A model of the "molecular vector machine" for protein folding

Vladimir.A. Karasev, Victor V. Luchinin

St.-Petersburg State Electrotechnical University, Prof. Popov str. 5, St.-Petersburg, 197376 Russia
e-mail:cmid@eltech.ru

Vasily E. Stefanov
Department of Biochemistry, St.Petersburg State University, Universitetskaya nab. 7/9, St.-Petersburg, 199034 Russia
e-mail: vastef@mail.ru

Earlier [1] we proposed a dodecahedral model of the canonical set of 20 amino acids based on antisymmetry of the side chains. Later [2] the model was refined. An additional anti-symmetry plane (III) was introduced (Fig.1, a). Side chains were arranged from the top to the bottom in the increasing order of their size. Shorter side chains are situated on the right whereas their heavier analogs on the left from the plane I. In this model four groups of chains can be distinguished: 1) chains symmetrical about plane I (e.g. Ser:Thr, etc.), 2) chains symmetrical about plane II (e.g. Ser:Cys, etc.), 3) chains symmetrical about plane III (e.g. Ser:His) and 4) chains symmetrical about the center of the dodecahedron (e.g. Ser:Trp). The aim of the present work was to search ways of practical application of the above model.

According to [3], the region of $N_iH\ldots O_{i-4}=C_{i-4}$ hydrogen bond in the 4-link protein cycle is the target of physical operators (amino acid side chains) which reconstruct the triplet-encoded structure. Following the principles assumed in the model we introduced three planes in the above region and considered 20 vectors representing the action exerted on this bond [2]. The model of the amino acid structure was brought into this region (Fig.1, b), the NH-group being placed into the vertex Gly and atom O_{i-4} into the center of the dodecahedron, so that the above vectors were directed to the dodecahedron vertices corresponding to the names of the side chains.

Thus, the amino acid side chains can be regarded as irreducible representations of the group composed by the vectors (dodecahedron diameters) and the structure itself as "molecular vector machine". The model accounts for certain features of the canonical set of amino acids, e.g. difference in the length of structurally similar Asp and Glu or oppositely charged Asp and Arg can be related to the effect of different orientation of vectors. Cyclic character of the side chains His, Trp, Phe, Tyr is likely to be associated with the rigid structure, which is required for the function of the vectors in the lower face of the dodecahedron could realize, etc.

The molecular vector machine (Fig.1, b) consists of two parts: dodecahedron and tetrahedron (block of C^{α}_i atom). Side chains connected with C^{α}_i (R_i) are directed towards the corresponding vertices of the dodecahedron and generate the encoded structure [3], whereas concomitant turn of the tetrahedron specifies the direction of the bond with the atom C^{α}_{i+1} of the chain [2]. We presume that

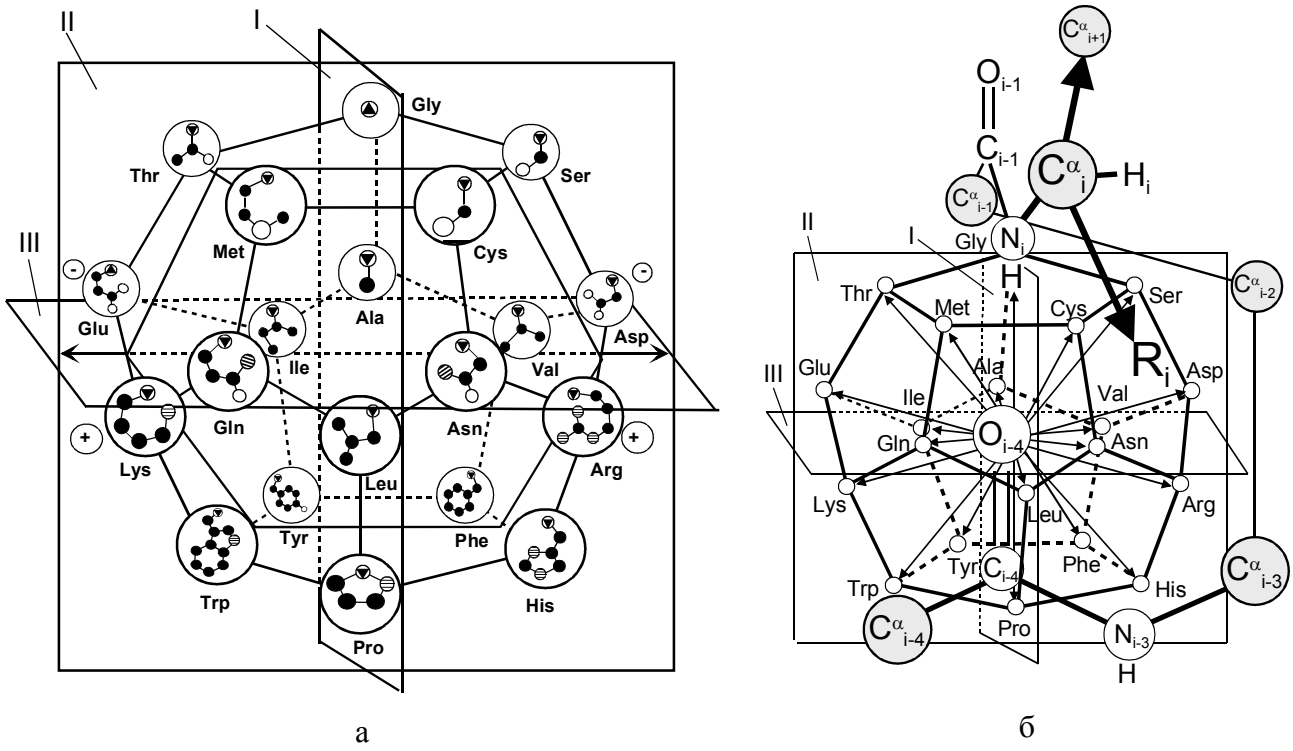

Fig. 1. Construction of the model of "molecular vector machine" for co-translational protein folding.

а – dodecahedral model of the canonical set of amino acids

I , II, III – symmetry planes. Side chains are oriented downwards; α-carbon atom is symbolized by triangle in the circle

б – transfer of the structure of the amino acids' canonical set into the region of the of $N_iH\ldots O_{i-4}=C_{i-4}$ hydrogen bond in the 4-link protein cycle.

operation of the molecular vector machine connected with the chain's growth, can provide co-translational protein folding. At present the model is being tested on different α-helical and β-sheeted fragments available from PDB.

References

1. V.A. Karasev et al. (2005). A dodecahedron-based model of spatial representation of the canonical set of amino acids. In: Mathematical Biology and Medicine, **Vol.8**. International Conference. "Advances in Bioinformatics and Its Applications", World Scientific Publ. Co., 482-493.

2. V.A. Karasev (2007). Amino acids of the canonical set as irreducible representation of the group of vectors – dodecahedron diameters. Dep. VINITI 25.04.07, № 461-B2007

3. V.A. Karasev, V.E. Stefanov (2001). Topological nature of the genetic code. J.Theor.Biol. **209:**303-317.