Assessment of applicability of the program Rosetta for the determination of membrane associated proteins structure. M1 influenza virus protein

E.N. BOGACHEVA*, A.L. KSENOFONTOV**, A.V. SHISHKOV*

*N.N. Semenov Institute of Chemical Physics Russian Academy of Sciences,
ul. Kosygina, 4, Moscow, 119991 Russia, e-mail: ben@chph.ras.ru
**Belozersky Institute of Physico-Chemical Biology of Moscow State University,
Leninskie Gory, 1, Moscow, 119992 Russia: e-mail: ksenofon@belozersky.msu.ru

If a protein is a part of multicomponent biological complexes such as viruses the reconstruction of proteins spatial structure remains extremely difficul and actual problem. The matrix M1 protein underlying the membrane of virion is the major structural component of influenza A virus. The atomic structure of the N-terminal two thirds of M1 protein was solved at acid and neutral pH [1].

We analyzed the structure of the M1 protein of the influenza virus A/Puerto Rico/8/34 (H1N1) strain in acidic solution using tritium planigraphy [2, 3]. The incorporation of tritium label into the domains of the M1 protein were studied. It was established that the C domain and the interdomain loops are preferentially accessible to tritium [4]. Bioinformatic analysis of the M1 protein sequence revealed intrinsically unstructured segments that were concentrated in the C domain and interdomain loops between the N-, M-, and C domains.

We've developed the computer algorithm imitating the anisotropic conditions of the bombardment of proteins in a membrane surrounding with the proper account of the protein molecule orientation in relation to the membrane surface for the beam of "hot" tritium atoms.

The preliminary model of spatial structure of M1 protein as a component of influenza virus was proposed [4]. This model is based on the data obtained by tritium labeling of intact virions and free M1 protein, theoretical prediction of the C-terminal domain secondary structure for M1 protein, and application of the developed computer algorithm.

The experimental and theoretical data obtained by tritium bombardment and simulation algorithm were compared with the Rosetta program prediction of the C-domain three-dimensional structure [5]. Analysis of the Rosetta algorithms on NM-domain has shown an opportunity of the tritium planigraphy experimental data usage for more correct construction 3D structures. The correlation between the methods were allocated but ussially coefficient of approximation for C domain varied in the interval 0.5-0.8. The application of the combined approach allowed reducing substantially the hypothetically possible spatial structures of the C-domain, but accuracy of

predictions was low.

In the present study we also utilised a number of different programs for intrinsic disorder prediction to study the structural characteristics of the full-length M1 protein. The results of the analysis of the M1 ordered/disordered regions using order/disorder predictors are presented. The regions in the M1 protein sequence that were highly accessible to tritium labelling largely coincided with the disordered regions predicted by computational analysis. Notably, all six algorithms similarly predicted disordered interdomain loops as well as high levels of disorder of the C-terminal domain. However, the predicted lengths of the disordered regions varied. Based on the results of the prediction algorithms, we suggest that pH-dependent structural change occurs not only in the M1 interdomain linker region [5] but also within the entire C-terminal domain, leading to the destruction of the M1 membranebound shell followed by the uncoating of the virions.

Comparison of data of a tritium planigraphy method and results of a prediction of disorder areas allowed to increase correlation of experimental and theoretical data to 0.8 and to specify protein structure. It is necessary to emphasize that application of the Rosetta program can lead to mistakes in case of a reconstruction of spatial structure of membrane and amiloid proteins as in a statistical set of the last water-soluble globular proteins are presented generally.

1.   B.Sha, M.Luo (1997) Structure of a bifunctional membrane-RNA binding protein, influenza virus matrix protein M1. *Nat. Struct. Biol.* **4:**239–244.

2.   L.A.Baratova, E.N.Bogacheva, V.I.Goldanskii, V.A.Kolb, A.S.Spirin, A.V.Shishkov (1999) Tritium planigraphy of biological macromolecules. Moscow.: Nauka, 175p.

3.   E.N.Bogacheva, V.I.Goldanskii, A.V.Shishkov, A.V.Galkin and L.A.Baratova (1998) Tritium planigraphy: from the accessible surface to the spatial structure of a protein. *Proc. Natl. Acad. Sci. USA.* **95:**2790–2794.

4.   Alexander Shishkov, Elena Bogacheva, Natalia Fedorova et all. 2011. Spatial structure peculiarities of influenza A virus matrix M1 protein in an acidic solution that simulates the internal lysosomal medium. FEBS J., v. 278, p. 4905–4916.

5.   A.L. Ksenofontov, E.N. Dobrov, N.V. Fedorova et all. (2011). Intrinsically Unstructured Regions in the C Domain of the Influenza Virus M1 Protein. Molecular Biology, **45**, No. 4, 634–640.