O. B. Ptitsyn, *Physical principles of protein structures.* Protein molecules are the most complicated of known macromolecules and at the same time one of the simplest biological objects. For this reason, the study of protein structure permits observing how nature uses the laws of molecular physics to create biologically useful systems. From the chemical point of view, a protein is a copolymer consisting of 20 different types of monomers (amino acids). All monomers (except one) have the same main chain and differ only by their side groups. The sequence of monomers in a protein (the chemical structure of the protein) is determined by the chemical structure of DNA in accordance with the well-known genetic code. In its turn, the chemical structure of a protein determines its spatial structure, which determines its biological function.

Due to the successes of x-ray structural analysis of protein crystals, the spatial structures of more than 100 different proteins are now known. Their chains are coiled into compact spherical or slightly stretched globules, whose structure at first glance looks very complicated and irregular. However, careful analysis has shown that the only complicated, irregular, and specific (for each protein) property is the *detailed* spatial structure of the protein (corresponding to close packing of atoms). The coarse structures of proteins, however, turned out to be surprisingly simple and regular. The main chain of the protein is made up of regular sections of two types, stabilized by hydrogen bonds between the atoms of the main chain: tightly coiled helices and stretched sections (Fig. 1). The average length of the sections of both types is ~ 15 Å, which is close to the radius of a typical protein globule. These sections are arranged in a regular manner in two or (more rarely) three layers and are connected by sections of the protein chain, not having a regular structure. In this case, proteins with the most varied chemical structure are characterized by similar arrangements of helical and stretched sections. This raises the following interesting question: how does the irregular and specific, for each protein, chemical structure lead to regular spatial structures that are similar for different proteins?

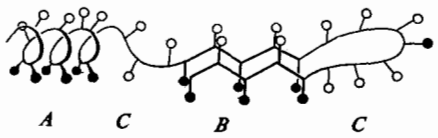

FIG. 1. Helical (*A*), stretched (*B*), and irregular (*C*) sections of a protein chain. The dark and light colored circles indicate nonpolar and polar side groups, respectively.

The attachment of regular sections in two or three layer complexes is due to the presence in the protein of approximately identical quantities of polar and nonpolar side groups, bound by a single main chain. Contact between the nonpolar groups and water increases the ordering of the water, which increases its free energy. This creates an effective "hydrophobic" attraction between the nonpolar groups in water, acting over distances not significantly greater than the usual van der Waals attraction. In order that the helical and stretched sections be able to form a regular complex, each of these sections must have at least one continuous nonpolar surface, which imposes limitations on the sequence of polar and nonpolar groups in such sections. It turned out, however, that the sequences of polar and nonpolar groups (specific for each protein), on being averaged over many proteins, are not distinguishable from random sequences. At the

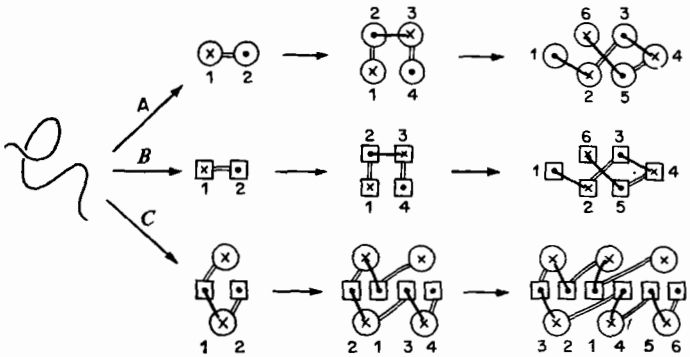

FIG. 2. Diagram showing the packing of chains consisting of helical (*A*), stretched (*B*), and alternating helical and stretched (*C*) regular sections. The helices are indicated by circles and the stretched sections are indicated by squares (view from above). The crosses and the dots indicate that the section is oriented away from and toward the observer. The numbers in *A* enumerate helical sections, while the numbers in *B* and *C* enumerate stretched sections. The length of the chains increases from left to right.

same time, an analysis of random sequences has shown that clusters of nonpolar groups capable of forming helical or stretched sections with continuous nonpolar surfaces can appear quite often in such sequences. Clusters of nonpolar groups arise with the synthesis of protein chains and then can no longer decompose, since they are fixed by chemical bonds. As a result, irregular (random) sequences of polar and nonpolar groups form regular packings of helical and stretched sections, similar to packings that chains consisting entirely of nonpolar groups would have. The energetically most favorable packings of helical and stretched sections in such chains can be easily obtained (Fig. 2), and they are surprisingly close to the typical packings of these sections in real proteins.

In contrast to the regular coarse structure of a protein, its detailed irregular structure is stabilized by the usual van der Waals interactions, leading to dense packing of the side groups. The presence of a regular "framework" consisting of helical and stretched sections, encompassing the entire molecule, makes any local breakdown of dense packing impossible, since such breakdown requires a displacement or breakdown of the entire framework. For this reason, breakdown of a densely packed protein structure is an intramolecular first-order phase transition. This fact has a great biological significance, since it makes the rigid structure of the protein stable with respect to thermal fluctuations. The rigid structure of the "active center" of the protein (i.e., the part of the protein that participates directly in its function) is guaranteed by the fact that this center is included in an extended highly efficient structure, capable of breaking down only as a whole. The architectural organization of this structure is not very important for biological function; the fact that it exists is in itself important. Such extended structures are then formed in random sequences of amino acid radicals almost without the participation of biological evolution. Biological evolution is primarily needed for "editing" these structures in order to create active centers based on them, capable of performing different biological functions.

This report was based on the following publications by the author and his colleagues:

[1]O. B. Ptitsyn, Vestn. Akad. Nauk SSSR, No. 5, 57 (1973).
[2]O. B. Ptitsyn and A. V. Finkelstein, Quart. Rev. Biophys. **13**, 339 (1980).
[3]O. B. Ptitsyn and A. V. Finkelstein in: Protein Folding, Ed. by R. Jaenicke, Elsevier North-Holland Biomed. Press, Amsterdam (1980), p. 101.
[4]O. B. Ptitsyn, FEBS Lett. **131**, 197 (1981).
[5]R. I. Gil'manshin, D. A. Dolgikh, O. B. Ptitsyn, A. V. Finkel'shtein, and E. I. Shakhnovich, Biofiz. **27**, 1006 (1982).
[6]E. I. Shakhnovich and A. V. Finkel'shtein, Dokl. Aka. Nauk SSSR **267**, 1247 (1982).
[7]O. B. Ptitsyn in: Conformation in Biology, Ed. by R. Srinivasan and R. H. Sarma, Adenine Press, New York (1983), p. 49.
[8]O. B. Ptitsyn and A. V. Finkelstein, Biopolymers **22**, 15 (1983).