in general since myoglobin is in many ways not a typical protein, having for example a much higher helix content than seems to be normal.

What of interactions between a protein molecule and other substances? These may be of many different kinds and must include all those special interactions responsible for protein function, such as enzyme-substrate and antigen-antibody interactions. Here we enter a very wide field, which would require much more than a week's meeting for adequate survey. The Study Week did, however, include accounts of a representative selection of interactions and of techniques for studying them.

The special role of metals in enzyme reactions has been long appreciated though it is perhaps only recently that the ubiquity of metal-protein interactions has been realized. The structural basis of such interactions needs precise definition; and here X-ray studies of complexes between metal ions and simple peptides are particularly important. LINDQVIST's studies of complexes between zinc or copper and amino-acid side chains or peptide bonds are relevant, for example, to the interaction involved in the active site of carboxypeptidase; this is an interaction which might in principle be studied both at the level of a model complex and in the intact protein. The role of zinc in stabilizing enzyme structure was further emphasized by THEORELL in his discussion of the bonding between nucleotide coenzymes and enzyme proteins, in which zinc plays an essential part.

Another widespread type of interaction is that between proteins and phosphate groups. This field, as reviewed by SILIPRANDI, revealed itself as one of undoubtedly wide significance, but one so far little studied. Considering how general is the distribution of phosphoproteins it is remarkable how little is known of their structure or function, and this — like the field of protein-metal complexes — is yet another example of a common tendency for protein research to be concentrated on the polypeptide part of the molecule, to the exclusion of quantitatively minor constituents.

As an example of a quite different type of protein interaction, we may note the remarkable reactions reported by KATCHALSKI and used by him to prepare water-insoluble enzymes suitable for incorporation in columns as analytical tools. Here indeed we see the beginnings of a direct control over the properties of macromolecules by means of deliberate modification — tailoring, almost — of their structures to suit the purpose in hand. These enzymes, covalently attached to an inert polymer, are prototypes of the kind of intentionally engineered molecule which we may expect to have available in the future, as a result of our rapidly increasing knowledge of protein structure and behaviour.

Finally we may note that aggregates of protein units, and sub-unit structures in protein molecules, should also be regarded as examples of protein interaction. The notion of sub-unit structures in protein molecules, should also be regarded as examples of protein interaction. The notion of sub-units in protein molecules dates back to the classical researches of SVEDBERG on the haemocyanins; and a recent example in which the precise structural basis has been revealed is provided by PERUTZ's X-ray analysis of haemoglobin, which has shown that this molecule — which only ten years ago was not generally credited even with the possession of four sub-units — is made up of four sub-units each of which closely resembles the molecule of myoglobin. Several other examples of protein

sub-units emerged in the deliberation at the Pontifical Academy — the plant viruses discussed by RUBIO HUERTOS, the sub-units of TMV described by FRAENKEL-CONRAT, and experiments on hybridization in haemocyanins and haemoglobin cited by TISELIUS. Sub-unit dissociation and recombination, it now appears, may be a most important biological device underlying some types of enzyme control mechanism. Here is a field too little explored, though both techniques and materials lie ready to hand, and undoubtedly rich with promise for the future.

In all these studies one is trying to develop the relationship between structure on the one hand, behaviour and function on the other. In the past the limiting factor has been the paucity of our knowledge of structure. Now that protein structure can in principle be determined we may look forward to a period in which all the classical methods can be calibrated, as it were, against a precisely known structure. It is likely, however, that protein structure determination by X-ray methods will remain a very tedious business, and for many proteins, available in very small quantities or uncrystallizable, an impossibility. It is for this reason that the researches of ANFINSEN, reported at this meeting, take on a special significance quite aside from their intrinsic interest and importance. For his results suggest that in general proteins are capable of folding *themselves* up into their specific three-dimensional conformation which must, therefore, be assumed to be thermodynamically stable; or, to put the problem another way, the genetic information determines only amido acid sequence — the conformation of the molecule is inherent in, and can in principle be deduced from, as *in vivo* it *is* derived from, that sequence. What we need today is a more intensive study of the structure, behaviour and function of selected proteins for which all three aspects are experimentally accessible, with the object of developing generalizations as to be relations between the three. In the long

run it may be that a knowledge of amino acid sequence — or even base sequence in the corresponding DNA? — will be enough for all our purposes. This is a distant gaol which will perhaps not be attained in our generation; nevertheless our recognition of it will serve to define the direction of our studies and to focus them on fruitful problems.

2) FROM A BIOCHEMICAL POINT OF VIEW (1)

The meeting of the Pontifical Academy of Science celebrates, as it were, the birth of molecular biology, and testifies to the dynamism of research in a field where physics, chemistry and biology meet and join forces.

The macromolecules of which living organisms are made possess a remarkably organized structure, they are organs of molecular dimensions. Organization, regulation, morphogenesis, heredity, evolution, most of the problems of general biology, are posed not only by the organisms, but also by the macromolecules themselves; these problems must now be approached at the molecular level, with the concepts and methods of physics and chemistry.

The first task of biochemistry was to isolate and describe the low molecular weight constituents of organisms. The isolation of the enzymes and the study of their action unveiled the chemical processes which take place within living organisms. A big advance was made with the elucidation of the principles of energy production, and of the mechanisms of transfer and delivery of energy. This permitted the discovery of biosynthetic processes, and led to the studies of the pathways of synthesis of cell constituents, which still continue. A further step is the current investigation of the fine structure of proteins and nucleic

(1) Compiled by H. CHANTRENNE.

acids, and of the nature of structural information, its conservation, transfer and expression.

The latter studies required the acceptance of a new type of biochemical reaction, the assembly of the building blocks of macromolecules on a template. Although the existence of molecular organizers was regarded for many years as a possibility and even as a necessity on purely logical grounds, many biochemists were long reluctant to consider their existence seriously. It is now experimentally established that DNA, protein and at least certain classes of RNA are synthetized in template processes. Biochemists might paraphrase VIRCHOW's aphorism in the following way: any « information rich » molecule comes from an « information rich » molecule.

The basic assumption, that genetic information consists in a meaningful arrangement of the four bases along the nucleic acid backbone, is now inescapable. The most sceptical chemists must accept it, even if they dislike the vocabulary and protest — with some reason — against the onesightedness of some molecular biologists. The present picture may be too schematic, there may be other sources of information than the nucleotide sequence (although this remains to be shown), complications and difficulties certainly remain. But although the emphasis is here placed on the brilliant success already obtained, the shortcomings of our present concepts are not forgotten; they are, on the contrary, the major preoccupation of the leading workers in the field, and a stimulating challenge for further research.

DNA synthesis was the first template process for which direct experimental evidence was obtained. It is not yet completely understood: is the sequence completely determined by base pairing, are there influences or restrictions from neighbouring nucleotides? How, in the absence of DNA primer is the adenine-thymine polymer made? The distribution of pyrimidine sequences in bromouracil-containing DNA, as reported by CHARGAFF, should be further investigated, because it ques-

tions the basic concept of information in nucleic acid sequence: on the one hand, it is known that certain bacteria can incorporate bromouracil in place of thymine without suffering abnormalities, therefore without changing the genetic information; yet CHARGAFF's results indicate, assuming there is no unknown flaw in the method for analysing the pyrimidine distribution, that bromouracil changes the base sequence.

Turning to less chemical, more biological aspects of information transfer, we must consider the process of recombination in bacteriophage. MESELSON's work illustrates the intelligent combination of several powerful methods of investigation now at our disposal: genetic analysis, radioactive tracer techniques, density gradient equilibrium. This analysis of a biological process at the molecular level demonstrates that genetic recombination actually involves exchange and reassociation of DNA segments. Not more than a year ago, the general belief was that genetic recombination, especially within a locus, did not reflect actual rupture and reassociation of genetic material. Rather, it was thought that recombination occurs during DNA duplication by a process called copy choice, in which the new strand of DNA is copied first from one DNA molecule and then continues from another. Although copy choice has not been eliminated as a possible mechanism of recombination, direct experimental demonstration that exchanges and reassociation occur is certainly a fundamental achievement. The results also emphasize that recombination is not an exceptionnal event, as accident of the process of replication, for it can occur in the absence of DNA replication. Biochemists, no doubt, will soon be searching for a recombination enzyme.

The insight into the expression of genetic information and the regulation of protein synthesis presented by JACOB is one of the most brilliant contributions in this field. One of the essential features of the theory is the role and properties of a special type of RNA, the « messenger ». The existence of an RNA receiving information from the gene was long postulated

by workers on animal and plant systems, but it was commonly assumed that this was the ribosomal RNA. In bacteria, such a view was difficult to reconcile with kinetic data. The novel feature of bacterial messenger-RNA is the assumption that it is short-lived, that it carries the message to the protein-making system and is destroyed thereafter.

Short-lived RNA fractions were indeed observed in bacteria. The most convincing evidence that such RNAs possess a base sequence similar to that of the DNA is provided by SPIEGELMAN's experiments. The rapidly labelled RNA, which can bind to ribosomes under suitable conditions, can be made to associate specifically with the DNA of the strain which produced it. The specific hybridization obtained by cooling a mixture of DNA and short-lived RNA is the best experimental evidence that the base sequence of this RNA is complementary to that of DNA and therefore that it contains genetic information and can rightly be called informational RNA. Hybridization of RNA with specific DNA might eventually make it possible to isolate the informational RNA which carries the message for a single protein, a possibility which was undreamt of only a few months ago.

The current theory of information transfer and regulation in bacteria clarifies the field and, furthermore, raises or helps formulate many questions. Restricting ourselves to the biochemical point of view let us mention a few questions which, no doubt, everyone has in mind at present. How is messenger RNA attached to the ribosome, how is it degraded, how many copies of a polypeptide chain can one molecule of messenger organize before it is inactivated, what is the function of ribosomal RNA and of ribosomal protein, what is the origin of ribosomal RNA, are all the ribosomes of a bacterium identical, or do they possess some degree of specificity with respect to the messenger they can accomodate, how is the synthesis of the ribosome controlled? What are the chemical symbols which mark the beginning or the end of a cistron, of an operon, of

an operator? How does the product of a regulatory gene differ from that of a structural gene? What is the chemical nature of the active repressor? At what level does the inducer act? Do the regulating mechanisms discovered in bacteria operate in higher organisms? To what extent do they account for regulation of enzyme level and for differentiation? Does differentiation include a chromosomal control of a type which does not exist in bacteria? There are more questions which concern genetic problems. Clearly, there remains much to be done, and the coming years will certainly be very rewarding.

Ever since their discovery, viruses, and especially TMV, have been a source of fundamental knowledge on macromolecules of biological significance. Viruses indeed possess several major attributes of organisms. They are parasites at the molecular level. A virus such as TMV, for example, seems to have reached the ultimate degree of simplicity, since it is reduced to informational RNA with a sheath to protect it from damage.

Chemical modification of virus RNA has provided much direct data on the nature of information and coding; investigations of this RNA will certainly continue to be useful. Several of the observations raise questions: does virus RNA code for other proteins besides the sheath protein (as phage DNA does)? What is the reason for the appearance of mutants with many amino acid replacements? How is virus RNA replicated? Is any DNA involved in this process?

Isolation of genetically activ virus RNA has prompted chemists to the bold project of synthetizing an active RNA, permitting it to reduplicate like the RNA of a virus and cause the production of proteins. SCHRAMM's experiments make one wonder whether synthetizing RNA is as formidable a task as organic chemists believed, and permit one to hope that shortcuts in the synthetic process may eventually be found in simple chemical reactions, which would be controlled or at least influenced by templates.

The simplest expression of genetic information is the appearance of a specific polypeptide or of a recognizable protein. Detailed knowledge of the process of protein synthesis is rapidly advancing. A thorough description of the current state was presented by LIPMANN, who also reported new experiments.

According to the present theory, amino acids react with ATP to form mixed carboxyl-adenyl anhydrides; the activated carboxyl groups are then bound in a high energy ester bond to the adenosine end of transfer RNAs, which are specific acceptors of individual amino acids. The individual transfer RNAs can be partially separated from one another by electrophoresis, counter current distribution or chromatography. The very ingenious purification and isolation method discussed by ZAMECNIK seems superior to those previously described. Good fractionation procedures for nucleic acids are urgently needed, for the very low efficiency of the physical methods now in common use greatly restricts the analysis of biochemical processes involving RNA.

Availability of transfer RNA simplifies the study of protein synthesis in vitro, and makes it possible to concentrate on the template process itself. Once the amino acids are attached to transfer RNAs, the problem of providing energy is solved; there remains the arrangement of the amino acids in the right sequence on the ribosomes. The reticulocyte ribosome is especially useful, for it contains the information required for the synthesis of haemoglobin, one of the best characterized proteins (cf. PERUTZ' report). Apparently the genetic message remains for a long time in these ribosomes. Most interesting is that the message can be deciphered and correctly interpreted by bacterial transfer RNA (although the process may require the help of some enzyme-like factor which seems to be specific to the ribosome). The genetic code must therefore be universal. The starting recent results of NIRENBER, mentioned by Dr. LIPMANN in his report, make it almost certain that the genetic code will be solved within a few years. These experi-

ments are at the same time a direct verification of part of the messenger model, since E. coli ribosomes can accomodate virus RNA, and participate in the synthesis of virus protein.

The exact mechanism by which transfer RNAs find their correct place on the template is not clear. Base pairing must in some way be involved in the recognition of coding sequences, but, in addition, GTP must somehow be included in the picture; more research is needed on this point. The action of puromycin, which apparently plays the part of an analogue of transfer RNA and falls from the template together with any piece of peptide which happened to bind to it, may prove a very useful tool in the study of the template process and of the interactions between sRNA and ribosomes.

Purine and pyrimidine analogues probably interfere with the template mechanism a better understanding of the mechanism by which they induce misreadings of the information is needed. Acellular systems of the kind used by NIRENBERG or by LIPMANNN are the obvious material to be employed at present.

When all the problems of transfer, transcription and translation of genetic information are solved, there will remain the question of how the genetic information which controls the sequence of the amino acids eventually expresses itself in an enzyme activity. It is quite clear that secondary and tertiary structures of proteins are all-important for enzyme activity or, in fact, for the physiological function of proteins in general.

An important question is whether the folding of polypeptide chains into a unique structure, and their association into active enzymes, present new problems of information and control, or whether they are automatic consequences of the amino acid arrangement. The simplest assumption is certainly that folding is entirely determined by the amino acid sequence. Studies on ribonuclease reported by ANFINSEN strongly suggest this, since certain unfolded polypeptide chains can spontaneously return to the proper configuration. Evidence in the same di-

rection was also presented by FRAENKEL-CONRAT for the case of virus proteins. A comparison of the structures of haemoglobin and myoglobin (PERUTZ and KENDREW) is very interesting in this respect. Here are two molecules which accomplish almost the same function; their tertiary structures are very similar, although their amino acid compositions differ widely. Closer examination, however, reveals that a few amino acids occupy identical positions in the two chains of hemoglobin and in myoglobin: in the tridimensional model, these amino acids are seen to occupy key positions which ultimately control folding of the chain, or interactions between regions of the chain or with the hematin group. It would seem that, for physiological activity, folding of the chain and positioning of a few amino acids are all-important, but variation is permissible in the rest of the structure. Comparison of the cytochromes (TUPPY) of different species also led to similar conclusions. This gives a first glimpse of the molecular basis of evolution. A darwinian process can indeed be outlined in precise molecular terms: mutation resulting in an amino acid replacement may change folding and lead to a better or worse enzyme, with the ultimate result of selective advantages or disadvantages.

A rational explanation of the emergence of the remarkably adapted structures of macromolecules like myoglobin and haemoglobin can thus be imagined.

Our knowledge of the mechanism of enzyme reactions is not very advanced, but studies like those reported by Dr. THEORELL indicate that progress may soon permit an understanding of how an enzyme functions. In the near future, chemists may feel that synthetizing a polypeptide endowed with enzymic activity is almost within their reach.

The far reaching progress made in understanding the mechanism of genetic control gives a solid background for approaching the next problems of biological research, differentiation and morphogenesis. Within the coming years, old problems familiar to embryologists will regain importance, reworded

in accordance with our present knowledge in the terms of molecular biology.

It is to be expected that, after some faltering and seaching in blind alleys, important clues will be obtained, and the molecular basis of differentiation will be uncovered just as that of heredity is presently being uncovered. We have seen in RANZI's paper examples of the way in which very complex processes of morphogenesis can be influenced by the experimenter. Embryological literature contains a wealth of observations waiting for interpretation. We are in a better position than ever before to approach this difficult task.

All the Participants unanimously agreed that colleagues CHANTRENNE and KENDREW should compile the above reported conclusions, arrived at during the Study Week.

They reflect the discussions which took place during the whole week according to what has been established by the Participants during their meeting on Thursday October 26th, 1961 in the peace and quiet of the Pontifical Gardens at Castelgandolfo.

ANFINSEN, CHANTRENNE, CHARGAFF, DEBYE, DE HEVESY, FRANKEL-CONRAT, GIACOMELLO, JACOB, KATCHALSKI, KENDREW, LINDQVIST, LIPMANN, LIQUORI, MESELSON, MIZUSHIMA, PERUTZ, PUTZEYS, RANZI, RICH, ROSSI-FANELLI, RUBIO-HUERTOS, SCHRAMM, SILIPRANDI, SPIEGELMAN, THEORELL, TISELIUS, TUPPY, ZAMECNICK.

CONCLUSIONS

1) DU POINT DE VUE DE LA BIOLOGIE MOLECULAIRE (*)

L'étude de la structure des macromolécules biologiques a passé à travers plusieurs phases distinctes mais interdépendantes. Sans trop simplifier, nous pouvons distinguer successivement une phase caractérisée par un carbo-hydrate, une phase protéique et une phase caractérisée par un acido nucléique. L'effort voué à l'étude d'une classe particulière de macromolécules est à tout moment une fonction complexe de son existence à l'état pur, ou pour le moins dans sa forme caractéristique de l'état du développement des techniques propres à son étude et du sens biologique du matériel en question. L'histoire de la recherche de l'acide nucléique offre des exemples de l'interdépendance de ces facteurs. Pendant plusieurs années, le sens biologique des acides nucléiques comme agent d'information n'a pas été généralement compris, out tout au plus vaguement entrevu par quelques-uns. Les méthodes pour l'obtention de spécimens originaux purs étaient primitives ou presque inexistantes. Tout ceci, parce que les occasions de développer pareilles méthodes ou techniques pour étudier les propriétés du matériel d'extraction faisaient défaut. Depuis qu'a été reconnue la prépondérance du rôle biologique des acides nucléiques, les techniques ont fait des progrès énormes, comme pourront s'en rendre compte tous ceux qui ont pris part à cette Semaine

(*) Rédigée par J.C. KENDREW.

d'Etude, s'ils retournent en arrière en pensée comme l'auteur l'a fait lui-même, lorsqu'il écoutait cette série des rapports, aux symposia sur l'acide nucléiques il y a bien dix ou quinze ans de cela. En effet, les développements dramatiques dans le domaine de l'acide nucléique, exposés dans l'article du Professeur CHANTRENNE, ont recueilli tant d'approbations qu'ils ont détourné l'attention et l'intérêt des développements moins spectaculaires, mais aussi importants dans l'étude des protéines.

Il est vrai qu'il a été démontré que les acides nucléiques ont un rôle unique et indispensable, étant responsables de la transmission et l'accumulation des informations, mais il est tout aussi vrai que les informations qu'ils transmettent ne peuvent être utilisées qu'à travers les protéines.

L'étude de la structure et de la fonction des protéines reste donc d'importance capitale dans nos efforts pour comprendre le travail des cellules vivantes. Le progrès dans ces études fut exposé largement dans les rapports présentés pendant cette Semaine d'Etude.

La base pour la compréhension du comportement des protéines doit être la détermination de la séquence amino-acides. Il y a à peine 10 ans la première détermination d'une séquence fut effectuée par SANGER et depuis lors, malgré les recherches dans plusieurs laboratoires, la technique est restée lente et laborieuse; aujourd'hui même le nombre des séquences des protéines, déterminé de façon complète, peut se compter sur les bouts des doigts. Un exemple de pareille détermination est celui du cytochrome *c* du cheval, étudié par TUPPY. La nature chimique de l'attachement du groupe de l'hémène du milieu de la protéine est connue, mais il est regrettable que jusqu'à présent on n'ait pas connaissance de la structure tridimensionale de cette ubiquitaire et très importante protéine. Toutefois, TUPPY dans ses études comparatives du cytochrome d'un nombre d'espèces différentes et des analogies entre celles-ci démontre qu'il est possible pour le moins de voir quels amino-acides sont importants pour le fonctionnement, même si la

raison de leur importance est encore inconnue. Toutefois, des études comme celles de KREIL et TUPPY et de MARGOLIASH et SMITH avec qui ils collaborèrent assez longtemps, resteront toujours d'importance capitale dans l'étude des protéines pour des raisons que nous exposerons plus bas. De nos jours, la nécessité la plus urgente dans ce domaine-là, est celle de la mise au point de méthodes plus rapides et plus sûres, de sorte que les séquences des amino-acides puissent être déterminées comme une routine et non pas comme il en est actuellement, demandant des années de travail patient. Logiquement, la détermination de la structure tridimensionale de la protéine devrait faire suite à la détermination de la séquence des amino-acides. La seule méthode directe, capable en principe de le faire, est l'analyse cristallographique aux rayons X. Et même dans ce cas, nous avons des problèmes concernant les relations entre la structure protéique dans la phase cristalline (comme déterminée aux rayons X) et en solution dans les conditions physiologiques. Toutefois, les principes ont été traduits en pratique tout récemment, pour quelques cas seulement, et cela même partiellement (comme décrit dans les contributions de PERUTZ et KENDREW). Il y a donc eu une forte stimulation au développement d'autres méthodes moins directes et capables de donner des informations moins complètes, mais au moins capables de mesurer d'une façon directe certains des paramètres des molécules en solution.

Parmi ces méthodes, une place de premier ordre occupent les méthodes hydrodynamiques qui donnent des valeurs pour toute une série de paramètres thermodynamiques. Un exemple fut rapporté au cours de la Semaine d'Etude dans la contribution de DEBYE qui montra comment les mesures de la dispersion de la lumière pouvaient fournir des informations non seulement sur le poids moléculaire et les dimensions d'une macromolécule, mais aussi sur l'interaction des forces *entre* les molécules, en effectuant des mesures au voisinage du point de mélange critique. On peut comparer ces mesures avec celles obte-

nues au moyen d'autres méthodes physico-chimiques. La difficulté surgit lorsqu'on essaie d'interpréter les « ellipsoides équivalents » en termes d'une structure moléculaire.

Une autre tentative est celle de faire usage de techniques, sensibles à des éléments particuliers d'une structure, plutôt qu'à des paramètres de l'ensemble de la structure. Les spectres infrarouges, ultraviolets et Raman, étudiés par MIZUSHIMA sont des exemples d'une pareille tentative; des lignes spectrales particulières peuvent être utilisées pour des raisons de diagnostic et leurs fréquences précises peuvent être une mesure d'une liaison particulière.

Ces méthodes peuvent être appliquées avec plus de succès à des composés-modèles (amino-acides et peptides) et les résultats utilisés pour définir les limites entre lesquelles la configuration des éléments de la chaîne du polypeptide peut varier, et en faisant usage de données mathématiques appropriées, l'on peut établir la configuration stable de la chaîne. De pareils calculs ne conduisent pas en général à une solution unique pour une structure bien déterminée et qui doit être établie par une étude expérimentale plus directe, comme p.ex. par l'analyse de la diffraction aux rayons X d'un modèle d'une fibre protéique.

L'emploie de méthodes physico-chimiques a conduit au développement de méthodes extrêmement sensibles pour la séparation de molécules très semblables, par exemple les méthodes de répartition et de filtration du gel, appliquées à la séparation des protéines au laboratoire de TISELIUS. Ce dernier, en particulier, a trouvé rapidement sa place dans chaque laboratoire de la chimie des protines. Le développement de pareilles techniques de haute sensibilité est essentiel, étant donné qu'il ressort que, dans plusieurs cas, les protéines ne sont pas homogènes et se présentent normalement comme des mélanges d'espèces légèrement différentes qui doivent être séparées avant d'être caractérisées d'une façon adéquate.

Un autre aspect de la structure des protéines, qui a fait

l'objet d'étude dans le passé, est la nature des forces qui stabilise l'arrangement tridimensional de la molécule, autrement dit l'interaction intramoléculaire.

Un certain nombre de types de forces peut être mentionné; liaisons covalentes (p. ex. disulfites), attractions électro-statiques entre des résidus chargés différemment, liaisons hidrogènes, et attractions de VAN DER WAALS entre des groupes non polaires. Une discussion du rôle de toutes ces forces a été faite dans le rapport de PUTZEY qui a souligné le rôle joué par chacune de ces forces; il est très improbable qu'une de ces forces ait un rôle primordial ou ne soit pas du tout mentionnée. La détermination quantitative peut cependant être au-delà du pouvoir de la technique classique et doit attendre une solution explicite des structures tridimensionales des protéines.

Ainsi, bien que dans la mioglobine l'analyse ait montré que les liaisons hydrofobes entre les résidus internes non polaires sont quantitativement les plus importantes, pratiquement tous les résidus polaires sont à la surface de la molécule et ont une interaction relativement peu prononcée entre elles, bien que seulement quelques-unes d'entre elles puissent avoir une certaine importance dans la détermination de la configuration de la chaîne du polypeptide. Cependant, il n'est pas encore très clairément établi jusqu'à quel point ces aspects de la mioglobine, sous différents aspects, n'est pas une protéine typique, ayant par exemple un plus grand contenu en hélices de ce qui doit être normalement.

Qu'en est-il de l'interaction entre la molécule d'une protéine et les autres substances? Celles-ci peuvent être de nature différente et doivent comprendre toutes les interactions particulières responsables de la fonction de la protéine, comme les substrats de l'enzyme et l'interaction antigène-anticorps,

Là, nous entrons dans un domaine très vaste qui demanderait plus d'une semaine pour un examen attentif. La Semaine d'Etude, toutefois, a inclus quelques extraits d'une sélection

représentative des interactions et des techniques afin de les étudier.

Le rôle spécial des métaux dans les réactions enzymatiques a été reconnu depuis longtemps, bien que peut-être ce n'est que tout récemment que l'interaction métal-protéine a été reconnue. Les bases structurelles d'une pareille interaction demandent des définitions précises et dans ce cas, l'étude aux rayons X des complexes entre les ions métalliques et les peptides simples est particulièrement importante. Les études de LINDQVIST sur les complexes entre le zinc ou le cuivre et les chaînes des amino-acides ou les liaisons peptidiques sont à ce sujet importantes, par exemple dans l'interaction interessée dans la partie active de la carboxypeptidase; celle-ci est une interaction qui peut être étudiée en principe soit au niveau d'un modèle complexe, soit dans la protéine intacte. Le rôle du zinc comme stabilisateur de la structure de l'enzyme fut ultérieurement souligné par THEORELL dans sa discussion sur la liaison entre nucléotide, coenzymes et enzymes protéiques, dans lesquels le zinc joue un rôle important.

Un autre type d'interaction très diffusé est celui entre les protéines des groupes des phosphates. Ce domaine, comme l'a d'ailleurs vu SILIPRANDI, se découvre être sans doute d'un grand intérêt, mais par contre très peu étudié. Compte tenu de la grand distribution des phosphoprotéines, il en ressort peu de connaissances que nous avons sur leurs fonctions et leurs structures — et ceci, comme dans le domaine des complexes métal-protéines est un exemple de la tendance commune, et a concentré la recherche pour les protéines sur la part polypeptidique de la molécule, en excluant de la sorte les constituents qui sont quantitativement inférieurs.

Comme exemple d'un type tout autre d'interaction protéique, nous pouvons rappeler les réactions qui ont été reportées par KATCHALSKI et utilisées par ce dernier pour préparer un enzyme insoluble dans de l'eau, et capable d'être incorporé dans des colonnes comme des outils analytiques. Là, effecti-

vement, nous voyons le commencement d'un contrôle direct des propriétés des macromolécules aux moyens de modifications délibérées (presque sur mesure) de leurs structures pour servir le but que l'on s'est proposé. Ces enzymes, liés de façon covalente à un polymer inerte, sont les prototypes de ces molécules expressément bâties, dont nous pouvons nous attendre à en disposer dans l'avenir, comme résultat de notre connaissance plus approfondie du comportement et de la structure des protéines. Enfin, nous pouvons relever que des agrégats d'unités protéiques et des sous-unités structurales dans les molécules pratiques peuvent être considérés comme des exemples d'interactions protéiques.

La notion de sous-unités dans les molécules protéiques datent des recherches classiques de SVEDBERG sur les hémocyanines; et un exemple récent, dans lequel les bases précises de la structure ont été révélées, est fourni par l'analyse au rayons X de l'hémoglobine effectuée par PERUTZ qui a montré que cette molécule — qui, il y a dix ans, n'était même pas reconnue, en général, de possèder quatre sous-unités — consiste en quatre sous-unités chacune ressemblant d'assez près à la molécule de la mioglobine.

Plusieurs autres exemples de sous-unités sont ressortis lors de la délibération à l'Académie Pontificale; les virus des plantes, étudiés par RUBIO HUERTOS, les sous-unités de TMV décrites par FRAENKEL-CONRAT, et les expériences sur les hybridations des hémocyanines et de l'hémoglobine citées par TISELIUS.

La dissociation et la recombinaison des sous-unités qui, comme il semble actuellement, est un des plus importants facteurs biologiques de ceux qui sont soumis au mécanisme de contrôle de certains genres d'enzymes. Nous avons là un domaine très peu exploré bien que le matériel et la technique se trouvent à portée de la main, et qui est riche de promesses pour l'avenir.

Dans toutes ces études on a essayé de développer la relation

entre la structure d'une part et le comportement et la fonction d'autre part.

Dans le passé, le facteur limitant a été notre pauvreté dans la connaissance de la structure. Maintenant que la structure des protéines peut en principe être déterminée, nous pouvons voir pour l'avenir une période, pendant laquelle les méthodes classiques pourront être étallonnées, comme il en a été, contre une structure bien connue et déterminée. De toute façon, c'est un fait que la détermination des structures par la méthode des rayons X sera un travail très ennuyeux et pour plusieurs protéines, disponibles en très petites quantités ou incristallisables, une impossibilité. C'est pour cette raison que les recherches d'ANFINSEN, rapportées à ce meeting, prennent une signification toute spéciale indépendamment de leur importance et de leur intérêt intrinsèque. Ses résultats suggèrent que généralement les protéines sont capables de se dédoubler *elles-mêmes* dans une spécifique conformation tridimensionale qui doit donc être acceptée comme étant thermodynamiquement stable, ou bien, pour placer le problème d'un autre point de vue, les informations génétiques déterminent uniquement des séquences d'amino-acides. La conformation de la molécule y étant alors inhérente et pouvant être en principe déterminée comme *in vivo,* elle en dérive par cette séquence.

Ce qu'il nous faut aujourd'hui, c'est une étude plus intense de la structure du comportement et de la fonction des protéines sélectionnées, pour lesquelles les trois aspects sont accessibles du point de vue expérimental ou ayant pour objet le développement des généralités, tout comme la relation entre les trois facteurs. Dans cette longue course il se peut que la connaissance d'une séquence d'amino-acides ou même une séquence de base dans le correspondant DNA soit suffisante pour tous les buts. Ceci est un but plutôt lointain qui ne sera peut-être pas atteint au cours de notre génération, néanmoins notre connaissance du problème servira à définir la direction de nos études et à les faire converger sur des problèmes profitables.

2) DU POINT DE VUE DE LA BIOCHIMIE (*)

Cette Semaine d'Etude de l'Académie Pontificale des Sciences témoigne du dynamisme de la recherche dans un domaine où la physique, la chimie et la biologie se rencontrent et se complètent; elle célèbre, pourrait-on dire, la naissance de la Biologie moléculaire.

Les macromolécules dont les êtres vivants sont faits possèdent une structure remarquablement organisée: ce sont des organes de dimensions moléculaires. Organisation, régulation, morphogenèse, hérédité, évolution, la plupart des problèmes de la biologie générale sont posés par les macromolécules elles-mêmes; c'est au niveau moléculaire, avec les concepts et les méthodes de la physique et de la chimie que ces problèmes doivent être attaqués.

La première tâche de la biochimie fut d'isoler et de décrire les constituants à poids moléculaire faible des organismes. L'isolement des enzymes et l'étude de leur action permirent de découvrir les processus chimiques qui se déroulent dans les organismes vivants. Un grand progrès fut accompli avec la découverte des couplages énergétiques et des mécanismes de transfert et de distribution de l'énergie, qui permit de découvrir les voies de synthèse des constituants cellulaires. Nous assistons à présent à un nouveau bond en avant avec l'étude de la structure fine des protéines et des acides nucléiques et la découverte de la nature de l'information structurale, de son transfert et de son expression. Un nouveau type de réaction vient de faire son apparition en biochimie: l'assemblage des éléments d'une macromolécule sur un modèle. Bien que l'existence d'organisateurs moléculaires fut considérée depuis longtemps comme une possibilité, voire comme une nécessité logique, les biochimistes hésitèrent longtemps à entreprendre sé-

(*) Rédigée par H. CHANTRENNE.

rieusement leur étude. Il est maintenant établi que le DNA, les protéines et certaines classes au moins de RNA prennent naissance sur des organisateurs moléculaires. Les biochimistes pourraient dire, en paraphrasant un aphorisme célèbre: « toute molécule riche en information vient d'une molécule riche en information ».

L'hypothèse de base selon laquelle l'information génétique réside dans l'arrangement des quatre bases le long de la chaîne nucléique est à présent si bien étayée que même les plus sceptiques doivent l'accepter, même s'ils protestent, non sans raisons peut-être, contre la pensée schématique de collègues trop enthousiastes. Les théories actuelles sont sans doute un peu trop schématiques, il pourrait y avoir d'autres formes d'information à côté de celle qui réside dans la séquence des nucléotides (bien que ceci reste à démontrer); il subsiste certainement des difficultés mais bien que l'on insiste en général sur les brillants succès, les difficultés ne sont pas oubliées, elles sont au contraire la principale préoccupation des chercheurs.

Le premier processus de synthèse sur modèle pour lequel des preuves expérimentales directes furent obtenues, la synthèse du DNA, n'est pas encore compris dans tous ses détails. La séquence est-elle simplement déterminée par l'appariement des bases, comme nous le pensons volontiers, ou bien des influences ou des restrictions sont-elles apportées par les nucléotides voisins? La distribution des pyrimidines dans le DNA contenant de la bromo-uracile mérite de nouvelles recherches, car les résultats rapportés par CHARGAFF remettent en question certaines idées de base sur la nature de l'information.

Le travail de MESELSON est une magnifique illustration de ce que peut donner la combinaison de quelques-unes des méthodes d'analyse les plus puissantes dont nous disposions à présent: analyse génétique, traceures, équilibre dans un gradient de densité. Il n'y a guère plus d'un an, l'opinion générale était que la recombinaison à l'intérieur d'un locus ne reposait pas sur une rupture du matériel génétique suivie de réas-

sociation mais qu'elle se produisait au cours de la reduplication du DNA par un processus appelé « copy choice », dans lequel la fibre naissante d'ADN est copiée d'abord sur une molécule et puis sur une autre. La preuve que des échanges et des réassociations se produisent réellement et que le DNA ne doit pas nécessairement se redupliquer pour se recombiner sont des faits fondamentaux. Sans doute cherchera-t-on bientôt l'enzyme de recombinaison.

Il y a longtemps que des chercheurs qui étudiaient des cellules animales ou végétales étaient arrivés à la conclusion que l'information génétique doit être transmise à des RNA avant de s'exprimer; mais on supposait généralement qu'il s'agissait de RNA ribosomial essentiellement stable. Chez les bactéries, cette idée était difficile à concilier avec la cinétique de la synthèse des enzymes. La grande nouveauté du RNA messager bactérien, tel qu'il apparaît dans la brillante théorie de JACOB, c'est qu'il a la vie courte, qu'il est détruit après avoir porté son message au système qui fait les protéines. Des fractions de RNA à vie brève ont été trouvées chez les bactéries et la preuve la plus convaincante que de tels RNA portent l'information génétique est fournie par les expériences de SPIEGELMAN. Le RNA à vie brève, qui peut se lier aux ribosomes dans des conditions adéquates, peut s'associer spécifiquement au DNA de la souche dont il provient. L'hybridation spécifique obtenue en refroidissant un mélange de DNA et de RNA à vie brève indique que ce RNA est complémentaire du DNA, qu'il contient donc de l'information génétique. L'hybridation de RNA avec le DNA spécifique permettra sans doute d'isoler le messager d'une protéine définie, une possibilité qu'on n'entrevoyait guère il y a seulement quelques mois.

La théorie de JACOB et MONOD permet de formuler clairement bon nombre de nouveaux problèmes. Du seul point de vue biochimique, une foule de questions viennent immédiatement à l'esprit: comment le messager s'attache-t-il au ribosome, comment est-il dégradé, combien de chaînes polypeptidi-

ques une molécule de messager organise-t-elle avant d'être inactivée, quelle est la fonction du RNA ribosomial et des protéines ribosomiales? Quelle est l'origine du RNA ribosomial? Tous les ribosomes d'une bactérie sont-ils identiques ou bien manifestent-ils quelque degré de spécificité? Comment la synthèse des ribosomes est-elle contrôlée? Quels sont les symboles chimiques qui marquent le début ou la fin d'un cistron, d'un opéron, d'un opérateur? En quoi le produit d'un gène régulateur diffère-t-il de celui d'un gène structural? Quelle est la nature chimique du répresseur? A quel niveau l'inducteur agit-il? Le mécanisme régulateur découvert chez les bactéries existe-t-il aussi chez les organismes supérieurs? Dans quelle mesure peut-il rendre compte de la différentiation? La différentiation comporte-t-elle un contrôle chromosomial d'un type inconnu chez les bactéries, etc... Bien d'autres questions apparaissent du côté génétique. De toute évidence, il reste beaucoup à faire et les années qui viennent promettent une belle récolte.

Depuis leur découverte, les virus et tout particulièrement le virus de la mosaïque du tabac furent une source de connaissances fondamentales sur les macromolécules biologiques. Les virus sont des parasites moléculaires qui semblent avoir atteint le degré ultime de simplification, puisqu'ils sont réduits à un RNA porteur d'information, enfermé dans un écrin protéique. Les tourments infligés par les expérimentateurs à l'acide nucléique des virus fournissent des données très directes sur la nature de l'information et sur le code et ils réservent encore bien des découvertes. De nouvelles questions se posent: l'acide nucléique des virus régit-il la synthèse de protéines autres que celles de l'enveloppe? Comment expliquer l'apparition de mutants dans lesquels de nombreux remplacements d'acides aminés se sont produits? Comment le RNA viral se reproduit-il? Le DNA est-il impliqué dans cette affaire? Les chimistes rêvent déjà de fabriquer un RNA, de le voir se reproduire comme celui d'un virus et diriger la synthèse de protéines. Les expériences de SCHRAMM font penser que fabriquer du RNA n'est peut-être pas

une tâche aussi formidable que les chimistes organiciens le croient. Elles permettent d'espérer que des raccourcis pourront être trouvés, et de penser que des réactions chimiques simples, qui seraient régies ou tout au moins influencées par des modèles, donneront peut- être la solution rêvée.

La manifestation la plus immédiate de l'information génétique est l'apparition d'un polypeptide spécifique ou d'une protéine parfaite. La connaisance du mécanisme de la synthèse des protéines progresse rapidement, LIPMANN nous en a donné une belle illustration en présentant, après un tableau d'ensemble, les résultats les plus récents. Selon la théorie actuelle, les acides aminés réagissent avec l'ATP en formant des anhydrides mixtes; les groupes carboxyle ainsi activés sont alors transférés à l'adénosine terminale des RNA de transfert, qui sont des accepteurs spécifiques des différents acides aminés. Les RNA de transfert ont été partiellement séparés l'un de l'autre par électrophorèse, distribution à contre-courant ou chromatographie. L'ingénieuse méthode de purification et d'isolement présentée par ZAMECNIK paraît supérieure à celles que l'on connaissait jusqu'ici. Depuis que l'on dispose de RNA de transfert purifié, il est possible d'étudier plus simplement la synthèse des protéines in vitro et de concentrer les efforts sur le mécanisme de l'organisation de la protéine sur le modèle. Lorsque les acides aminés sont sur les ARN de transfert, le problème énergétique est en effet résolu, il reste à observer les acides aminés qui s'arrangent à la surface du ribosome. Le ribosome du réticulocyte est particulièrement intéressant car il contient l'information nécessaire à la synthèse de l'hémoglobine, l'une des protéines le mieux connues (voyez la contribution de PERUTZ). Le messager génétique reste longtemps attaché à ces ribosomes et le message peut être déchiffré et interprété correctement par de l'ARN de transfert bactérien, bien que quelque enzyme spécifique du ribosome puisse être requis. Si le RNA de transfert d'une bactérie comprend le message d'un réticulocyte de lapin, c'est que la code génétique est universel.

Les résultats récents de NIRENBERG, que LIPMAN a mentionnés brièvement dans son rapport, permettent de penser que le code génétique sera déchiffré dans les quelques années qui viennent.

Nul ne sait comment les RNA de transfert trouvent leur place sur le modèle. Sans doute l'appariement des bases intervient-il, mais il y a plus: par exemple, GTP est nécessaire et sa fonction est encore complètement obscure. La puromycine, qui se comporte comme un analogue du RNA de transfert et se détache du modèle en entraînant avec elle un morceau de polypeptide inachevé, sera sans doute un précieux outil pour l'étude des dernières étapes de la synthèse des protéines. Les analogues de purine et de pyrimidines perturbent aussi ces dernières étapes. Une meilleure compréhension du mécanisme par lequel ils provoquent des erreurs de lecture de l'information serait bien souhaitable. Les systèmes acellulaires utilisés par NIRENBERG ou par LIPMANN donneront sans doute la solution.

Lorsque tous les problèmes de transfert, de transcription, de lecture et de traduction de l'information génétique seront résolus, il restera à comprendre comment elle s'exprime en une activité enzymatique définie. Il est clair que l'activité des enzymes et la fonction physiologique des protéines en général dépendent des structures secondaires et tertiaires des protéines. Nous nous demandons à présent si le reploiement des chaînes polypeptidiques en une structure unique et leur association en enzymes actifs poseront de nouveaux problèmes d'information et de régulation ou s'ils résultent automatiquement de l'arrangement des acides aminés. L'hypothèse la plus simple est certainement que la structure est déterminée complètement par la séquence des acides aminés. Les études sur la ribonucléase qu'ANFINSEN nous a présentées vérifient cette dernière hypothèse puisqu'elles montrent que certains polypeptides déroulés peuvent retrouver spontanément le reploiement correct. Des arguments dans la même sens furent présentés par FRAENKEL-CONRAT pour la protéine du virus; la comparaison des structures de l'hémoglobine et de la myoglobine (PERUTZ et KENDREW)

sont aussi très suggestives: nous voyons que deux molécules qui accomplissent à peu près la même fonction ont une structure tertiaire très semblable, bien que leurs compositions en acides aminés soient fort différentes. En y regardant de plus près cependant, nous voyons que quelques acides aminés se retrouvent aux mêmes positions dans les deux chaînes de l'hémoglobine et dans la myoglobine et qu'ils occupent des positions clés qui déterminent le reploiement de la chaîne ou bien qui assurent les interactions entre certaines régions des chaînes polypeptidiques et le groupe hématinique. Mais des variations considérables sont permises dans le reste de la structure. La comparaison de cytochromes de différentes espèces conduit à des conclusions semblables (TUPPY). Cela nous donne un premier aperçu des bases moléculaires probables de l'évolution. On peut en effet, dès maintenant, esquisser en termes moléculaires un mécanisme darwinien: une mutation qui entraîne le remplacement d'un acide aminé par un autre pourra changer le mode de reploiement et faire apparaître un meilleur ou un plus mauvais enzyme; la sélection assurera la perpétuation des enzymes les plus réussis.

Nos connaissances sur le mécanisme des réactions enzymatiques sont encore fort limitées, mais des études comme celles qui furent présentées par THEORELL montrent qu'elles progressent et que nous pourrons peut-être dire bientôt que nous comprenons vraiment comment un enzyme fonctionne. Les chimistes ont le sentiment que la fabrication d'un polypeptide doué d'activité enzymatique est presque à leur portée. Certains essaieront d'y parvenir directement; d'autres prépareront un RNA et feront fabriquer la protéine par des enzymes selon le plan qu'ils auront fourni.

Enfin, l'analyse du contrôle génétique donne une base solide à de nouvelles recherches sur la différentiation et la morphogénèse. Dans les années qui viennent, les vieux problèmes chers aux embroyologistes reparaîtront, mais ils seront formulés en termes de biologie moléculaire. On peut penser qu'après

des tâtonnements et des efforts dans des voies sans issue, quelques données essentielles apparaîtront et que les bases moléculaires de la différentiation apparaîtront aussi clairement que celles de l'hérédité. Nous avons vu dans le rapport de RANZI des exemples de l'extrême complexité du processus de morphogenèse et de la façon dont l'expérimentateur peut l'influencer. La littérature embryologique abonde en observations qui attendent une interprétation claire. Nous sommes mieux armés que jamais pour y parvenir.

Tous les Participants sont tombés d'accord, à l'unanimité, pour charger les Collègues CHANTRENNE et KENDREW de rédiger les conclusions auxquelles on est arrivé pendant la Semaine d'Etude et imprimées ci dessus.

Ces conclusions reflètent les discussions qui ont eu lieu au cours de la Semaine et se trouvent en accord avec ce qui a été établi par tous les Participants pendant leur réunion de jeudi 26 octobre 1961 dans la paix et la quiétude des Jardins Pontificaux de Castelgandolfo.

ANFINSEN, CHANTRENNE, CHARGAFF, DEBYE, DE HEVESY, FRANKEL-CONRAT, GIACOMELLO, JACOB, KATCHALSKI, KENDREW, LINDQVIST, LIPMANN, LIQUORI, MESELSON, MIZUSHIMA, PERUTZ, PUTZEYS, RANZI, RICH, ROSSI-FANELLI, RUBIO-HUERTOS, SCHRAMM, SILIPRANDI, SPIEGELMAN, THEORELL, TISELIUS, TUPPY, ZAMECNICK.

INDEX

TRAVAUX SCIENTIFIQUES