Graphical Representation of Proteins as Four-Color Maps and Their Numerical Characterization

Milan Randić[1], Ketij Mehulić[2], Damir Vukičević[3], Tomaž Pisanski[4], Dražen Vikić-Topić[5], Dejan Plavšić[5]

[1]*National Institute of Chemistry, P.O. Box 3430, 1001 Ljubljana, Slovenia,* [2]*School of Dental Medicine, University of Zagreb, Gundulićeva 5, 10000 Zagreb, Croatia,* [3]*Department of Mathematics, University of Split, Nikole Tesle 12, 21000 Split, Croatia,* [4]*IMFM, Department of Theoretical Computer Science, University of Ljubljana, Jadranska 19, 1000 Ljubljana, Slovenia,* [5]*Ruđer Bošković Institute, NMR Center, P.O.Box 180, HR-10002 Zagreb, Croatia*

We put forward a novel compact 2-D graphical representation of proteins based on the concept of virtual genetic code and a four-color map. The novel graphical representation uniquely represents proteins and allows one to easily and quickly visually observe and inspect similarity/dissimilarity between them. It also leads to a novel protein descriptor being a 10-vector derived from a novel structure matrix ***S*** associated with the map. The approach is illustrated with the A-chain of human insulin and the A-chain of human insulin analogue glargine.