*AMINO ACID SEQUENCE SELECTION IN PROTEIN SYNTHESIS**

By Herbert Jehle

PHYSICS DEPARTMENT, GEORGE WASHINGTON UNIVERSITY† AND UNIVERSITY OF NEBRASKA

Communicated by Paul Doty, July 28, 1959

A living organism is able to synthesize an enormous variety of particular macromolecules—macromolecules of a type which that organism already contains—out of a great variety of different material ("food" in a wider sense). From the point of view of information theory the explanation of this capability of a living organism will be very much simplified if that "food" is considered as being chopped up and then reassembled by a process of replica formation.

We consider the sequence selection in protein synthesis first, before getting into the problem of RNA synthesis, because the assembly of the highly complex and diversified amino acid sequences in proteins calls for an explanation in terms of a general simple scheme; if one can show the possibility of a protein replication scheme which satisfies the physical, chemical, and structural aspects, such a scheme will have evident advantages over other schemes which only too often are overloaded with *ad hoc* postulates.

The selection of a particular sequence of amino acids is one of the main problems in understanding protein synthesis. In the present paper[1] this problem is reduced to the problem of replica formation: The essential information content necessary for the formation of a (filial) protein chain is assumed to be stored in a parental protein chain which is firmly attached to an RNA chain. We shall try to understand this sequence formation as a single step phenomenon rather than as a process involving the subsequent use of hundreds of specific enzymes, synthesizing the chain stepwise, amino acid by amino acid.

In protein synthesis we have to consider all the variety of forces usually involved in biochemical reactions: covalent bonds, ionic bonds, hydrogen bonds; furthermore there is the electrostatic interaction between charged or polar molecules which is strongly modified by gegen-ions from the surrounding ionic medium, and then there are London-van der Waals forces between polarizable molecules.

The latter "polarizability forces" are due to fluctuations of charge distributions in the molecules. Charge fluctuations can be caused by the quantum mechanical zero point motion due to the uncertainty principle. Charge fluctuations may also be due to thermal motion if excited electronic states are in thermal reach. A somewhat different type of thermal charge fluctuations is the fluctuation of proton distributions over the surface of molecules, investigated by Kirkwood and Shumaker. All these charge fluctuation forces have been shown [1, 2] to contribute to an association of like molecules as nearest neighbors which is energetically more favorable than association of unlike ones by an amount of the order of magnitude of a few kT. This property has the important significance that it may account for the selection of the right set of amino acids from the surrounding medium, and for their juxtaposition to the corresponding amino acids of a parental chain.[2] In this paper we shall propose several schemes in which the information about a filial amino acid sequence is stored in an (unfolded) parental protein chain; or, in a separate analysis, one may consider the information stored in a sequence of individual

parental amino acids, independently attached to the RNA chain. In either case the amino acid sequence constitutes the essential part of the "template." The important role of the nucleic acids shall be described later in this paper. In the following schemes we try to leave the task of sequence selection entirely to the specificity property of the charge fluctuation forces without taking recourse to complementarity considerations[3] nor to a multitude of specific enzymes postulated *ad hoc* to accomplish selection of the right sequence.

We do not assume that these filial amino acids (activated amino acids) from the medium will get attracted by their like counterparts in the parental protein at long range from far away; it is Brownian motion which is the agent which shuffles these molecules around rapidly. On the microscopic scale Brownian motion is an extremely rapid phenomenon as the time needed to accomplish a given net displacement is proportional to the square of that displacement. Once filial amino acids come close to the corresponding amino acids in the parental chain, they will be retained there. The selection of the right kind of sequence is achieved in this way.

Some 20 years ago, Hamaker[4] recognized the significance of the equilibrium between static electrostatic repulsions between molecules (in particular between identical molecules) modified by gegen ions (depending on the ionic concentrations in the medium) and the attractive (specific) London-Eisenchitz-Wang (van der Waals) forces. A change of the ionic concentrations in the medium is thus capable of making the nonspecific repulsion predominate over the specific London-van der Waals attraction. Such a change provides for a mechanism for regulating over-all motion of molecules, a motion which might be reversible.

In this paper the question is raised as to how these various forces fit into the process of sequence selection. We want to illustrate the application of the property of specificity of charge fluctuation forces by giving several model schemes. It is evident that the interacting molecules will have to be sterically compatible. In particular, the monomer repeat distances of the nucleic acid chain and of the parental protein chain attached to it should be compatible. It is furthermore necessary to keep in mind that the steric arrangements between filial amino acids which are supposed to be joined by peptide bonds should be such that the peptide bond forming loci would be very close to each other. We definitely want to avoid schemes in which the filial amino acids are many Angstroms apart. The specific interaction due to charge fluctuation forces was discussed in parts I and II. It involves the mutual orientation of the interacting molecules which can be characterized as parallel orientation as regards the z direction (the direction of the line connecting the centers of the two interacting molecules cf. Fig. 3) and antiparallel in the x and y directions. The pictures show this orientation between filial and parental amino acid side chains at several instances. The pictures of space filling models have the purpose of illustrating the spatial arrangements and of showing their compatibility with the steric requirements and the requirements of closeness of interacting loci at the time of peptide bond formation. We have, of course, also to see that the energy conditions for peptide bond formation are properly taken care of. The backbones of parental and of filial chains are parallel oriented. Subsequently, the filial protein chain may be made to peel off from the nucleoprotein template through changes in the ionic constitution of the medium.

As regards the biochemical evidence, the filial amino acids which are to be used

for the synthesis of a new protein, are available in activated form. It is to be emphasized that in our schemes it is quite irrelevant what kind of activation actually is involved. The important steps in the formation of replica proteins are practically the same whether the amino acid activation consists in an ester linkage of an amino acid to the first ribose of a soluble RNA or in a phosphate bond of an amino acid to an AMP, or whether activated small peptides may be the units from which the filial protein chain is assembled. In view of the fact that the detailed steps of the activation mechanism are not yet all settled, we can best perform the molecular model building by using the simplest prototype of activation: we represent an activated amino acid by an amino acid attached, through its carbonyl group, to the phosphate of an AMP, or the monophosphate of another purine or pyrimidine base, i.e., a ribonucleotide.

We further make the usual assumption that the presence of RNA, or DNA, or other substitutes is required before protein synthesis may occur. We discuss RNA schemes; the others might perhaps be similar. We shall use the name RNA when referring to any sequence of nucleotides. In building models, we sometimes use a polyadenylic acid chain or a polyuridylic acid chain, or we use a genuine mixture of nucleotides. The form which the RNA chain or chains will assume and the way of attachment of the parental protein chain (or the parental individual amino acids) to the RNA chain will be discussed along with each of the proposed schemes described in this paper.

This brings us to the concept of a "template." We talk about "templates" even though that word has often been used to designate steric complementarity, complementarity of van der Waals contacts, complementarity of electrostatic patterns of charge distribution, complementarity of hydrogen bond systems—or all of them together. In our case we understand by "template" a molecule or a complex (which contains the proper sequence of parental amino acids, individually or as a protein chain) on which the filial sequence of amino acids is thereupon assembled by specific charge fluctuation forces. The functioning of this template requires the parental amino acid sequence to be arranged so that it forms some kind of a linear array of exposed parental amino acid side chains to which the filial amino acids have access. This arrangement may be achieved by an attachment of the parental sequence to a nucleic acid chain or helix which has the function of a backbone to hold the amino acid sequence, or protein chain. The specific function of the nucleic acid will be discussed below. It may be well to remember that highly charged groups like the phosphates of the nucleotides are probably well neutralized by an additional chain of basic amino acids (as found in microsomal nucleoproteins), by Mg^{++}, or other small ions. The charges on the nucleotides and on the parental and filial amino acids are therefore not expected to completely dominate the mode of attachment in question.

The problem of protein synthesis is really a twofold one: the problem of formation of a template and the problem of formation of a replica protein. In discussing the possible models we concentrate on the latter of these two processes because the schemes of formation of a template are likely to be more dependent on detailed experimental results (for which we still have to wait) than are the schemes of formation of filial protein chains.

In discussing the formation of a replica protein we give, in this paper, detailed

attention only to schemes in which the template is an extended parental protein chain attached to a single strand extended RNA chain by hydrogen bonds (denoted by H). There might be similar schemes involving double strand RNA. There may also be schemes in which the template is a sequence of individual amino acids, attached to the extended (single or double strand) RNA by covalent bonds (denoted by C).[5] Nature provides for an exceedingly accurate conservation of information. This seems to imply that the information content either be preserved by virtue of the parental protein chain remaining tightly hydrogen bonded together with the RNA (in schemes H), or be preserved by virtue of the individual parental amino acids being covalently bound to the RNA chain (in schemes C). In any case we would like to regard nucleoproteins rather than nucleic acids as carriers of genetic information.

Considering the various possibilities of template formation from an RNA chain and a parental protein chain or a parental sequence of individual amino acids, and considering only those sections of the RNA chain which carry parental material, the question evidently arises: what fraction of the set of parental amino acids is attached to what fraction of the set of nucleotides? We will discuss in detail only schemes in which every monomer of the protein chain is attached to the RNA chain. On the one hand that includes the case where the ratio of the number of amino acids to the number of nucleotides in the template is one to one (all the schemes listed in reference (5)). This 1:1 ratio for the template, formed by H bonds between RNA and peptide chain, is supported by the experiments of E. T. Bolton and his co-workers[6] and other groups. On the other hand, every amino acid might perhaps be attached to every nucleotide pair in the wide groove of a Watson-Crick-Rich RNA double helix ($H^{+}{}_{B}2p$), a double helix assumed to be similar to the Watson-Crick-Wilkins DNA helix.

These attachments between protein and RNA chains imply fairly stringent condition as regards the repeat distance of attachment sites on the RNA: they have to match the repeat distance along the protein chain (schemes H). The corresponding conditions in schemes C would not be less stringent: in either case the arrangement of the parental amino acids has to be such that the filial amino acids have a chance to become assembled at the proper peptide repeat distance from each other, i.e., some 3.4 Angstrom apart. Otherwise peptide bond formation in the filial chain could only occur simultaneously with the filial chain peeling off and that seems to be most unlikely to result in a complete filial protein chain.

These schemes may help the biochemist to select the actual process which is at work at protein synthesis. They also pose a number of experimental problems, the solution of which will decide in favor of one or against another of the hypotheses made in regard to protein synthesis theories.

Protein Chain Hydrogen Bonded to Riboses of Single Strand, Stretched-out RNA ($H^{+}{}_{R}1a$).[5]—In this scheme we assume the presence of a single chain RNA molecule whose bases are stretched out alternately from the midline of the chain like the oars of a racing boat. With models it is possible to build an RNA chain with a monomer repeat distance equal to that of an extended protein chain and having the $C_2'OH$ groups on the riboses accessible all from one side as shown in upper Figure 1 (top view). We also assume a protein chain in unfolded form with its carbonyl groups pointing alternately to somewhat opposite sides of the chain's

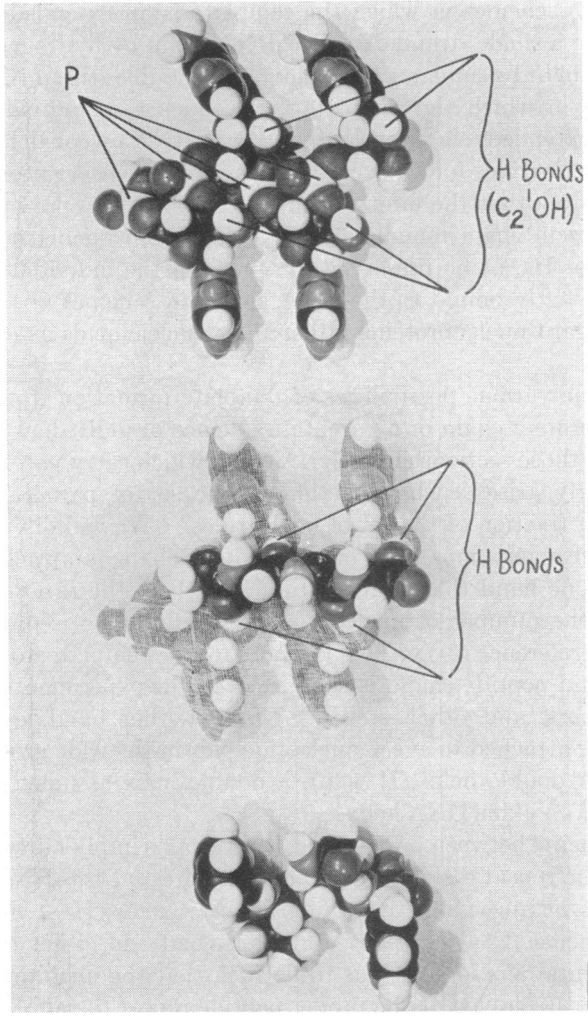

FIG. 1.—All three pictures are "top views" showing the ingredients for the template for H^+_R1a. The upper picture shows the bare four monomer RNA chain with the OH groups on the C'_2 of the four riboses clearly exposed. The middle picture shows this RNA under a veil; over the veil, a four monomer protein backbone is hydrogen bonded to the four C'_2OH of the RNA. The veil makes it possible to distinguish the protein backbone from the RNA. The lower picture shows the protein chain in the same view as the protein backbone of the middle picture; the backbone is mostly hidden behind the side chains. From left to right these are phenylalanine$^+$, valine, glutamine, tryptophane$^-$. The same sequence is used in Figures 2, 3, 4, 8, 9, 10.

C is black; H is white, an almost complete sphere, except in an H bond where it is about a cylindrical disk; P is beige, appearing white, tetrahedral: O is red, appearing as a darker gray; double bond O is shaped like a flattened sphere with two indentations, single bond O looks more like a sector of a sphere; N is blue, appearing as a lighter gray.

midline as shown in middle Figure 1 (repeat distance along the midline axis is 3.4 Å). A short but fairly bulky section of the glucagon chain, i.e., tryptophane$^-$, glutamine, valine, and phenylalanine$^+$, is used to illustrate this protein chain, lower Figure 1. The attachment of the protein chain to the nucleic acid chain is through hydrogen bonds between the oxygens of the carbonyl groups of the peptide backbone

and the hydrogens of the OH groups on the C_2' of the riboses of the RNA. Such an attachment would not be possible with DNA. The negative charges on the phosphates may be compensated by Mg^{++} ions and by some basic protein side chains which can get at the phosphates. The RNA with attached (parental) protein chain would act as a template for the synthesis of a filial protein chain. Figure 2 shows this template, its RNA part again being covered by a veil (side view).

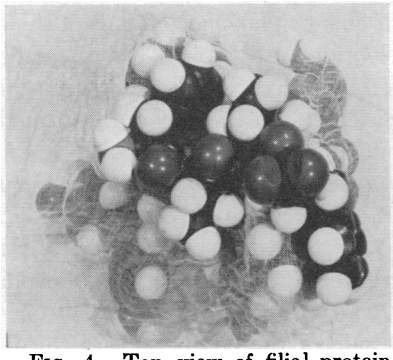

FIG. 4.—Top view of filial protein chain lying on top of the template which is altogether covered by a veil, $H^+{}_R1a$. Backbone of filial chain (four oxygens) is visible on top. Relative position of filial to parental protein chain is like that of Figure 10 which is an end view of the chains, showing the filial and parental tryptophanes after the filial chain peeled off.

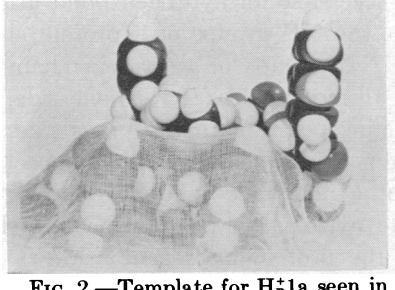

FIG. 2.—Template for H^+_R1a seen in side view. RNA is under the veil, the protein chain of Figure 1 is over it, again tryptophane$^-$ on the right.

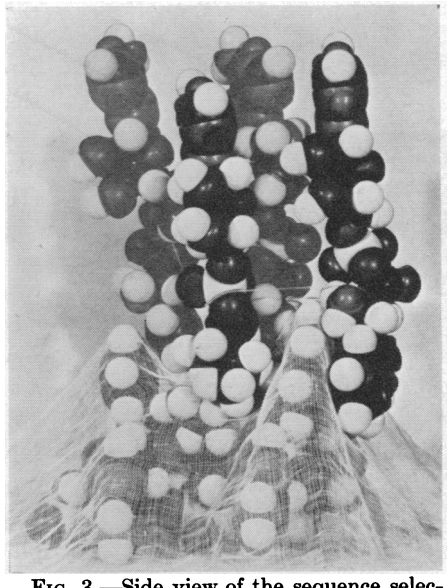

FIG. 3.—Side view of the sequence selection of filial amino acids in $H^+{}_R1a$. The template of Figure 2 is entirely covered by a veil; four activated amino acids (AA-AMP), i.e., filial amino acids, lie, in proper orientation, side by side next to the respectively identical parental amino acids on the template, held on by specific charge fluctuation forces, after Brownian motion has brought a variety of amino acids close by. The rear row (phenylalanine and glutamine) of the activated amino acids are represented shaded to distinguish them from the others.

We assume activated amino acids to be available in the medium, in the case of the illustration simply amino acids phosphate bonded to AMP. Four of those activated amino acids are seen in Figure 3. They are retained next to the corresponding parental amino acid side chains, in proper orientation, after Brownian motion has thoroughly shuffled them around. In this Figure 3 (again, a side view), the entire template of Figure 2 is covered by one veil; the activated amino acids are bare.

Brownian motion also brings an amino group of one activated amino acid frequently near to the carbonyl group of the adjacent activated amino acid. The polarity of the filial chain is parallel to that of the parental chain. At this instant

the bonds connecting the amino acids to the ribonucleotides may be broken with simultaneous formation of peptide bonds between the filial amino acids. When the broken-off ribonucleotides have drifted away, the filial protein chain lies attached, as shown in Figure 4, to the parental protein chain. A later change in ionic constitution of the medium permits the filial chain to peel off, as in Figure 10.

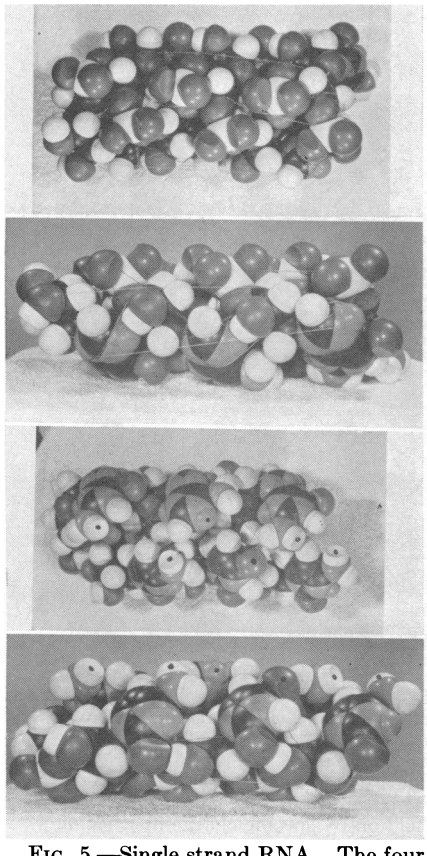

FIG. 5.—Single strand RNA. The four views are related to each other as four views in an engineering drawing. The internal hydrogen bonding is where the disk shaped H bond hydrogens connect between riboses and bases. Figure 6 whose top corresponds to the right-hand side of Figure 5 shows the details. The monomer repeat distance is about 3.3 Å. The sites where the protein chain of Figure 7 is hydrogen bonded to this RNA chain are marked by black dots.

One has also to envisage the possibility of other processes following Figure 3. When the high energy phosphate bonds of the AA-ribonucleotides are broken, there might be the possibility that, besides the formation of a protein chain, an RNA chain is formed, or that a ribonucleoprotein results.

Protein Chain Hydrogen Bonded to Phosphates of Single Strand, Stretched-out RNA ($H^{-}{}_{P}1a$).—One can arrange the RNA in a slightly different fashion, with its phosphates in quite an exposed position, all on one (upper) side of the chain, so that the RNA has a phosphate backbone resembling the row of heads of the oarsmen of a racing boat and the bases like the oars which in this case are halfway pulled in. Such an RNA (or DNA) single chain may attain a monomer repeat distance as short as of the order of magnitude 3.2 Å and may have all its phosphates accessible for hydrogen bond attachment to the NH groups of a protein chain. (If the NH groups of a protein chain are turned all more or less toward one side, the side chains pointing to the other side, one has a repeat distance of 3.2Å to expect for the peptide monomers.) In this case $H^{-}{}_{P}1a$ it is the oxygens on the phosphates which accept the protein hydrogens: Mg^{++} and some basic proteins might again neutralize negative charges on the phosphates. DNA might function according to the same scheme.

Protein Chain Hydrogen Bonded to Bases of Single Strand RNA with Alternate Base Arrangement ($H^{\pm}{}_{B}1a$).—The single strand RNA discussed in the first instance ($H^{+}{}_{R}1a$) had its bases outstretched, alternately from one and the other side of the phosphate backbone. The next scheme ($H^{-}{}_{P}1a$) brought the bases closer to the backbone. The present scheme ($H^{\pm}{}_{B}1a$) brings the bases very close together. Their closest packing is seen in the illustrations (Fig. 5). The phosphates again form a zigzag backbone with indentations between them which make easy access to the side chains of some (nonspecific) basic protein

and to Mg^{++} ions which compensate the negative charges on the phosphates. This RNA single chain is very compact and has a monomer repeat distance (measured along the midline) of about 3.3 Å. The $C_2'OH$ (which evidently exists only in RNA, not in DNA!) of every ribose forms a hydrogen bond to the structurally adjacent base[7] which is the second-to-next base along the sequence of nucleotides (Fig. 6). The H bond acceptors are the N_7 in case of purines and the C_2O in case of pyrimidines. On the side of the bases there are hydrogen bonding sites, some H donors, some H acceptors to which a protein backbone may be attached (black dot

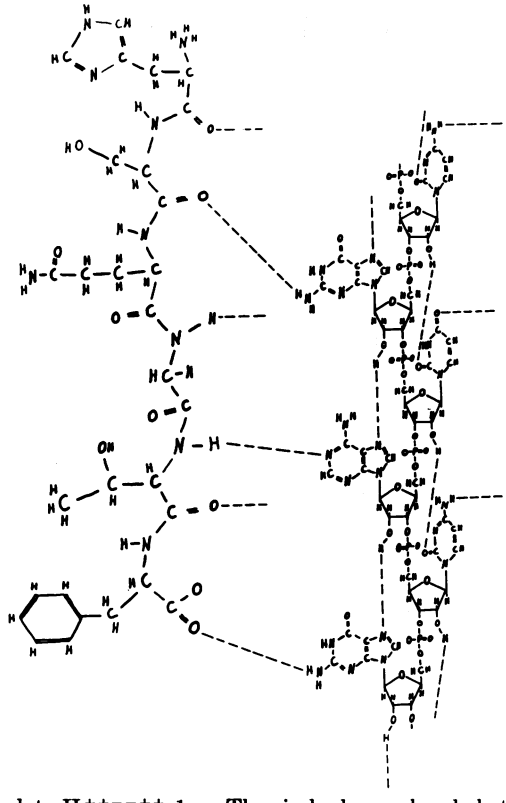

FIG. 6.—Template $H^{++--++}{}_B1a$. The six hydrogen bonds between protein and nucleic acid are marked by dashes, three of which are interrupted in order not to disturb the RNA formula. The protein chain consists of the first six amino acids of glucagon, i.e., histidine$^+$, serine, glutamine, glycine, threonine, phenylalanine.$^-$

markings in Fig. 5 and 7). But such attachment is conditional on an adequate sequence of nucleotides, in particular when the protein's side chains are taken into account. These side chains not only influence the possible shapes of the protein backbone: their bulkiness also implies direct compatibility conditions with the available space on the rough terrain which the bases of the RNA have to offer. The degree of specificity between this RNA and the parental protein compatible with it will be discussed in a later note.

This compatibility condition between RNA chain and parental protein chain is not of such a kind that three adjacent nucleotides would determine one amino acid

(Gamow's overlapping code) but something related to that coding idea: the compatibility condition is much more relaxed than Gamow's. But it implies that if a foreign nucleic acid is injected into a cell, it will in general no longer be possible to attach the cell protein to that foreign nucleic acid, and a rearrangement of peptide chains into a new parental protein chain might come about which then produces that same new type filial protein. The phenomenon of transforming principle might in this way be understood to operate. According to that view the new protein produced will not only be dependent on the kind of nucleic acid injected but also on the type of protein present in the original cell. (This of course presupposes that the injected transforming principle does not carry any information-containing nucleoprotein with it.)

In this connection, where we have a stereospecific attachment between nucleic acid and protein, we are to be reminded that the degree of conservation of information lies in the degree of stability of that tightly hydrogen bonded nucleoprotein structure, i.e., the template (and in its accurate replication which we have so far

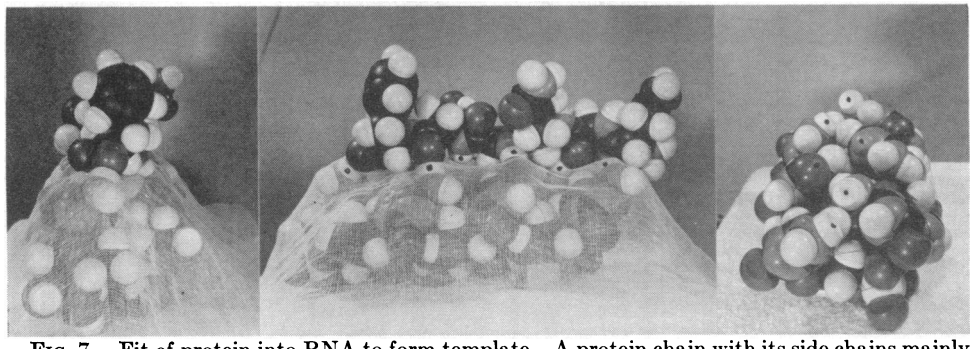

FIG. 7.—Fit of protein into RNA to form template. A protein chain with its side chains mainly pointing in the upward direction, has an average amino acid repeat distance of 3.2 to 3.5 A; it can be shortened, but not lengthened. The RNA structure is very compact, and can, given the internal H bonding, and in view of its compactness, not be shortened as long as it is assumed to be a straight rod. The fit, in the present example, is perfect; in the general case both protein and nucleic acid have about 3.3 A as their average monomer repeat distance. The fit is conditional on the detailed sequence in the two kinds of molecules.

sidetracked). We are also to be reminded that mistakes in synthesis of filial protein chains are of a random character and therefore much less critical than a change in a nucleoprotein which is a highly permanent structure.

Protein Chain Hydrogen Bonded to Bases of Single Strand, Stacked-up RNA ($H^{+}{}_{B}1p$ *and* $H^{-}{}_{B}1p$).—To complete the survey, two schemes should be mentioned. In these schemes the RNA chain is single strand, and assumed to be of the general form of a half of a Watson-Crick-Rich double helix. The function of the RNA is again to provide for a structure on which the stretched-out parental protein chain may be held by hydrogen bonds. A variety of attachments might be possible between either the CO or the NH groups of the protein chain to several sites on the bases of the RNA. The molecule model $H^{+}{}_{B}1p$ shows H bond attachments of the CO groups of the peptides to the NH_2 groups of the adenine bases: the ingredients on Figure 8, the template on Figure 9, and the template with the filial protein chain nearby, after it has been peeled off (Figure 10). The scheme $H^{-}{}_{B}1p$ is analogous: the protein's NH attach to the uracil's bases to their O atoms.

The negative charges on the phosphates are again assumed to be compensated. Other attachments of a protein chain to this nucleic acid chain, monomer by monomer, fail because of the high repeat distance between riboses or between phosphates. DNA would permit a similar scheme.

Double Attachments of Protein Chains.—It does not seem possible that a peptide chain would be hydrogen bonded to an RNA chain by hydrogen bonds both over the NH and the CO groups of the same monomers. Such an otherwise interesting bonding would imply serious obstruction to amino acid side chains.

Covalent Attachments of Individual Parental Amino Acids to the RNA Chain.—Space does not permit elaborating on those schemes. The more important schemes

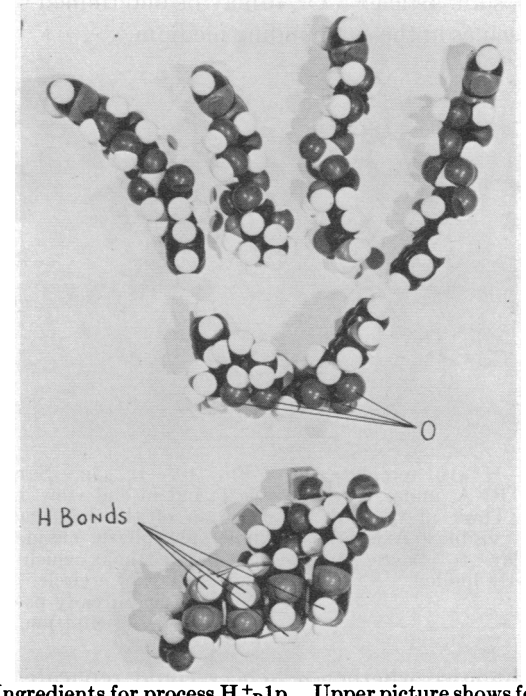

FIG. 8.—Ingredients for process $H^{+}{}_{B}1p$. Upper picture shows four activated filial amino acids; middle picture, the parental protein chain in side view; lower picture, the RNA chain (half of Watson-Crick-Rich double helix) in top view: the axis of this helix is about perpendicular to the four base planes. Bases with H bonds are all visible.

seem to be C_R1a, C_P1a, C_B1a, C_B1p, C_B2p, C_R2p, C_P2p. The latter two would imply attachment, in a zigzag fashion, of one sequence of amino acids in the narrow groove of a Watson-Crick-Rich helix, to all riboses or to all phosphates respectively; the filial protein chain would come to lie along the narrow groove. C_B2p might perhaps be a possibility though not a sterically very clean one, where the amino acids are attached to the bases in the wide groove of a Watson-Crick-Rich helix with the corrected helix parameters given by Wilkins. The first four of these schemes are analogous to the corresponding hydrogen bonded schemes which we have described.

Further Remarks.—It is the purpose of this paper to bring the current biochemical

findings in the field of protein synthesis into proper relation to the physical phenomenon of specificity of charge fluctuation forces so as to account for the right sequence selection of amino acids.

Protein synthesis involves (1) not only formation of a protein chain of proper amino acid sequence, it also involves (2) formation of α helices or sheets, and (3) folding of the helices or sheets into a three dimensional globular or other structure. Only the first of these steps has been discussed in this paper. The second step seems not to raise too difficult problems: an α helix might be easily formed from a chain. As regards the third step, the folding of an α helix into a three dimensional structure may be determined by the sequence of amino acids of the α helix. Or, this folding may be an assembly process of the type described in Figure 5 of our symposium article.[2] Or, it may be determined by other biologically important macromolecules in the surrounding medium.

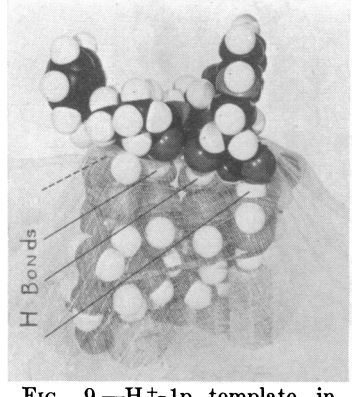

FIG. 9.—$H^{+}_{B}1p$ template in side view, RNA under, protein over veil. Three of the four H bonds are visible. Axis about perpendicular to bases, phosphates mostly hidden.

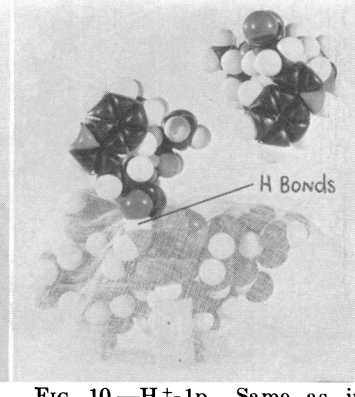

FIG 10.—$H^{+}_{B}1p$. Same as in Fig. 9, but in end view, i.e., in the direction of the RNA axis. The ribose phosphate chain forms a right-handed helix which shows up as a segment of a circle. The filial protein chain, already peeled off, is seen in proper orientation.

The question might arise whether α helices could replicate without unfolding, or not. Suppose a parental α helix has collected filial amino acids from the surrounding medium, in a mantle region surrounding the helix, corresponding to the stage represented by Figure 3. Consecutive amino acids are now, however, very far apart along a helix surrounding the parental helix, and peptide bond formation might come about between amino acids number 1, 5, 9, 13, . . . rather than between number 1, 2, 3, 4, 5, . . . as it should. This clearly does not permit an orderly replication of a protein.

This investigation received important help from many colleagues to whom I would like to express my appreciation and thanks. The work could not have been done without the molecule models for which I am greatly indebted to Dr. Linus Pauling, Dr. Robert B. Corey, and Dr. Richard E. Marsh. At Nebraska I owe thanks to Dr. Robert B. Johnston, Dr. J. R. Matton, Dr. Patricia Weymouth, Miss Audrey Fosbrooke, and Mr. Donald E. McArthur who contributed many ideas in connection with the model building.

* Research supported by the Research Corporation and the National Cancer Institute (National Institutes of Health grant C-3304 BBC), and by the University of Nebraska Research Council.

† Permanent address.

[1] We shall refer to this note as part III. Part I by Yos, Jerrold M., William L. Bade, and Herbert Jehle, these PROCEEDINGS, **43**, 341 (1957); part II by the present author, these PROCEEDINGS **43**, 847 (1957), "Specificity of Intermolecular Forces due to Charge Fluctuations."

[2] Muller, H. J., *Am. Naturalist*, **56**, 32 (1922), *Cold Spring Harbor Symp. Quant. Biol.*, **9**, 290 (1941), and *Proc. Roy. Soc.*, **B 134**, 1 (1947); Haurowitz, F., *Chemistry and Biology of Proteins*, (Acad. Press, 1950), p. 340 ff.; Hamaker, H. C., *Physica*, **4**, 1058 (1937); Jehle, H., J. M. Yos and W. L. Bade, *Phys. Rev.*, **119**, 793 (1958); Yos, J. M., *Phys. Rev.*, **119**, 800 (1958); Yos, J. M., W. L. Bade, and H. Jehle, *in Symposium on Molecular Structure and Biological Specificity*, ed. L. Pauling and H. Itano, (Am. Inst. of Biol. Sciences, 1957), and *Proceedings of the First National Biophysics Conference* (Yale University Press, 1959).

[3] Cf. the remarks in part IV.

[4] Hamaker, H. C., *Rec. Trav. Chim. Pays Bas*, **56**, 3 (1937); *Chem. Weekblad*, **35**, 47 (1938); Overbeek, J. Th. G., in Kruyt, H. R., *Colloid Science*, (Elsevier, 1952) **1**, pp. 276, 277; Derjgauin, B., and L. D. Landau, *Acta Physicochim. USSR*, **14**, 663 (1941); Lifshitz, E. M., *JEPT*, **2**, 73 (1956); Debye, P., and E. Hueckel, *Phys. Z.*, **24**, 185 (1923); Onsager, L., *Chem. Revs.*, **13**, (1933); Klotz, I. M., in *The Proteins*, ed. Neurath and Bailey (Academic Press, 1953), p. 727 ff.; Prigogine, I., and A. Bellemans, *Disc. Faraday Soc., Nonelectrolytes*, **15**, 80 (1953).

[5] $H^{+}_{R}1a$ stands for "protein chain *H*ydrogen bonded to the $C_2'OH$ of the hydrogen donor (+) *R*iboses of a *single* (*1*) strand RNA whose bases are stretched out *a*lternately left and right from the phosphate-ribose backbone, like the oars of a racing boat." C_P2p stands for "individual amino acids *C*ovalently bonded to the *P*hosphates of a *double* (*2*) strand RNA, the bases of each strand being stacked-up *p*arallel as in a Watson-Crick-Wilkins-Rich helix." The schemes to be considered are: $H^{+}_{R}1a$ protein H bonded to H donor riboses of single strand alternate base RNA, $H^{-}_{P}1a$ protein H bonded to H acceptor phosphates of single strand alternate base RNA, $H^{\pm}_{B}1a$ protein H bonded to H donor or acceptor bases of single strand alternate base RNA, $H^{+}_{B}1p$ protein H bonded to H donor bases of single strand parallel base RNA, $H^{-}_{B}1p$ protein H bonded to H acceptor bases of single strand parallel base RNA. The first, third, and fourth of these schemes are represented in the illustrations.

[6] I am indebted to Dr. E. T. Bolton and his colleagues in the Department of Terrestrial Magnetism of the Carnegie Institution, Washington, D. C., for telling me about the not yet published results which they obtained. They have found a ratio of two amino acids per nucleotide in ribonucleoproteins from *E. coli*, and by means of a strong Na^{+} solution they have been splitting off one protein chain, the remainder being constituted in a protein-RNA chain with one amino acid per nucleotide. Subsequent treatment with urea resulted in splitting off the second protein chain from the RNA.

[7] I am indebted to Dr. Chr. Jardetzky who has drawn my attention to this stabilizing hydrogen bond.

---

*PYRIDOXAL KINASE OF HUMAN BRAIN AND ITS INHIBITION BY HYDRAZINE DERIVATIVES**

BY DONALD B. McCORMICK† AND ESMOND E. SNELL

DEPARTMENT OF BIOCHEMISTRY, UNIVERSITY OF CALIFORNIA, BERKELEY

Communicated July 13, 1959

Although the mechanism of conversion of vitamin B_6 to pyridoxal phosphate is not fully known for any tissue,[1] it requires in every case participation of a kinase. Such kinases must be of almost ubiquitous occurrence and, indeed, have been detected in brain,[2] liver,[3] *Streptococcus faecalis*,[4, 5] *Escherichia coli*,[6] and yeast[6]. Despite the