Mössbaun Spectroscopy in Biological Systems Proceedings (1969)
Univ. of Illinois Bulletin
67, No. 41 pp. 22-24(1969)
Urbana, IL 61801

HOW TO FOLD GRACIOUSLY

Chairman: R.A. Marcus
Speaker: Cyrus Levinthal
Notes by: A. Rawitch
Retranscribed: B. Krantz

Proteins are macromolecules which possess several unique properties. They are very large (containing 2,000 or more atoms) and complex. Their structures show no obvious regularity but a very subtle regularity is apparent upon close examination. We know from the fact that proteins may be crystallized and further from x-ray crystallography that each atom occupies a unique place in the relative 3-dimensional space of the molecule. If we consider a protein containing 2,000 atoms with no structural restrictions, such a macromolecule would possess 6,000 degrees of freedom. We know, however, from x-ray studies and other techniques as well, that there are indeed certain structural restrictions in a polypeptide structure. For example, if we schematically indicate a polypeptide chain as in Figure 1, we find that the 6 atoms in each unit indicated by the dotted lines lie in a common plane. Considerations of such factors allow us to predict only 450

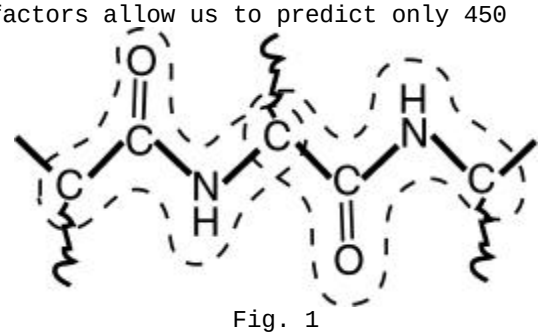


Fig. 1

degrees of freedom in a protein structure containing 150 amino acids, for example. Of these 450 degrees of freedom, 300 would be due to rotations and 150 would be due to relative bond angles of the side chains.

We know that we can take such a structure, i.e., a protein, and place it in an environment which causes unfolding to a random coil (no secondary structure or tertiary structure) and when we restore the protein to an aqueous medium at neutral pH containing some salt or buffer, the protein will refold to its original structure as judged by biological activity and physical parameters. Such experiments have been supported fully by experiments with synthetic polypeptides and, more recently, a synthetic enzyme.

Let us ask ourselves how proteins fold to give such a unique structure. By going to a state of lowest free energy? Most people would say yes and indeed, this is a very logical assumption. On the other hand, let us consider the possibility that this is not so. We began thinking along these lines several years ago while we were attempting to predict the 3-dimensional structure of some polypeptides from primary sequence information.

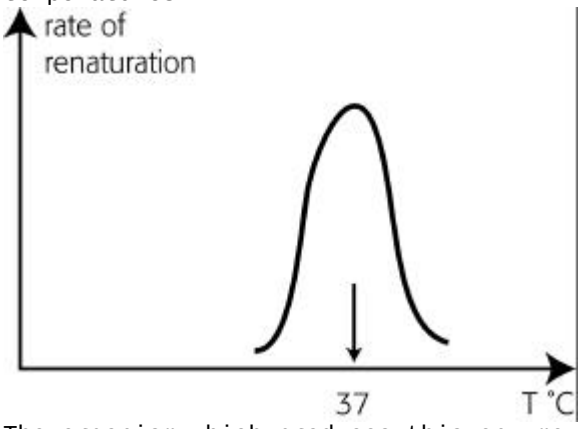
If we begin with a set of bond angles and bond lengths and go to 3-dimensional coordinates (via vector matrix multiplications), we can build a 3-dimensional image and display it on a computer controlled oscilloscope. If we know the coordinates of any two atoms and their interaction energy functions, could we extend this treatment to sum the total energy of a given polypeptide or protein structure?

Well, let us consider the various parameters involved. How accurately must we know the bond angles to be able to estimate these energies? Even if we knew these angles to better than a tenth of a radian, there would be 10^{300} possible configurations in our theoretical protein. In nature, proteins apparently do not sample all of these possible configurations since they fold in a few seconds, and even postulating a minimum time from one conformation to another, the proteins would have to try on the order of 10^{8} different conformations at most before reaching their final state. We feel that protein folding is speeded and guided by the rapid formation of local interactions which then determine the further folding of the polypeptide. This suggests that local amino acid

sequences which form stable interactions and serve as nucleation points in the folding process.

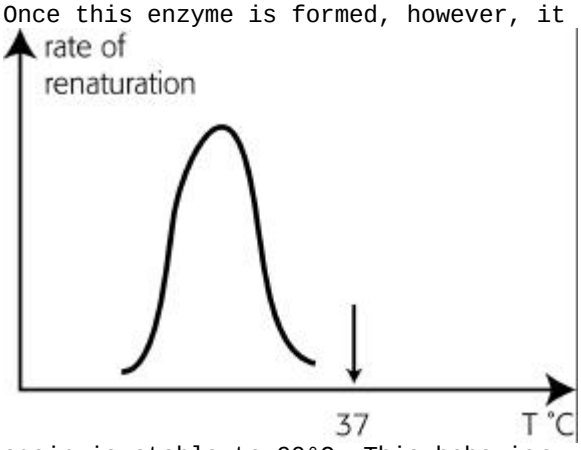
Then, is the final conformation necessarily the one of lowest free energy? We do not feel that it has to be. It obviously must be a metastable state which is in a sufficiently deep energy well to survive the possible perturbations in a biological system. If it is the lowest energy state, we feel it must be the result of biological evolution; i.e., the first deep metastable trough reached during evolution happened to be the lowest energy state. You may then ask the question, "Is a unique folding necessary for any random 150-amino acid sequence?" and I would answer, "Probably not." Some experimental support for this statement comes from the difficulty many of us are all too aware of in attempting to crystallize proteins.

I would like to illustrate some of these points by telling you about some work we have done on an alkaline phosphatase enzyme. This enzyme has a molecular weight of 40,000 and consists of two similar or identical subunits. We have unfolded this enzyme and then followed the rate of renaturation under appropriate conditions as a function of temperature. As can be seen in the figure below, the optimum rate of renaturation occurs at 37°C and falls rapidly at higher and lower temperatures.



The organism which produces this enzyme grows optimally at 37°C also. Although the renaturation rate drops off above 37°C, the native intact enzyme, or the refolded enzyme is stable up to 90°C. Thus, once the folding is complete, the resulting structure is quite stable.

We have isolated mutants of this organism which produce active enzyme only when grown at temperatures below 37°C and we have found that the protein renatures only at temperatures below 37°C, as shown in the figure below. Once this enzyme is formed, however, it



again is stable to 90°C. This behavior is obviously not expected in an equilibrium situation. As it turns out upon closer study, the limiting rate in the formation of active enzyme is the formation of the dimeric species of the enzyme. We can, however, say that at least in the assembly of protein subunits, it matters in which order what events occur.

We may be helped ultimately by sufficient data from x-ray crystallographic work to find clues as to the kinds of local interactions which are most important in protein folding.

What then, can we derive from computer calculations? We know very accurately:

1. bond lengths in polypeptides,
2. planar groupings in the polypeptide structure.

For small molecules, it is possible to analyze x-ray diffraction data by means of the direct methods. For large molecules, this is generally beyond our ability at present and we must obtain phase information in order to reconstruct the reflected intensities. We hope to look for reflections from certain postulated substructures by having our computer search in Fourier space for such groupings and then refine these data by means of the tangent formula and then relate other intensities to these.

Professor Levinthal then showed a short motion picture which illustrated the synthesis of polypeptide structure and the process of then forming a desired interaction via the most favored energy path as displayed on the computer controlled oscilloscope. The relevance of these studies to Mössbauer spectroscopy may be in the understanding of small perturbations of polypeptide structures and their effect on the Mössbauer nucleus.

Discussion:

Q: Is a protein really ever truly unfolded, i.e., devoid of secondary and tertiary structure?

A: Both physical measurements and synthetic polypeptide work suggest the answer is yes.

Q: The tangent formula requires phase information first in order to refine the data. Are you implying this is not the case?

A: Since we are looking for known substructures within the patterns, we can use the tangent formula.

Q: Have you used your method to produce a known structure and looked for the most likely thermal perturbation of the structure?

A: No, we haven't done calculations of that sort.