Biophysical Chemistry 3 (1975) 1–20

A MODEL OF MYOGLOBIN SELF-ORGANIZATION

O.B. PTITSYN and A.A. RASHIN
Institute of Protein Research, Academy of Sciences of the USSR, Poustchino, Moscow Region, USSR

Received 12 November 1973
Revised manuscript received 8 August 1974

The self-organization of helical regions of myoglobin into a compact tertiary structure is considered on the basis of the hypothesis on the step-wise mechanism of self-organization of protein molecules. It is assumed that the self-organization begins with the formation of "centers of crystallization" and proceeds with the growth of one such center or by a sequential collapse of two or more grown centers.

Different pathways of self-organization of myoglobin are considered; the most favourable structures corresponding to the greatest number of dehydrated bulky hydrophobic groups and to all the strongly hydrophilic groups exposed to water are selected at every stage of the given pathway and the others are neglected. One of the two most favourable structures obtained in such a way coincides in rough resolution with the native tertiary structure of protein.

1. Introduction

During the last years a conviction is becoming more and more wide-spread that self-organization of the native three-dimensional structure of a globular protein cannot proceed by a random search of this structure among all the possible conformations of the protein chain. Therefore it is assumed [1–4] that self-organization of globular proteins is a directed or at least a partly directed process. This process can be hypothetically divided into three main steps [5]:

1. At the first step fluctuating regions of the secondary structure are formed in the unfolded polypeptide chain and their formation is determined by local interactions, i.e., by interactions of amino acid residues adjacent along the chain (in the first place by hydrogen bonds).

2. At the second step fluctuating regions of the secondary structure collapse, forming a single compact globule, the formation of which is determined by long-range non-specific interactions of side groups with the surrounding medium (water and the hydrophobic core of the globule formed).

3. At the third step this "intermediate" compact structure is re-arranged into a unique native structure of globular protein, the formation of which is determined by specific long-range interactions of spatially close amino acid residues.

Determinacy of the process of self-organization may be accounted for by the fact that fluctuating structures of every region of the chain forming at the early stages of the process direct the formation of more complex and more extent structures at the following stages, the fluctuating structures themselves being at the same time only maintained and not rearranged [2–5]. For this to be so, it is necessary that the interactions within every region of a protein molecule stabilize the same structure of this region which is stabilized by interactions of the regions with each other [3]. Such a concordance of local and long-range interactions represents a necessary condition for the "technological" process of self-organization and therefore must be programmed in the primary structures of all the really existing proteins selected in the course of biological evolution.

Considerable progress has been lately achieved in the development of a theory of the first step of self-organization. At first a semi-empirical theory of fluc-

tuating helical region formation in unfolded protein chains was developed [5–9] and then a molecular theory of local secondary structures of unfolded protein chains [10,11]. However, up to the present no model of the globule self-organization from regions with a fluctuating secondary structure has been suggested permitting a theoretical calculation of the globular protein three-dimensional structure. An attempt to develop such a model is just the purpose of this paper.

In the simplest case of highly helical proteins of the globin type, the first step of the self-organization process is reduced to formation of fluctuating helical regions and the second step to the assembly of a three-dimensional globule from them. Namely this simplest case is considered in concrete application of the general model of the second step of self-organization.

2. Scheme of self-organization of regions with a secondary structure into a compact globule

The suggested model of compact globule self-organization from regions with a fluctuating secondary structure includes the following five assumptions:

1. Self-organization of regions with a secondary structure into a compact globule begins with the formation of one or several complexes from neighbouring along the chain fluctuating regions with a secondary structure*.

2. These complexes serve as "centers of crystallization" for a further process which consists in the growth of centers by a sequential joining of regions adjacent along the chain.

3. A single compact globule is formed by the growth of one such center or by the sequential collapse of two or more grown centers.

4. Joining of the new regions at each stage of each pathway of self-organization occurs only to the "most favourable" structures formed at the preceding stage and does not re-arrange but only maintains these structures.

5. The most favourable of all the final structures corresponds to the native structure and the pathway leading to this structure corresponds to the real pathway of self-organization.

The first three assumptions determine the way of possible intermediate and final structure formation. The necessity of postulating the formation and growth of centers of crystallization instead of the *simultaneous* collapse of all the regions with a secondary structure into a single globule is clear from the fact that the probability of pair collision of regions is always far greater than the probability of simultaneous collision of several regions. It is also obvious that the formation and growth of centers of crystallization must proceed by the collapse of regions with a secondary structure *adjacent along the chain*. In fact the collapse of every pair of helices is connected with the formation of a loop from the chain region which joins them, and the probability of the formation of such a loop rapidly decreases with the increase of the number of residues j in the region forming the loop (for the gaussian chain this probability decreases inversely proportional to $j^{3/2}$ and even more rapidly for more rigid chains [12]).

The fourth assumption determines the most probable ("technological") pathways of self-organization and the last assumption determines the real pathway of self-organization and the sought native structure. *The native structure is assumed to be the most favourable of all the final structures formed as the result of the most probable pathways of self-organization.* Thus, using our approach the search for the native structure is not made among all the possible conformations of the protein chain but only among those of them which lie on the most probable (technological) pathways of self-organization.

As known, the stability of the tertiary structure of water-soluble globular proteins is connected to a great extent with shielding from water of the overwhelming majority of bulky hydrophobic groups and the exposure to water of practically all the charged and other strongly hydrophilic groups [13]. Namely this general principle of the structure of water-soluble globular proteins was taken as a basis for defining the "most favourable" intermediate and final structures in this paper: the "most favourable" structures are implied

* It should be noted that in some cases (e.g., at formation of the β-structure from far along the chain regions) the secondary structure of some regions may be formed at this or following stages of self-organization on the neighbouring, along the chain already structurized parts of the molecule as on matrices.

to be those corresponding to the maximal dehydration of bulky hydrophobic groups and the minimal dehydration of charged and other strongly hydrophilic groups. More concretely, we searched for those intermediate and final structures in which all the strongly hydrophilic amino acid residues (Asn, Asp, Arg, Gln, Glu, Lys) localized, as a rule, at the surface of globular proteins, were exposed to water and the maximal number of strongly hydrophobic residues (Cys, Val, Met, Leu, Ile, Phe, Tyr, Trp) localized, as a rule, in the hydrophobic cores of globular proteins, were shielded from contacts with water. Dehydration of other amino acid residues which can be located both in the cores of proteins and on their surfaces was not taken into account in this paper. Other factors affecting the stability of intermediate and final structures were not taken into account either, in particular, the change of free energies of fragments connecting the helical regions.

In the present paper the above-suggested scheme of self-organization is applied to the myoglobin molecule, and it is shown that this scheme allows to obtain the native structure of myoglobin in rough resolution at least as one of the few most favourable structures and to predict the pathway of its self-organization.

3. Description of the model

A myoglobin molecule was modelled by nine helical regions (A, B, C, CD, D, E, F, G and H) and by flexible fragments connecting them: AB (see below), EF, FG and GH (see [14] for the nomenclature). Strongly hydrophobic groups on helical regions are enumerated in table 1. The ends of helical regions were fixed in accordance with the data of the X-ray analysis [15]. The CD region containing three bulky hydrophobic groups was considered to be helical since according to data of X-ray [15] and neutron [16] analysis it contains three hydrogen bonds of the 1–4 type in the backbone and therefore can be regarded as a fragment of the helix 3_{10}. At the same time it was assumed that conformations of the terminal residues CD1 and CD7 of this helix not fixed by hydrogen bonds can be arranged in such a way that the side groups of these residues could be shielded from water simultaneously with the side group of the residue CD4. The connecting fragments EF, FG and GH consist correspondingly of 8, 5 and 5 residues which is quite sufficient for the formation of a loop allowing the pairs of helices connected by them to be packed anti-parallel to each other into a hair-pin. Though the connecting fragment AB is absent from the data of X-ray analysis [15], we considered for generality also structures which contain the hair-pin A–B. For this, the helix B was shortened from the N-terminus by four residues with the break of four corresponding hydrogen bonds CO···NH. (An alternative formation of the connecting fragment AB by shortening the A-helix from the C-terminus is certainly less favourable, inasmuch as in the obtained anti-parallel hair-pin A–B the hydrophobic terminus of the B helix is not shielded by the shortened A-helix.) Structures

Table 1
Strongly hydrophobic groups in helical regions of sperm whale myoglobin

A	Trp 5, Leu 7, Val 8, Leu 9, Val 11, Trp 12, Val 15
B	Val 2, Ile 9, Leu 10, Ile 11, Leu 13, Phe 14
C	Leu 5
CD	Phe 1, Phe 4, Leu 7
D	Met 5
E	Leu 4, Val 9, Val 11, Leu 12, Leu 15, Ile 18, Leu 19
F	Leu 1, Leu 4
G	Ile 2, Tyr 4, Leu 5, Phe 7, Ile 8, Ile 12, Ile 13, Val 15, Leu 16
H	Met 8, Leu 12, Leu 14, Phe 15, Ile 19, Tyr 25

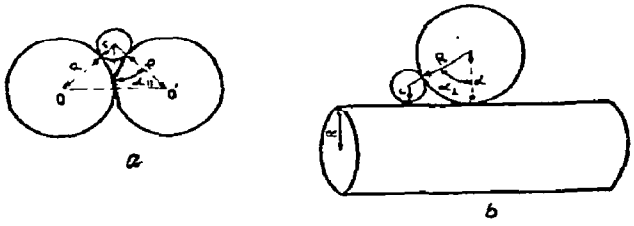


Fig. 1. Scheme of contact of two mutually parallel (a) and perpendicular (b) cylinders. $\alpha_{\parallel}$ and $\alpha_{\perp}$ are angle half-widths of dehydrated zones.

with the hair-pin A–B (i.e., with a reduced number of hydrogen bonds) were considered as favourable if the number of dehydrated bulky hydrophobic groups in them was greater than in structures without such a hair-pin.

Each helical region is modelled by a cylinder 10–12 Å in diameter and a length equal to the length of the helical region. Side groups are regarded to be localized in the places of projections of their C_{α}-atoms on the side surface of the cylinder. The contact of helical regions corresponds to the contact of cylinders modelling these regions.

From simple geometrical considerations the part of the side surface of every cylinder shielded from contacts with water by side surfaces of other cylinders (the "dehydrated zone") was determined for all the examined structures. Thus, at a contact of mutually parallel cylinders the length of the dehydrated zone of each is evidently equal to the length of the contact strip of the cylinders, and the angle half-widths of dehydrated zones $\alpha_{\parallel}$ are determined from the simple equation $\cos\alpha_{\parallel} = R/(R+r)$, where R is the radius of the cylinder and r is the effective radius of a water molecule (see fig. 1a). Taking $R = 5$ Å and $r = 1.8$ Å* we obtain $\alpha_{\parallel} = 42°$. An increase of R to 6 Å decreases the angle half-width only to 40°. It is easy to see that a change of the angle between the axes of the contacting cylinders within the limits of ± 20° also does not essentially influence the dimensions of dehydrated zones for helices consisting of less than 20–25 residues: the angle half-widths of the dehydrated zones at the ends of the contacting strip decrease with this only to 36°, while the lengths of zones can be taken as previously to be equal to the length of the contacting strip.

When the mutually perpendicular cylinders are in contact, the lengths of their dehydrated zones may be taken to be equal to the diameter of the cylinder (10 Å). The angle half-widths of zones $\alpha_{\perp}$ are determined from the equation $\cos\alpha_{\perp} = (R-r)/(R+r)$, which at $R = 5$ Å and $r = 1.8$ Å gives $\alpha_{\perp} = 62°$ (see fig. 1b).

It is easy to determine the dimensions of dehydrated zones in other cases as well. For example, for the packing of four cylinders shown in fig. 2 the length of the dehydrated zone for each cylinder is equal to 20 Å, and its angle half-width is equal to $90° + \alpha_{\parallel} + \alpha_{\perp} \approx 190°$ (see fig. 2b). In practical calculations the dehydrated zones for any mutual disposition of cylinders were approximated by zones corresponding either to parallel or to perpendicular dispositions. The angle half-widths of dehydrated zones at a parallel disposition of two cylinders were taken as 40° and at their perpendicular disposition as 60°. The left and the right boundaries of the dehydrated zone parallel to the helix axis were situated at an equal angle distance from the nearest to them C_{α}-atoms of bulky hydrophobic groups lying inside the zone.

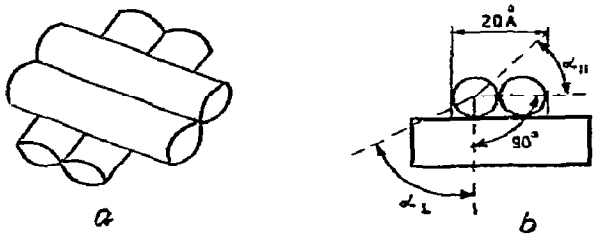


Fig. 2. An example of packing of four cylinders (a) and scheme of their contact (b).

* Maximal dimensions of a water molecule are 3.6 Å. The choice of the effective radius $r = 1.8$ Å means that the effect of dehydration becomes essential at such an approach of cylinders when the free motion of a water molecule between them becomes impossible.

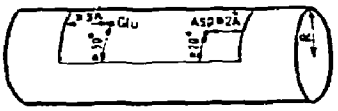


Fig. 3. Schematic picture of the dehydrated zone on the surface of the cylinder. Positions of strongly hydrophilic groups are shown in which they are considered to be dehydrated.

Side groups of hydrophobic residues located near the boundaries of the dehydrated zone will strive to enter this zone, while side groups of hydrophilic residues will, on the contrary, try to leave the zone. Therefore the side groups Asn and Asp are considered to be dehydrated only if the points representing them lie within the dehydrated zone at an angle distance of $\geqslant 20^\circ$ from the boundary of the zone parallel to the axis of the helix and at a distance of $\geqslant 2$ Å from the boundary of the zone perpendicular to the helix axis (see fig. 3). Longer hydrophilic side groups Arg, Gln, Glu and Lys are taken to be dehydrated if the points representing them lie within the zone at distances correspondingly of $\geqslant 30^\circ$ and $\geqslant 3$ Å from its boundaries (fig. 3). The bulky hydrophobic side groups are regarded to be dehydrated if the points representing them lie within the zone at a distance of $\geqslant 1$ Å from its boundary perpendicular to the helix axis.

Dehydration of non-helical regions each containing only one bulky hydrophobic group was not considered as well as dehydration of short helices C and D each containing also only one such a group. These short helices were regarded as connecting fragments correspondingly between helices B and CD, and between helices CD and E, but their shielding effect on the surfaces of other helices was taken into account. Owing to very small lengths of the fragments between helices B, C, CD, D and E it was assumed that the angles between the axes of these helices as well as the angle between the unshortened helices A and B cannot be smaller than 90°, and the point of contact of corresponding cylinder bases must lie within one of them. Possible shifts of helices A, B, CD, E, F, G and H relative to each other were determined taking into account the maximal lengths of the fragments connecting them. These maximal lengths were calculated for connecting fragments AB, EF, FG and GH in the extended state (3.5 Å per residue) and for the connecting fragments C and D in the helical state (1.5 Å per residue).

In order to take into account the geometry of the model and really possible situations, we assumed that at formation of a hair-pin from two cylinders with a connecting region of five amino acids the base of one cylinder must touch the side surface of the other and the point of contact must be apart from the base of the latter at a distance of not more than 3 Å. Shielding of the cylinder base by the side surface of the other cylinder not neighbouring along the chain was assumed to be unfavourable because of the possibility of steric overlaps with the non-helical region joined to this base.

4. Method of calculation

In accordance with the above-postulated scheme it was assumed that the second stage of self-organization (following the formation of helical regions) consists in the formation of complexes from helices neighbouring along the chain. These complexes can serve as centers of crystallization for the further process. To simplify the calculations it was considered that every center of crystallization is formed from two long helices (A, B, E, G and H) and one short helix adjacent with them along the chain (CD and F). Thus, pathways of self-organization were examined which begin with the formation of the following centers of crystallization: A–B–CD, B–CD–E, E–F–G and F–G–H. This gives four "one-center" pathways of self-organization beginning with the formation of any one of these centers, and three "two-center" pathways beginning with the simultaneous formation of any pair of these centers (A–B–CD, E–F–G; A–B–CD F–G–H and B–CD–E, F–G–H) (see fig. 4).

The next, third stage of the process of self-organization consists in the growth of centers of crystallization at the expense of helical regions neighbouring with them along the chain, and the fourth stage consists in the formation of a single compact structure of a protein molecule by the growth of one center of crystallization or by a consecutive collapse of two or more grown centers. It was supposed that during the growth of each center it adjoins firstly the long helix neighbouring with it along the chain and then the short one, i.e., the B–CD–E center adjoins firstly the

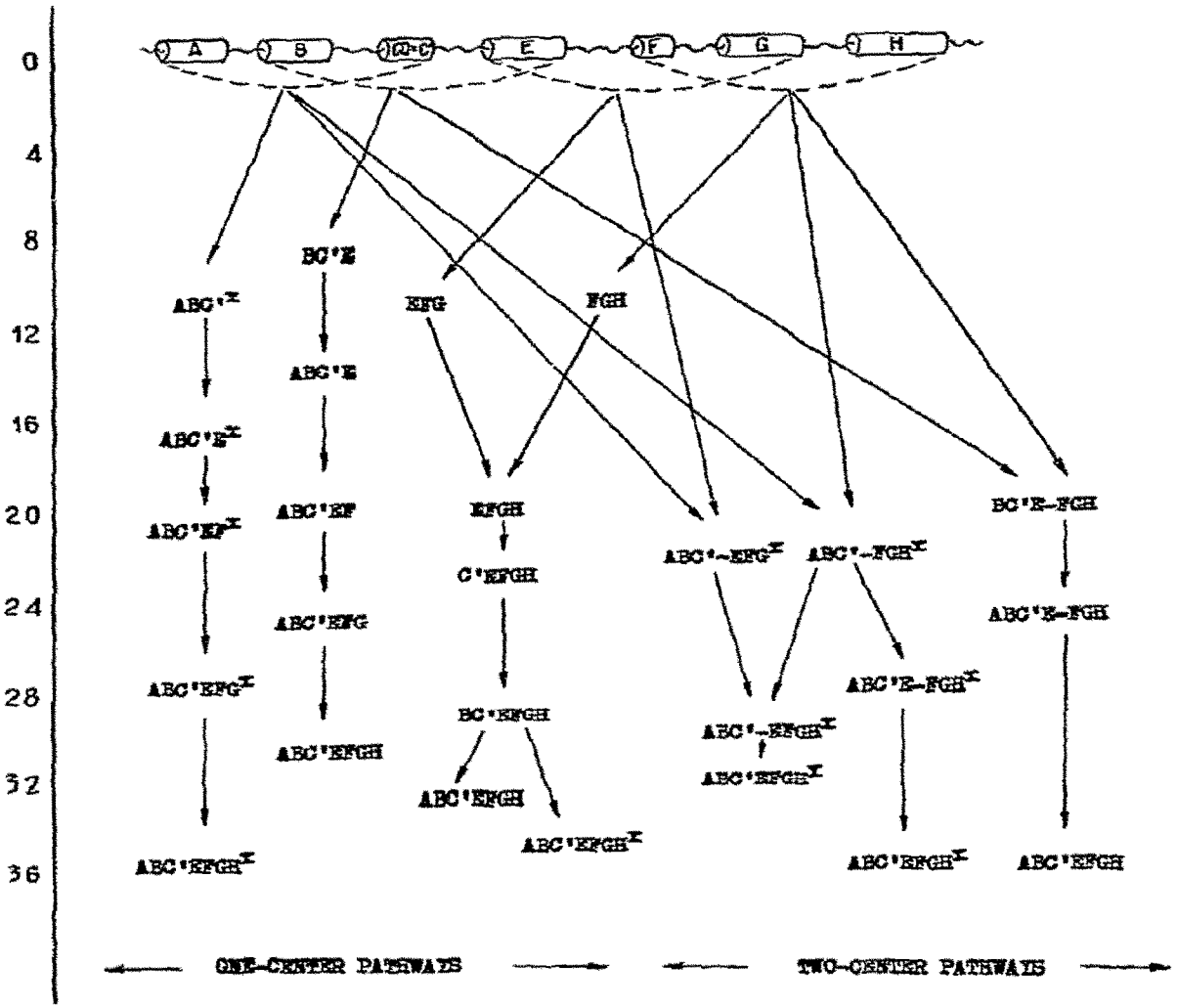


Fig. 4. Scheme of pathways of self-organization of a myoglobin molecule. Letters denote helices contained in the structure at a given stage of self-organization. The hyphen separates two centers which are not yet joined. Crosses mark the structures with the hair-pin A–B. A scale showing the number of dehydrated bulky hydrophobic groups is given at the left.

A-helix, and then the F-helix, while the E–F–G center firstly adjoins the H-helix and then the CD-helix. It was also supposed that in the presence of two centers of crystallization their building up firstly proceeds at the expense of the neighbouring helical regions and only then the joining of these centers into a single structure takes place. It is easy to see (fig. 4) that with these assumptions the branching takes place only for one two-center pathway from the seven pathways of self-organization, i.e., for the pathways beginning with the formation of the centers A–B–CD and F–G–H (these two centers "compete" with each other for the E-helix).

All intermediate and final structures were described in terms of the *type of packing* of the given set of helices and the more detailed *conformation* of the given set. The type of packing was characterized by an approximate mutual arrangement of cylinders modelling the helices (see figs. 5–10), and the conformation was determined by small transfers of cylinders

and their rotation around their axes within the limits of the given type of packing. In accordance with the scheme of self-organization accepted by us it was assumed that the most favourable conformation of centers of crystallization is not re-arranged with the growth or joining of these centers.

Therefore the joining of helices to the already formed intermediate structures was performed only for the most favourable structures with a great enough number of dehydrated bulky hydrophobic groups. However, the roughness of our method makes it impossible to determine quite reliably the real degree of dehydration of helices in all the considered structures. To increase the reliability of the method we treated as most favourable not only the conformations with the maximal number of bulky hydrophobic groups but also the conformations with this number decreased by one. Among the latter conformations we rejected all those with the number of dehydrated Val greater than the minimal number of Val in the conformation with the maximal number of dehydrated bulky hydrophobic groups as unfavourable for taking into account the lesser hydrophobicity of Val in comparison with the more bulky groups.

The searches of the most favourable structures were performed practically in the following way. The search for possible types of packing for every center of crystallization was done manually (on plasticine models or graphically) within the limits of the model described in the previous section. We considered only those types of packing which do not lead to dehydration of strongly hydrophilic groups (see above) or to formation of long internal channels open for water. Then for every type of packing the conformation corresponding to the maximal number of dehydrated bulky hydrophobic groups of all the helices in the center was considered as well as the other above-mentioned "most favourable" conformations. The searching of conformations for every type of packing was also done manually. To this aim an evolvent of the surface of the corresponding cylinder was plotted on the plane for every helix and the

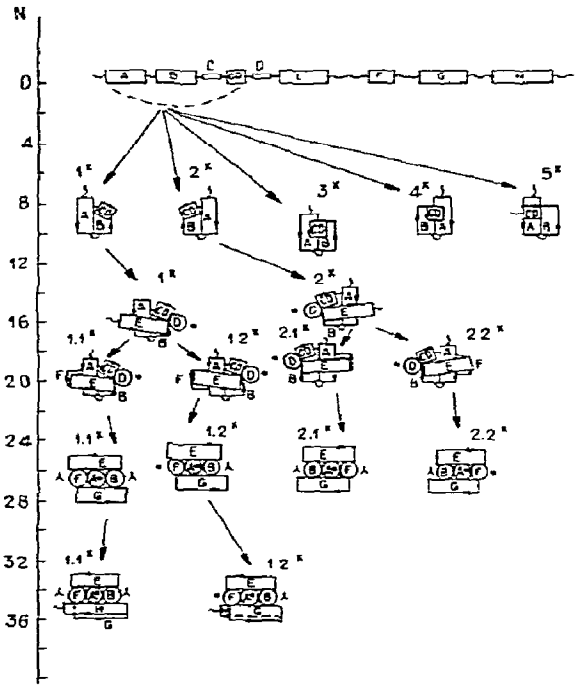


Fig. 5. The most favourable intermediate and final structures of the one-center pathway of self-organization with the center of crystallization A–B–CD.

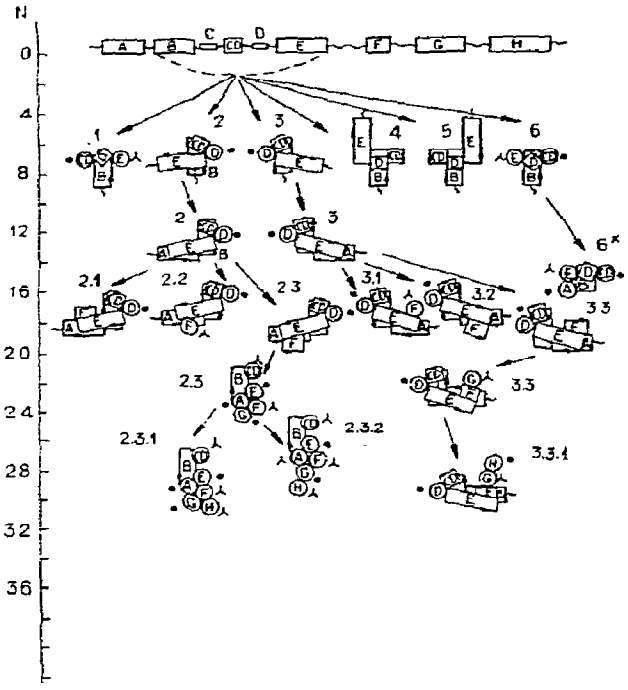


Fig. 6. The most favourable intermediate and final structures of the one-center pathway of self-organization with the center of crystallization B–CD–E.

positions of C_α-atoms of all the residues of the helix were marked on it. The given type of packing of helices imposes definite dimensions and a certain form of their dehydrated zones with an accuracy to the parameters dependent on the small transfer of helices (the rotation of the helices around their axes does not, evidently, change the dimensions and the form of the dehydrated zone), while the conformation of the center of crystallization determines these parameters and fixes the position of each zone on the evolvent of the corresponding cylinder. Therefore for every type of packing of all cylinders such positions of dehydrated zones were chosen on the evolvent at which the maximal number of bulky hydrophobic groups (Cys, Ile, Leu, Met, Phe, Trp, Tyr, Val) enters all these zones but neither of strongly hydrophilic groups (Asn, Asp, Arg, Gln, Glu, Lys) does so. The parameters of zones dependent on parallel transfer of cylinders varied within the limits permitted by maximal lengths of connecting fragments. After this, the number of dehydrated bulky hydrophobic groups was calculated for every position of zones on the evolvents, i.e., for every conformation of the center of crystallization, and only the "most favourable" conformations were considered.

The same criterion was also applied to the conformations corresponding to different types of packing, i.e., only those types of packing of each center were considered which have the "most favourable" conformations (see above).

An analogous consideration was applied to the packings and conformations of centers of crystallization at all the stages of their growth. To the conformation of the center, considered at the previous stage, the helical region neighbouring along the chain was joined in all the possible ways and the packing and conformation of the new structure were determined by the method described above. During this, the conformation of the center to which joining proceeds did not vary, while only the type of joining, the small transfer and the rotation of the joined helical region around its axis varied. Again only the "most favourable"

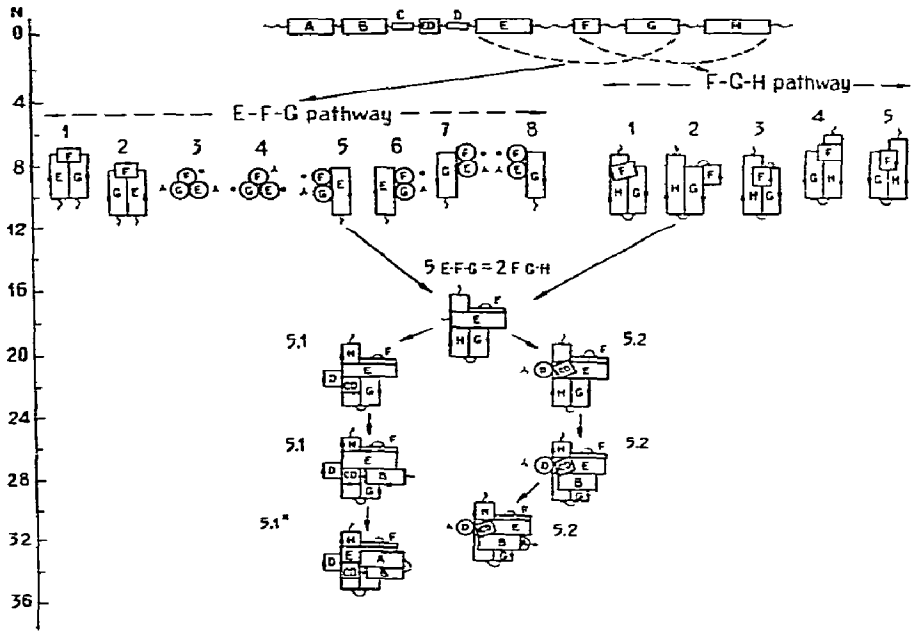


Fig. 7. The most favourable intermediate and final structures of the one-center pathways of self-organization with the centers of crystallization E–F–G and F–G–H. At the formation of the E–F–G–H complex both pathways become confluent.

conformations of the new structure were considered.

For two-center pathways the indicated criterion was applied to each center individually at all the stages preceding the joining of these centers into a single structure. The joining of two centers into a single structure was regarded as the growth of one

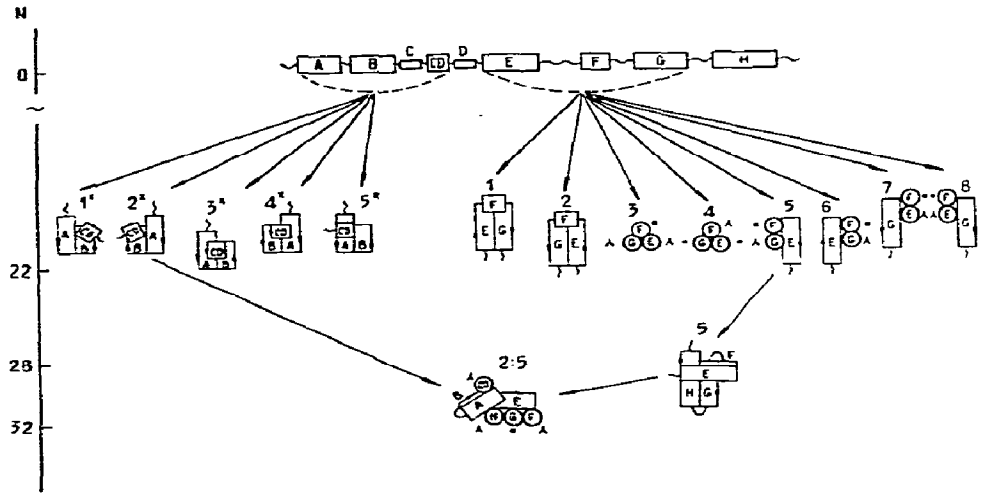


Fig. 8. The most favourable intermediate and final structures of the two-center pathways of self-organization with centers of crystallization A–B–CD + E–F–G.

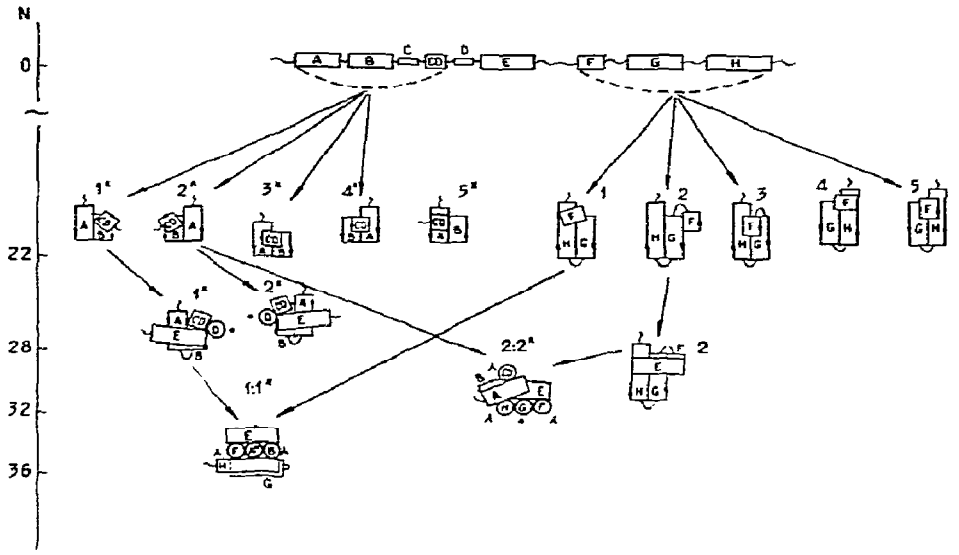


Fig. 9. The most favourable intermediate and final structures of the two-center pathway of self-organization with the centers *A–B–CD + F–G–H*. The E-helix can be joined either to the first (structures 1 and 2) or to the second (structure 3) centers. In case of the formation of the complex E–F–G–H the further pathway of self-organization coincides with the pathway A–B–CD + E–F–G (see fig. 8).

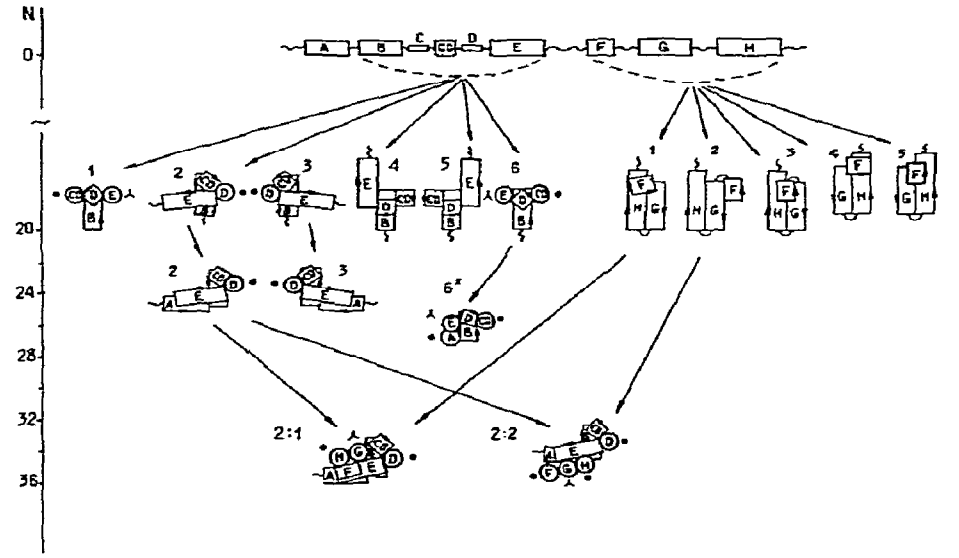


Fig. 10. The most favourable intermediate and final structures of the two-center pathway of self-organization with the centers of crystallization *B–CD–E + F–G–H*.

center, i.e., the conformations of every center did not vary and it was only their mutual packing that varied. Only the "most favourable" final structures were considered.

5. Results

Figures 5–10 show the conformations of the most favourable intermediate and final structures for all the pathways of self-organization presented in fig. 4. The helical regions are shown by projections of cylinders to the plane of the figure (by circles or by rectangles). The direction from the N-terminus of the chain to the C-terminus is shown in the rectangles with an arrow and in the circles with a point if the corresponding vector is directed from the plane of the figure and with the ⅄ sign if the vector is directed beyond the plane of the figure. So as not to overburden the figures, the helix C (and sometimes also helix D) is not shown. The types of packing of each center of crystallization are denoted in figs. 5–10 by numbers. If some type of packing leads at the following stage to several different types of packing its number is maintained and the number of the type of packing grown from it is added after the point. The scale of the number of dehydrated bulky hydrophobic groups is given at the left of each of figs. 5–10; the types of packing containing hair-pin A–B (and correspondingly a smaller number of hydrogen bonds) are marked with crosses. The bottom of the picture of each structure indicates the maximal number of dehydrated bulky groups in the corresponding type of packing.

Table 2 lists for each type of packing the maximal numbers of dehydrated bulky hydrophobic groups on helices and the numbers of considered most favourable conformations (the latter numbers listed in column 5 should be considered as approximate because of difficulties of manual search of conformations) as well as the numbers of the most favourable conformations of the preceding stage leading to this type of packing and the numbers of the most favourable conformations of this type of packing leading to the most favourable conformations at the following stage. If this type of packing does not lead to any favourable structure at the following stage the table gives as an example one of its most favourable conformations and lists bulky hydrophobic groups dehydrated in this conformation. If different conformations of this type of packing lead to different types of packing at the following stages, the table gives for each of the new types of packing one conformation leading to it. Different conformations of the given type of packing are denoted by adding a letter after the number corresponding to this type of packing. All the possible favourable conformations can be obtained in a systematic way as shown for the example of the center F–G–H in the appendix.

For example, fig. 6 gives six most favourable types of packing of the center of crystallization B–CD–E. Four of these six types (types of packing, 1, 4, 5 and 6) have each four most favourable conformations and each of two others only one most favourable conformation. The last two types of packing (two and three) as well as all the four conformations of type of packing 6 give the most favourable conformations after joining of the A-helix. Types of packing 2 and 3 have at this stage of self-organization eight most favourable conformations in each and type of packing 6 containing the hair-pin A–B has six most favourable conformations. The favourable joining of the F-helix is possible to five different conformations of type of packing 2 leading to three new types of packing: 2.1, 2.2 and 2.3 with one most favourable conformation in types 2.1 and 2.2 and four most favourable conformations in type 2.3. Type of packing 3 produces the same number of new types of packing and conformations in them, while type of packing 6 gives no favourable conformations at this stage. The favourable joining of the G-helix is possible to three conformations of type of packing 2.3 and to one conformation of type of packing 3.3 giving three new conformations for each type of packing 2.3 and 3.3. Joining of the H-helix to the type of packing 2.3 gives two new types of packing 2.3.1 and 2.3.2 with two and five different favourable conformations. One conformation of the type of packing 3.3 leads after the joining of the H-helix to type of packing 3.3.1 with three different favourable conformations. One of the most favourable conformations of each type of packing is given in table 2. If some type of packing leads to several new types of packing at the following stage, table 2 gives for each new type of packing one favourable conformation leading to it. Thus, at the stage A–B–CD–E–F–G two favourable conformations

Table 2
Intermediate and final structures on all the pathways of self-organization

Centers of crystallization	Helices included in a growing center	Type of packing	Max. number of dehydrated groups	Number of conformations in the type of packing	Number of preceding conformations	Number of continuing conformations	Examples of conformations	Dehydrated Bulky Hydrophobic Groups on Helices						
								A	B	CD	E	F	G	H
A-B-CD-E	A-B-CD	1	10^{X}	4		3	1a	8, 12,15	10,11,13,14	1,4,7				
		2	10^{X}	2		1	2a	5, 8, 12,	9,10, 13,14	-"-				
		3	11^{X}	12		0	3a	5, 8,9, 12,15	10, 13,14	-"-				
		4	10^{X}	2		0	4a	5, 8, 11,12	- " -	-"-				
		5	10^{X}	3		0	5a	5, 8,9 12,15	9, 13,	-"-				
	A-B-CD-E	1	17^{X}	5	3	5	1a	5, 8,9 12,15	10,11,13,14	1,4,7	4,11,12,15			
		2	16^{X}	1	1	1	2a	5,7,8, 11,12	9,10, 13,14	-"-	- " -			
	A-B-CD-E-F	1.1	21^{X}	9	5	9	1a.1	5, 8,9, 12,15	10,11,13,14	1,4,7	4,11,12,15,18,19	1,4		
		1.2	21^{X}	9	5	5	1a.2	- " -	- " -	-"-	- " -	-"-		
		2.1	20^{X}	1	1	1	2a.1	5,7,8, 11,12	9,10, 13,14	-"-	- " -	-"-		
		2.2	20^{X}	1	1	1	2a.2	- " -	- " -	-"-	- " -	-"-		
	A-B-CD-E-F-G	1.1	28^{X}	22	9	7	1a.1	5,7,8,9,11,12,15	9,10,11,13,14	1,4,7	4,11,12,15,18,19	1,4	4,5, 8,12, 15	
		1.2	27^{X}	3	3	3	1a.2	5, 8,9,11,12,15	- " -	-"-	- " -	-"-	2, 5, 12,13, 16	
		2.1	27^{X}	3	1	0	2a.1	5,7,8,9,11,12,	- " -	-"-	- " -	-"-	- " -	
		2.2	29^{X}	3	1	0	2a.2	- " -	- " -	-"-	- " -	-"-	- " -	
	A-B-CD-E-F-G-H	1.1	36^{X}	8	7		1a.1	5,7,8,9,11,12,15	9,10,11,13,14	1,4,7	4,11,12,15,18,19	1,4	2,4,5, 8,12,13,15,16	8,12,14,15,19
		1.2	36^{X}	3	3		1a.2	- " -	- " -	-"-	- " -	-"-	- " -	- " -
M-B-A	B-CD-E	1	9	4		0	1a		9,10,11,13,14	1,4,7	4,			
		2	8	1		1	2		9,10,11, 14	-"-	-"-			
		3	8	1		1	3		9,10,11,13	-"-	-"-			
		4	9	4		0	4a		9,10,11,13,14	-"-	-"-			
		5	9	4		0	5a		- " -	-"-	-"-			
		6	9	4		4	6a		- " -	-"-	-"-			
	A-B-CD-E	2	14	8	1	8	2a	5, 8, 12	9,10,11, 14	1,4,7	4, 12,15, 19			
							2b	8, 12,15	- " -	-"-	4,11, 15,18			
							2c	8, 11, 15	- " -	-"-	- " -			
		3	14	8	1	4	3a	5, 8, 12	9,10,11,13	-"-	4, 12,15, 19			
							3b	8, 12,15	- " -	-"-	4,11, 15,18			
							3c	8, 11, 15	- " -	-"-	- " -			
		6	16	6	4	0	6a	5, 8, 11,12,15	9,10,11,13,14	-"-	4, 12,15			
	A-B-CD-E-F	2.1	19	1	1	0	2a.1	5, 8, 11,12	9,10,11, 14	1,4,7	4,11,12,15,18,19	1,4		
		2.2	19	1	1	0	2b.1	5, 8,9, 12,15	- " -	-"-	4,11, 15,18,19	-"-		
		2.3	20	4	3	1	2c.3	5, 8, 11,12,15	- " -	-"-	4,11,12,15,18,19	-"-		
		3.1	19	1	1	0	3a.1	5, 8,9, 12,15	9,10,11,13	-"-	4,11, 15,18,19	-"-		
		3.2	19	1	1	0	3b.1	5, 8, 11,12	- " -	-"-	4,11,12,15,18,19	-"-		
		3.3	20	4	3	1	3c.1	5, 8, 11,12,15	- " -	-"-	- " -	-"-		

Table 2 (continued)

	A-B-CD-E-F-G	2;3 3;3	25 24	3 3	1 1	3 1	2a,3a 2a,3b 3a,3a	5; 8,9,11,12,15 " " " " " "	9,10,11, 14 " " " 9,10,11,13	1,4,7 " " " " " "	4, 11,12,15,18,19 " " " " " "	1,4 " "	2, 5, 13 4, 7,8 " " "	
	A-B-CD-E-F-G-H	2;3;1 2;3;2 3;3;1	31 31 31	2 3 3	2 3 1		2a,3a,1a 2a,3b,2a 3a,3a,1a	5; 8,9,11,12,15 " " " " " "	9,10,11, 14 " " " 9,10,11,13	1,4,7 " " " " " "	4, 11,12,15,18,19 " " " " " "	1,4 " "	2,4,5, 8, 13,15 2,4, 7,8, 13, 16 " " "	8,12, 15,19 " " " " " "
E-F-G	E-F-G	1 2 3 4 5 6 7 8	10 11 10 10 11 11 10 10	11 11 3 3 11 11 2 2		0 0 0 0 5 0 0 0	1a 2a 3a 4a 5a 6a 7a 8a				12,15,18,19 4, 11, 15,18,19 11,12,15,18,19 " " " " " " " " " 12,15,18,19 11, 15,18,19	1,4 " " " " " " "	2, 5, 13, 16 2, 5, 8,12, 2, 5, 8 4, 7,8 4,5,7,8 2,4,5, 8 2, 5, 12,13 " " "	
	E-F-G-H	5	20	4	4	4	5a				4, 11,12,15,18,19	1,4	2,4,5,7,8, 13, 16	8,12,14,15,19
	CD-E-F-G-H	5;1 5;2	23 23	4 6	4 4	4 6	5a,1 5a,2			1,4,7 " " "	4, 11,12,15,18,19 " " "	1,4 "	2,4,5,7,8 13, 16 " " "	8,12,14,15,19 " " "
	B-CD-E-F-G-H	5;1 5;2	29 29	5 7	4 6	4 7	5a,1 5a,2		10,11,13,14 9,10, 13,14	1,4,7 " " "	4, 11,12,15,18,19 " " "	1,4 "	2,4,5,7,8,12,13,15,16 " " "	8,12,14,15,19 " " "
	A-B-CD-E-F-G-H	5;1 5;2	35[x] 33	5 7	5 7		5a,1 5a,2	5; 8, 11,12,15 8, 11,12,15	9,10,11,13,14 9,10, 13,14	1,4,7 " " "	4, 11,12,15,18,19 " " "	1,4 "	2,4,5,7,8,12,13,15,16 " " "	8,12,14,15,19 " " "
F-G-H	F-G-H	1 2 3 4 5	11 11 11 10 10	8 5 5 10 2		8 5 0 0 0	1a 2a 3a 4a 5a					1,4 " " " "	2, 5, 13, 16 2,4, 7 13, 16 4,5, 8,12, 16 2, 5, 12, 16 " " "	8,12,14,15,19 8,12, 15,19 8,12, 15,19 8,12, 15,19 " " "
	E-F-G-H	2	coincides with 5 on the pathway E-F-G											
	CD-E-F-G-H	2;1 2;2	" "											
	B-CD-E-F-G-H	2;1 2;2	" "											
	A-B-CD-E-F-G-H	2;1 2;2	" "											
A-B-CD + E-F-G	A-B-CD-E-F-G-H	2 : 5	32[x]	4			2a : 5a	5; 8,9, 12	9,10,11,13,14	1,4,7	4, 11,12,15,18,19	1,4	2,4,5,7,8, 13, 16	8,12,14,15,19
A-B-CD + F-G-H	A-B-CD-E-F-G-H	1 : 1 2 : 2	36[x] 32[x]	8 4	coincides with 2 : 5 on the pathway A-B-CD-E-F-G		1a : 2a	5,7,8,9,11,12,15 " " "	9,10,11,13,14 " " "	1,4,7 " " "	4, 11,12,15,18,19 " " "	1,4 "	2,4,5, 8,12,13,15,16 " " "	8,12,14,15,19 " " "
B-CD-E + F-G-H	A-B-CD-E-F-G-H	2 : 1 2 : 2	36 36	7 9			2a : 1a 2a : 2a	5,7,8, 11,12,15 5, 8,9,11,12,15	9,10,11,13,14 9,10,11, 14	1,4,7 " " "	4, 11,12,15,18,19 4,9,11,12,15,18,19	1,4 "	2,4,5,7,8,12,13,15,16 " " "	8,12,14,15,19 " " "

The designation under each helix gives numerals of the bulky hydrophobic groups on them which are shielded from contact with water (see table 1). The sign - " - means that the same groups are shielded as in the preceding line.

Table 3
Comparison of final structures

Centers of crystallization	Final structures	Exposed bulky hydrophobic groups on helices					Total number of exposed bulky hydrophobic groups	Number of Val among them
		A	B	E	G	H		
Structures without AB hairpin								
B–CD–E + F–G–H	2a : 1a 2c : 2a	Leu 9 Leu 7	Val 2 Val 2, Leu 13	Val 9 -	- -	Tyr 23 Tyr 23	4 4	2 1
E–F–G (F–G–H)	5a.2 (2a.2)	Trp 5, Leu 7 Leu 9	Val 2, Leu 13	Val 9	-	Tyr 23	7	2
B–CD–E	2c.3a.1a 2c.3b.1a 3c.3a.1a	Leu 7 Leu 7 Leu 7	Val 2, Leu 13 Val 2, Leu 13 Val 2, Phe 14	Val 9 Val 9 Val 9	Phe 7, Ile 12, Leu 16 Leu 5, Ile 12, Val 15 Leu 5, Ile 12, Val 15	Leu 14, Tyr 23 Leu 14, Tyr 23 Leu 14, Tyr 23	9 9 9	2 3 3
Structures with AB hairpin								
A–B–CD	1a.1 1a.2	-	Val 2	Val 9	Phe 7	Tyr 23	4	2
E–F–G (F–G–H)	5a.1 (2a.1)	Leu 7, Leu 9	Val 2, Ile 11	Val 9	-	Tyr 23	6	2
A–B–CD + F–G–H	1a : 2a	-	Val 2	Val 9	Phe 7	Tyr 23	4	2
A–B–CD + E–F–G (A–B–CD) + F–G–H	2a : 5a (2a : 2a)	Leu 7, Val 11 Val 15	Val 2	Val 9	Ile 12, Val 15	Tyr 23	8	4

Note: In case two structures coincide, designation of one of them is given in brackets.

2c.3a and 2c.3b are given for type of packing 2.3 leading to new types of packing 2.3.1 and 2.3.2 of the complex A–B–CD–E–F–G–H.

Centers of crystallization E–F–G and F–G–H (fig. 7) have eight and five different types of packing with a different number of most favourable conformations in each (all the conformations of the center of crystallization F–G–H are listed in detail in the appendix). Both these centers of crystallization lead to the same most favourable type of packing at the joining to them of the helix E, and starting from this moment the two pathways merge. The joining of the helix CD leads to two comparably favourable types of packing. At a further growth these two types of packing lead to two different final types of packing, one of them containing the hair-pin A–B.

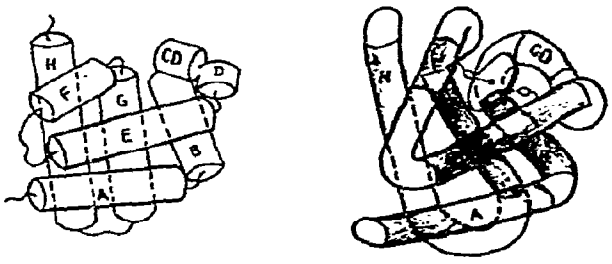
Table 3 lists the final structures obtained on all the pathways of self-organization. Designations of the most favourable conformation, exposed bulky hydrophobic groups on each helix, the total number of the exposed bulky hydrophobic groups in each of the listed conformations and the number of Val among them are given for each structure. Of the eleven obtained final structures six do not contain the hair-pin A–B and have from four to nine exposed bulky hydrophobic groups. Fig. 4 shows that the most favourable among all the final structures are the structures formed on the two-center pathway of self-organization beginning with the formation of the B–CD–E and F–G–H centers (fig. 10). Naturally the formation of the B–CD–E center and the joining of the A-helix to it, as well as the formation of the F–G–H center proceed on this pathway in the same manner as on the corresponding one-center pathways of self-organization. As a result three comparably favourable types of packing of the complex from helices A, B, CD, E (2, 3 and 6) and five comparably favourable types of packing of the complex from helices F, G and H are formed. The joining of these centers without a change of their conformations can give fifteen different final structures. However, only joining of type of packing 2 of the complex A–B–CD–E with types of packing 1 and 2 of the center of crystallization F–G–H leads to two final structures comparably favourable with types of packing 2:1 and 2:2. Table 3 shows that these final structures have the minimal number of exposed bulky hydrophobic groups, namely four.

It is seen from table 3 that the most favourable structure obtained on the pathways of self-organization with the centers E–F–G and F–G–H has seven exposed bulky hydrophobic groups and therefore is considerably less favourable than the structures obtained on the two-center pathway with the centers B–CD–E and F–G–H. The moreso it is true for the structures obtained on the one-center pathway with the center of crystallization B–CD–E which have nine exposed bulky hydrophobic groups each. The most favourable structures with the hair-pin A–B have from four to eight exposed bulky hydrophobic groups and therefore a considerable loss of free energy at disruption of a few hydrogen bonds necessary for formation of the hair-pin A–B is not compensated in them by the gain in free energy of dehydration in comparison with the structures not containing the hair-pin A–B. Consequently all these structures also are noticeably less favourable in comparison with the structures obtained on the two-center pathway with the centers B–CD–E and F–G–H.

The two final structures obtained on the pathway of self-organization B–CD–E + F–G–H according to the criteria used in our work are equally favourable and preference cannot be given to any. However, it is seen from table 3 that structure 2:1 has a somewhat lower free energy of dehydration than alternative structure 2:2 as it has two Val exposed to water and two more bulky hydrophobic groups instead of one Val and three more bulky hydrophobic groups in the alternative structure. Examination of the schematic representation of these structures in fig. 10 shows that this structure also possesses a greater general compactness and that the location of helices F and E in it is considerably more favourable for joining of the heme to His E7 and F8 than in structure 2:2.

6. Discussion

Fig. 11 shows the final structure 2:1 presented in fig. 10, as well as the well-known structure of native myoglobin in a 5 Å resolution [17] shown for comparison. It is seen from the figure that the theoretically obtained structure practically coincides with the native structure of myoglobin in a rough resolution: in both cases helices A and B form a nearly right angle, helices C, D and the region CD joining them form a



(a) (b)

Fig. 11. (a) The most favourable final structure of sperm whale apomyoglobin obtained as a result of the "technological" pathway of self-organization (structure 2:1 in fig. 10); (b) the structure of native sperm whale myoglobin in a 5 Å resolution [17].

loop, helix E goes approximately in anti-parallels to the A-helix simultaneously contacting with the middle of the B-helix, helix F is directed to the E-helix at an acute angle, and the G- and H-helices form an approximately anti-parallel hair-pin lying in a plane parallel to the plane of helices A and E. Therefore we think that the pathway shown in fig. 10 and leading to the final structure 2:1 in the main features coincides with the "native" pathway of self-organization of a myoglobin molecule although the details of this pathway will possibly require a more precise definition (e.g., the A-helix can be joined after and not before the collapse of the two halves of the molecule).

Thus beginning with the consideration of 24 favourable types of packing with more than 100 different favourable conformations corresponding to formation of different centers of crystallization we obtained, using the formulated above criterion of the "most favourable" structures, at each stage of each of the seven pathways of self-organization 11 different final types of packing from which only two types of packing were found to be comparably favourable according to our criterion. One of these two most favourable structures turned out to approximately coincide with the myoglobin native structure. This result can be hardly treated as a chance one and it should probably be considered as a confirmation of the applicability of the criterion proposed by us for selection of intermediate and final structures.

It should be underlined that in this communication we have considered only the self-organization of the protein part of myoglobin and not the joining of heme, therefore the structure obtained by us does not pertain to myoglobin but to apomyoglobin. Therefore the approximate coincidence of this structure with the structure of native myoglobin may mean that the joining of heme does not lead to the re-arrangement of the most stable structure of apomyoglobin, but to a further stabilization of its already pre-existing structure which may differ from the structure of holoprotein not in principle but only by a less stability and greater susceptibility to fluctuations [5]. The circumstance that apomyoglobin is characterized by a less degree of helicity [18] and larger dimensions of the molecule [19] than those of myoglobin does not contradict this point of view, since the fluctuations of the secondary structure must lead to a decrease of the mean degree of helicity while the fluctuations of the tertiary structure must lead to an increase of mean dimensions of the molecule.

The pathway of self-organization of a myoglobin molecule obtained by us consists of an independent formation of the structures of both halves of the molecule with their subsequent collapse into a single compact structure. This conclusion is in good agreement with the analysis of intramolecular contacts in a myoglobin molecule [20–22] which implies that intramolecular interactions in the protein part of the molecule stabilize the internal structures of both its parts to a greater extent than the structure of the molecule as a whole. On the basis of these results it can be presumed that the locking and unlocking of both halves of the molecule represents an essential part of the processes of heat fluctuations in the molecule of apomyoglobin and that the building-in of the heme between these halves stabilizes their locked structure. Thus, the structure of apomyoglobin obtained by us and the pathway of its self-organization seem to be in good accordance with the main function of apomyoglobin, i.e., the joining of the heme. In this connection it is interesting to note once again that in our final structure 2:1 (in contrast to the alternative final structure 2:2 on the same pathway of self-organization) His E7 and His F8 are located on the internal surfaces of helices E and F opposite each other and can be simultaneously joined to the heme.

The model used by us for computing is certainly very rough. Therefore we have studied the dependence of the result on small changes of parameters involved in computing. Thus an analogous result was obtained

on the two-center pathway of self-organization B–CD–E + F–G–H at a somewhat different localization of the termini of helices and without taking into account the F-helix. Another criterion of the "most favourable" structures was also used [23]. Tanford's data on the free energy of amino acid transfer from water to ethanol [24] were used for calculating the free energy of bulky hydrophobic group dehydration in different conformations. We attributed to the "most favourable" structures those structures, the free energy of dehydration of which differed from the maximal not more than by 4.1 kcal/mole which corresponds to the difference in the exposure of two Val or one Val and one Leu. With this only one new type of packing appears comparably favourable with the native on the same pathway of self-organization. And, finally, the transfer of His to the group of strongly hydrophilic amino acids also had no effect on the result.

This evidences that the main result of the study (namely, that the myoglobin native structure is found to be nearly the same as one of the two or three structures obtained at the assembly of apomyoglobin from two halves with the centers of crystallization B–CD–E and F–G–H) is invariant with changes of parameters involved in calculation. Examination of amino acid replacements in eight known primary structures of different myoglobins [25] has shown that this result is true for all of them.

The present paper represents the first attempt to obtain theoretically the tertiary structure of globular protein and to predict the pathway of its self-organization. At this stage of development of the theory our aim was not to derive the tertiary structure directly from the primary one and we proceeded from helical regions localized according to the data of X-ray analysis. The progress achieved lately in the theory of the secondary structure of globular proteins [10,11,26, 27] allows to hope that the theoretical localization of the secondary structure will become in the nearest future a sufficiently reliable basis for the theoretical prediction of the tertiary structure.

The criteria for selection of the most favourable structures and the method of searching for different compact structures used in the present paper are of course very rough. Although they work for the myoglobin molecules consisting of long helical regions, they can hardly lead to success for other proteins containing only comparatively short helical regions or regions with β-structure. Therefore, they should in no way be considered as a general method for the theoretical obtainment of tertiary structures of globular proteins. The results obtained above should be regarded rather as an illustration of the usefulness of the *approach suggested by us to the theoretical obtainment of the tertiary structure of globular protein as a result of the "technological" pathway of its self-organization.* It seems to us that namely this approach and not its concrete realization in the given paper can be of general importance for the problem of theoretical prediction of the tertiary structure of globular proteins.

Appendix: Example of the search for the most favourable conformations

To render the method of search for the most favourable structures, described above, more clearly we will give it in detail here for the center of crystallization F–G–H.

As indicated in the paper the search was carried out manually; at first all possible types of packing were searched graphically or on plasticine models and then conformations differing one from another by small transfer and rotations of helices around their axis were specified within the framework of the given type of packing. In the dominating majority of cases a graphic study was quite sufficient and plasticine cylinders connected by flexible fragments were needed only when a two-dimensional drawing did not give a clear idea of the three-dimensional structure.

In the hair-pins of the perpendicular type only the N-terminus of the H-helix can be shielded from water because of the small length of the GH fragment. Since the N-terminus of the H-helix does not contain bulky hydrophobic groups it is evident that the hair-pins of such a type will not be favourable enough. Therefore all favourable packings must include the hair-pin G–H of the anti-parallel type. Fig. 12 represents all six possible types of packing of the three helices F, G and H including the G–H hair-pin of the anti-parallel type. Types of packing 1 and 4, 2 and 6, 3 and 5 are connected by an approximate mirror symmetry relation. There are no other types of packing with a sufficiently favourable shielding of hydrophobic regions.

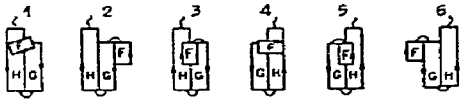


Fig. 12. Possible types of packing of the center of crystallization F–G–H.

Fig. 13 represents evolvents of the side surfaces of helices G and H on the plane. In order to find all possible rotations of the helix G in the hair-pin G–H we select in fig. 13a a bulky hydrophobic group lying at the right edge of the hydrophobic region of the helix G. This is Ile G13. We include it by all possible means into the strip 80° wide shielded by the helix H so as to dispose the right and the left boundaries of the dehydrated zone parallel to the helix axis at an equal angle distance from the C_α-atoms of bulky hydrophobic groups lying inside the zone nearest to them. Then we move this strip from right to left including into it consecutive bulky hydrophobic groups. Then the dehydrated groups to the extreme right will consecutively be the groups Ile G2, Leu G16, Leu G5, etc. Upon reaching the left boundary of the hydrophobic region of the helix G we will have all possible rotations of the helix G relative to the helix H. All possible rotations of the helix H relative to the helix G can be found in the same way. Possible shifts of helices relative to each other are determined by the length of the non-helical region GH. Dehydrated zones

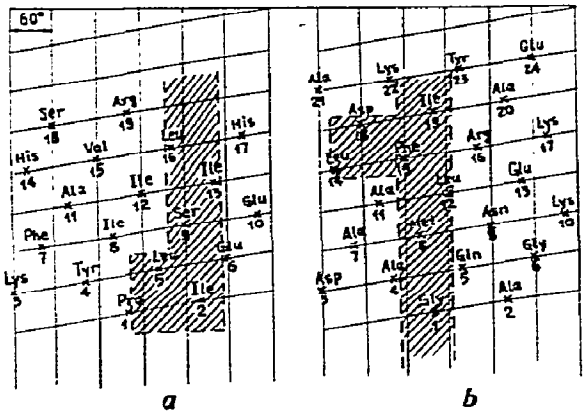


Fig. 13. The evolvents of the G and H cylinders on the plane. The hatched zones correspond to dehydrated zones of the conformation 1a.

on helices G and H corresponding to conformation 1a are hatched in fig. 13. A hair-pin of the anti-parallel type leads to the hatched zone parallel to helix axes, and hair-pins F–G and F–H of the perpendicular type lead to zones perpendicular to axes of the helices G and H also hatched in fig. 13.

Table 4 lists dehydrated bulky hydrophobic groups of all the conformations obtained by the method described above with ten and more dehydrated bulky hydrophobic groups for all the six types of packing except types 2 and 6 for which only a part of the conformations is listed. For types of packing 2 and 6 location of the axes of helices F, G and H in one plane is considerably more favourable for the further growth of the structure than non-plane location since the obtained plane permits a simultaneous contact of the helix E with three helices without formation of internal cavities or channels as well as joining of the hydrophobic plane formed by the other half of the molecule. Examination of the growth of non-planar conformations of types of packing 2 and 6 confirms this conclusion. Thus, non-planar conformations of types 2 and 6 can be considered as unfavourable. All conformations of type 6 with ten or eleven dehydrated bulky hydrophobic groups pertain to the non-planar type as well as all the unlisted conformations of type 2. Thus, conformations 6a and 6b listed as an example and unlisted conformations of type 2 can be rejected (the rejected conformations are marked with the sign "–" in the corresponding lines of the last right column of table 4).

It is seen from table 4 that the maximal number of dehydrated bulky hydrophobic groups on the helices of the center of crystallization F–G–H is equal to 11 and in some conformations there is no Val among these 11 groups. According to our criterion of the "most favourable" structure all the conformations with ten dehydrated bulky hydrophobic groups with one Val among them are considered as unfavourable and can be rejected. These are conformations 1i, 1j, 3f, 3g, 4j, 4k, 4l, 4n, 4o, 4p, 5c, 5d, All the other conformations not marked with the sign "–" are favourable and should be taken into account in considering further growth of the structure.

In the packing of type 1 there are eight favourable conformations differing one from another by rotations around the axes or by small shifts of helices in the packing. It should be noted that each conformation

Table 4
All possible conformations of the crystallization center F–G–H with 10–11 dehydrated bulky hydrophobic groups

Type of packing	Conformation	Dehydrated Bulky Hydrophobic Groups on Helices			Total number of dehydr. groups	Number of Val among them	
		F	G	H			
1	1a	1,4	2, 5, 13, 16	8,12,14,15,19	11	0	
	1b	–"–	– " –	8, 14,15,19	10	0	
	1c	–"–	2, 5, 13	8,12,14,15,19	10	0	
	1d	–"–	2, 5, 16	– " –	10	0	
	1e	–"–	5, 12, 16	– " –	10	0	
	1f	–"–	4,5, 8,12	– " –	11	0	
	1g	–"–	– " –	8, 14,15,19	10	0	
	1h	–"–	4, 8,12, 15	8,12,14,15,19	11	1	
	1i	–"–	– " –	8, 14,15,19	10	1	–
	1j	–"–	4, 8, 15	8,12,14,15,19	10	1	–
2	2a	1,4	2,4, 7, 13, 16	8,12, 15,19	11	0	
	2b	1,	– " –	– " –	10	0	
	2c	4	– " –	– " –	10	0	
	2d	1,4	– " –	8, 15,19	10	0	
	2e	–"–	– " –	8,12, 19	10	0	
3	3a	1,4	2, 5, 13, 16	8,12, 15,19	10	0	
	3b	–"–	2, 5, 8, 16	– " –	10	0	
	3c	–"–	4,5, 8,12, 16	– " –	11	0	
	3d	–"–	5, 8,12, 16	– " –	10	0	
	3e	–"–	4,5, 8,12,	– " –	10	0	
	3f	–"–	4, 8,12, 15	– " –	10	1	–
	3g	–"–	4, 7,8, 15	– " –	10	1	–
4	4a	1,4	2, 5, 12, 16	8,12, 15,19,23	11	0	
	4b	–"–	– " –	8,12, 15,19,	10	0	
	4c	–"–	– " –	8, 15,19,23	10	0	
	4d	–"–	– " –	8,12, 19,23	10	0	
	4e	–"–	2, 5, 8,12	8,12, 15,19,23	11	0	
	4f	–"–	– " –	8,12, 15,19	10	0	
	4g	–"–	– " –	8, 15,19,23	10	0	
	4h	–"–	– " –	8,12, 19,23	10	0	
	4i	–"–	5, 8,12, 15	8,12, 15,19,23	11	1	
	4j	–"–	– " –	8,12, 15,19	10	1	–
	4k	–"–	– " –	8, 15,19,23	10	1	–
	4l	–"–	– " –	8,12, 19,23	10	1	–
	4m	–"–	4,5, 8, 15	8,12, 15,19,23	11	1	
	4n	–"–	– " –	8,12, 15,19	10	1	–
	4o	–"–	– " –	8, 15,19,23	10	1	–
	4p	–"–	– " –	8,12, 19,23	10	1	–
5	5a	1,4	2, 5, 12, 16	8,12, 15,19	10	0	
	5b	–"–	2, 5, 8,12	– " –	10	0	
	5c	–"–	5, 8,12, 15	– " –	10	1	–
	5d	–"–	4,5, 8, 15	– " –	10	1	–
6	6a	1,4	2,4,5,7, 15	8,12, 15,19	11	1	–
	6b	–"–	2,4 8, 15	– " –	10	1	–

Designations the same as in table 2.

with eleven dehydrated bulky groups among which there is no Val can "generate" several conformations with ten dehydrated bulky groups differing from the "generating" conformation by the exposure of one additional hydrophobic group on one of the helices due to *small* rotations or shifts of this helix. In the packing of type 2 there are five favourable conformations, four of them are "generated" by conformation 2a. Packing 3 has five favourable conformations, packing 4 has ten and packing 5 has two favourable conformations. The enumerated conformations represent a complete set of favourable conformations of the center of crystallization F–G–H.

All the conformations for any other structure can be obtained in an analogous way.

References

[1] C. Levinthal, J. Chim. Phys. 65 (1968) 44.
[2] P.N. Lewis, F.A. Momany and H.A. Scheraga, Proc. Natl. Acad. Sci. USA 65 (1971) 2293.
[3] O.B. Ptitsyn, V.I. Lim and A.V. Finkelstein, in: Proc. VIIIth FEBS Meeting, Analysis and Simulation of Biochemical Systems, eds. H. Hess and H.C. Hemker (North-Holland, Amsterdam, 1972) p. 421.
[4] C.B. Anfinsen, Biochem. J. 128 (1972) 737.
[5] O.B. Ptitsyn, Dokl. Akad. Nauk SSSR 210 (1973) 1213.
[6] P.N. Lewis, N. Gō, D. Kotelchuk and H.A. Scheraga, Proc. Natl. Acad. Sci. USA 65 (1970) 810.
[7] P.N. Lewis and H.A. Scheraga, Arch. Biochem. Biophys. 144 (1972) 576, 588.
[8] O.B. Ptitsyn, A.I. Denesuyk, A.V. Finkelstein and V.I. Lim, FEBS Letters 34 (1973) 55.
[9] A.I. Denesyuk, O.B. Ptitsyn and A.V. Finkelstein, Biofizika (USSR) 19 (1974) 549.
[10] A.V. Finkelstein and O.B. Ptitsyn, Biopolymers, in press.
[11] A.V. Finkelstein, O.B. Ptitsyn and S.A. Kozitsyn, Biopolymers, in press.
[12] P.J. Flory, Statistical mechanics of chain molecules (Interscience, New York, 1969).
[13] D.M. Blow and T.A. Steitz, Ann. Rev. Biochem. 39 (1970) 63.
[14] M.F. Perutz, J. Mol. Biol. 13 (1965) 646.
[15] H.C. Watson, Progr. Stereochem. 4 (1968) 299.
[16] B.P. Schoenborn, Cold Spring Harbor Symp. Quant. Biol. 36 (1972) 569.
[17] J.C. Kendrew, H.C. Watson, B.E. Strendberg, R.E. Dickerson, D.C. Phillips and V.C. Shove, Nature 190 (1961) 663.
[18] S.C. Harrison and E.R. Blout, J. Biol. Chem. 240 (1965) 299.
[19] M.J. Crumpton and A. Polson, J. Mol. Biol. 11 (1965) 722.
[20] K. Nishikawa, T. Ooi, Y. Isogai and N. Saitô, J. Phys. Soc. Japan 32 (1972) 1331.
[21] N. Saitô, Y. Isogai and K. Kosuge, Abstracts of Contributed Papers, IV Biophys. Congress, Moscow, 2 (1972) 32. EVI a 1/6.
[22] S.A. Kozitsyn and O.B. Ptitsyn, Molek. Biol. (USSR) 8 (1974) 536.
[23] O.B. Ptitsyn and A.A. Rashin, Dokl. Akad. Nauk SSSR 213 (1973) 473.
[24] Ch. Tanford, J. Amer. Chem. Soc. 84 (1962) 4240.
[25] M.O. Dayhoff, in: Atlas of protein sequence and structure, Vol. 5, ed. M.O. Dayhoff (NBRF, Silver Spring, Md., 1972).
[26] V.I. Lim, Biofizika (USSR) 19 (1974) 562.
[27] V.I. Lim, J. Mol. Biol. 88 (1974) 857, 873.