Modeling of proteins structures with the coarse–grained UNRES force field in the CASP14 experiment

Anna Antoniak[1*], Patryk A. Wesołowski[1], Adam K. Sieradzan[1], Cezary Czaplewski[1], Emilia A. Lubecka[2], Agnieszka G. Lipska[1], Artur Giełdoń[1], Rafał Ślusarz[1], Paweł Krupa[3], Mateusz Kogut[1], Iga Biskupek[1], Krzysztof K. Bojarski[1], Małgorzata M. Kogut[1], Mateusz Marcisz[1], Martyna Maszota-Zieleniak[1], Sergey A. Samsonov[1], Magdalena J. Ślusarz[1], Karolina Zięba[1], Adam Liwo[1]

1 - Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland,
2 - Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, G. Narutowicza 11/12, 80-233 Gdańsk, Poland,
3 - Institute of Physics, Polish Academy of Sciences, Aleja Lotników 32/46, Warsaw, PL-02668, Poland

anna.antoniak@etoh.chem.ug.edu.pl

Abstract

The UNRES force field was tested in the CASP14 experiment of protein-structure prediction, in which larger monomeric and multimeric targets were present compared to previous editions. Three prediction modes were tested (i) *ab initio* (the UNRES group), (ii) contact-assisted (the UNRES-contact group), and (iii) template-assisted (the UNRES-template group). The average Global Distance Test Total Score (GDT_TS) of the 'Model 1' predictions were 29.17, 39.32, and 56.37 for the UNRES, UNRES-contact, and UNRES-template predictions, respectively, increasing by 0.53, 2.24, and 3.76, respectively, compared to CASP13. It was also found that the GDT_TS of the UNRES models obtained in *ab initio* mode and in the contact-assisted mode decreases with the square root of chain length, while the exponent in this relationship is 0.20 for the UNRES-template group models and 0.11 for the best performing AlphaFold2 models, which suggests that incorporation of database information, which stems from protein evolution, brings in long-range correlations, thus enabling the correction of force-field inaccuracies.

Scheme of prediction of CASP14 targets

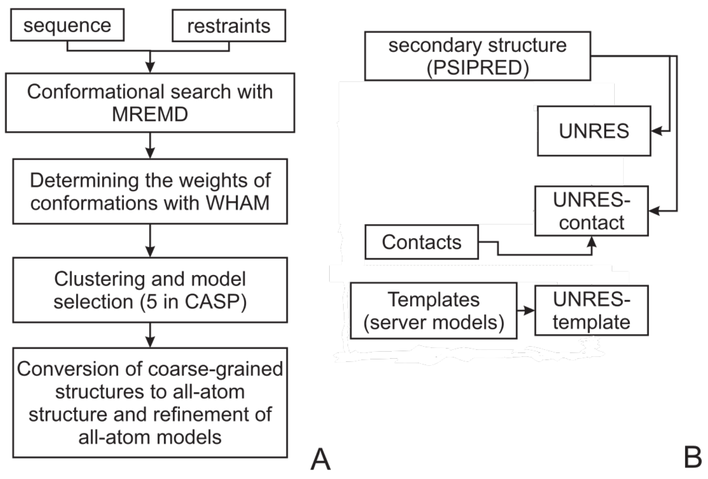

Figure 1: Scheme of prediction of CASP14 targets by the UNRES, UNRES-contact, and UNRES-template groups. Panel A summarizes the calculation flow and panel B shows the restraints used by the respective UNRES-based groups. The arrows in panel B connect the restraints to the groups which applied them.

UNRES model and force field

Coarse grained models seem to be the right to handle large systems and time scales [1,2,3].In order to prove that they yiels reliable results, assessment against the experimental data is necessary. One of the hardest tests is the biannual Community Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction (CASP). In this communication, we report the results of the evaluation of the UNited RESidue (UNRES) force field developed in our laboratory [4,5], with Multiplexed Replica Exchange Molecular Dynamics (MREMD) as the conformational-search engine [6-8] in the CASP14 experiment.

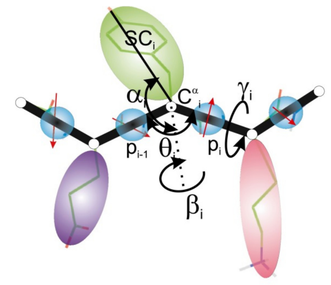

Figure 2: The UNRES model of polypeptide chains. Blue spheres: peptide groups (p); spheroids: side chains (SC); white spheres: α-carbon atoms (which are not interaction sites but only serve to assist in chain-geometry definition).

GDT TS

The violin plots of the Global Distance Test Total Score (GDT_TS) distributions of the first, the best, and all models, where the maximum, the minimum, and the mean values, as well as the approximate distributions are displayed, for all EU categories (TBM-easy, TBM-hard, FM/TBM, FM, MultiDom, and D0)[6,7,8].

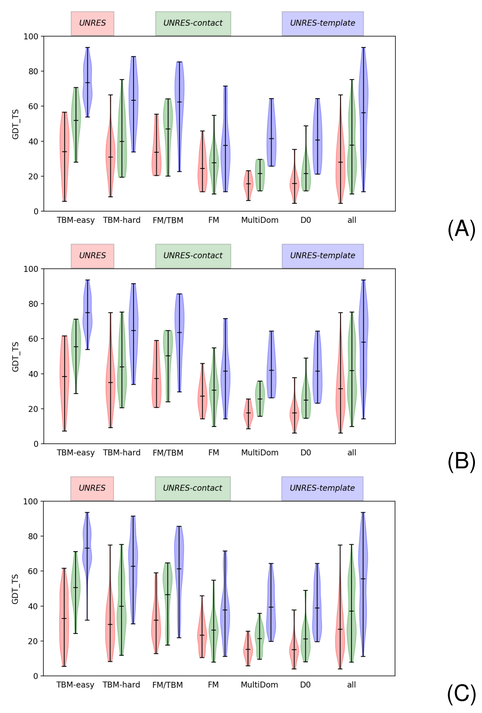

Figure 3: Violin plots of the GDT_TS of the 'Model 1' (A) and best (B) predictions, and GDT_TS averaged over all models (C) of the respective CASP14 EU types and all EUs, for the UNRES, UNRES-contact, and UNRES-template groups. The mean, minimum, and the maximum values are shown as black horizontal lines, the GDT_TS ranges are shown as vertical lines, and approximate distributions are shown as color-filled shapes.

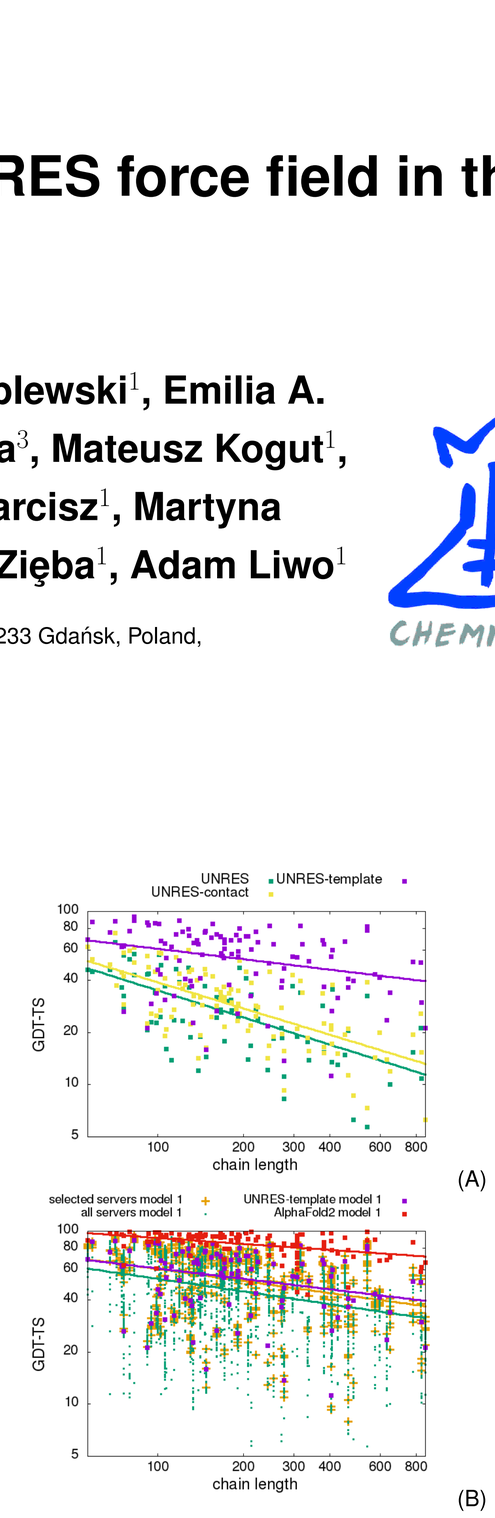

Figure 4 : Plots of the GDT_TS of the 'Model 1' predictions of the CASP14 EUs obtained by the UNRES, UNRES-contact, and UNRES-template groups (A) and those from the server models selected to derive the restraints for UNRES-template predictions, all servers, UNRES-template, and AlphaFold2 (the winner group) (B). The logarithmic scale is used on both axes. The straight lines correspond to fitting the power function to the respective points.

Prediction example of UNRES-based groups

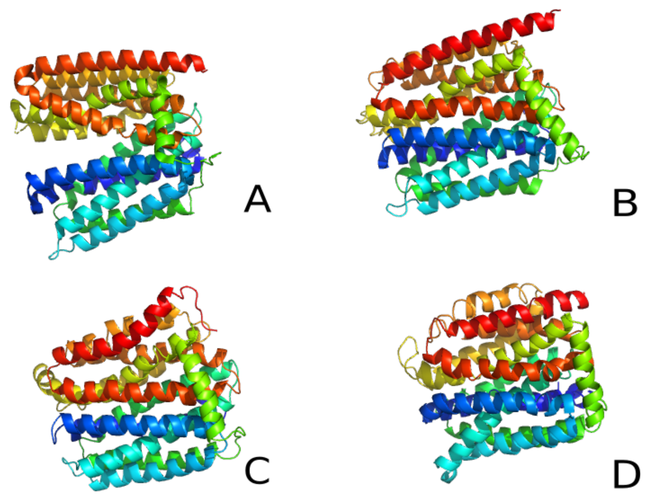

Figure 5 : The experimental structure of target T1024 (PDB: 6T1Z) A, the best UNRES-template model (model 4; GDT_TS=60.23, RMSD=3.22 Å), the best UNRES-contact model (model 2; GDT_TS=49.04, RMSD=4.974 Å), and the best UNRES model (model 2; GDT_TS=37.79, RMSD=6.13 Å). All structures are rainbow-colored from the N- to the C-terminus.

Acknowledgments

Supported by grants UMO-2017/25/B/ST4/01026, UMO-2017/27/B/ST4/00926, UMO-2017/26/M/ST4/00044 from the National Science Center of Poland (Narodowe Centrum Nauki). Computational resources were provided by (a) the Interdisciplinary Center of Mathematical and Computer Modeling (ICM) the University of Warsaw under grant No. GA76-11, (b) the Centre of Informatics - Tricity Academic Supercomputer & Network (CI TASK) in Gdańsk, (c) the Academic Computer Centre Cyfronet AGH in Krakow under grants: asunres18 and asunres19, and (d) our 796-processor Beowulf cluster at the Faculty of Chemistry, University of Gdańsk. [a0]sciposter

References

[1] Voth, G., Coarse-Graining of Condensed Phase and Biomolecular Systems, CRC Press, Taylor and Francis Group, 1st edition (2008).
[2] Marrink, S. J. and Tieleman, D. P., Chem. Soc. Rev., 42, 6801–6822 (2013).
[3] Kmiecik, S, Gront, D, Koliński, M, Wieteska, L, Dawid, A and Koliński, A, Chem. Rev., 116, 7898–7936 (2016).
[4] Liwo, A., Baranowski, M., Czaplewski, C., Gołaś, E., He, Y., Jagieła, D., Krupa, P., Maciejczyk, M., Makowski, M., Mozolewska, M. A., Niadzvedtski, A., Ołdziej, S., Scheraga, H. A., Sieradzan, A. K., Ślusarz, R., Wirecki, T., Yin, Y. and Zaborowski, B., J. Mol. Model., 20, 2306 (2014).
[5] Liwo, A., Sieradzan, A. K., Lipska, A. G., Czaplewski, C., Joung, I., Żmudzińska, W., Hałabis, A. and Ołdziej, S., J. Chem. Phys., 150, 155104 (2019).
[6] M. Khalili, A. Liwo, A. Jagielska, H. A. Scheraga, Molecular dynamics with the united-residue (UNRES) model of polypeptide chains. II. Langevin and Berendsen-bath dynamics and tests on model α-helical systems, J. Phys. Chem. B 109 (2005) 13798–13810
[7] Y. M. Rhee, V. S. Pande, Multiplexed-replica exchange molecular dynamics method for protein folding simulation, Biophys J. 84 (2003) 775–786.
[8] C. Czaplewski, S. Kalinowski, A. Liwo, H. A. Scheraga, Application of multiplexing replica exchange molecular dynamics method to the UNRES force field: Tests with alpha and alpha + beta proteins, J. Chem. Theory Comput. 5 (2009) 627–640.