STRUCTURE, AFFINITY AND SPECIFICITY RIDDLES IN BIOMOLECULAR INTERACTIONS.

Alexandre M.J.J. Bonvin

Computational Structural Biology Group, Department of Chemistry, Faculty of Science, Utrecht University, 3584CH, Utrecht, The Netherlands.

a.m.j.j.bonvin@uu.nl

Motivation

The prediction of the quaternary structure of biomolecular macromolecules is of paramount importance for fundamental understanding of cellular processes and drug design. In the era of integrative structural biology, one way of increasing the accuracy of modelling methods used to predict the structure of biomolecular complexes is to include as much experimental or predictive information as possible in the process.

Methods

We have developed for this purpose a versatile information-driven docking approach HADDOCK (http://www.bonvinlab.org/software/haddock2.2) [1-3]. HADDOCK can integrate information derived from biochemical, biophysical or bioinformatics methods to enhance sampling, scoring, or both [4]. The information that can be integrated is quite diverse: interface restraints from NMR, mutagenesis experiments, or bioinformatics predictions; shape data from small-angle X-ray scattering [5] and, recently, cryo-electron microscopy experiments [6].

Results

In my talk I will illustrate HADDOCK's capabilities with various examples, including results from our participation to CAPRI [7,8]. HADDOCK has demonstrated sustained prediction and scoring performance since the start of its participation to CAPRI. This is due, in part, to its ability to integrate experimental data and/or bioinformatics information into the modelling process, and also to the overall robustness of the scoring function used to assess and rank the predictions. Thirteen years after the original publication, the HADDOCK scoring function remains a simple linear combination of OPLS [9] intermolecular van der Waals and Coulomb electrostatics energies, an empirically-derived desolvation energy term [10], and one or more restraints energy terms reflecting the agreement between models and experimental/prediction information. Our simple scoring scheme successfully selected acceptable/medium quality models for 18/14 of the 25 targets during the combined CASP-CAPRI prediction round, making us rank at the top. Considering that for only 20 targets acceptable models were generated our effective success rate reaches as high as 90% (18/20)! These results underline the success of our simple but sensible prediction and scoring scheme.

I will end by discussing the problem of binding affinity prediction, showing that current scoring functions in macromolecular docking fail at predicting the affinity of protein-protein complexes and introduce a simple, contact-based predictor that outperforms complex, energy-based predictors [11].

References

1. G.C.P van Zundert, J.P.G.L.M. Rodrigues, M. Trellet, C. Schmitz, P.L. Kastritis, E. Karaca, A.S.J. Melquiond, M. van Dijk, S.J. de Vries and A.M.J.J. Bonvin. The HADDOCK2.2 webserver: User-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.*, **428**, 720-725 (2015).
2. S.J. de Vries, M. van Dijk and A.M.J.J. Bonvin The HADDOCK web server for data-driven biomolecular docking. *Nature Protocols*, **5**, 883-897 (2010).
3. C. Dominguez, R. Boelens and A.M.J.J. Bonvin HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.*, **125**, 1731-1737 (2003).
4. J.P.G.L.M Rodrigues and A.M.J.J. Bonvin Integrative computational modeling of protein interactions. *FEBS J.*, **281**, 1988-2003 (2014).
5. E. Karaca and A.M.J.J. Bonvin. On the usefulness of Ion Mobility Mass Spectrometry and SAXS data in scoring docking decoys. *Acta Cryst. D.*, **D69**, 683-694 (2013).
6. G.C.P. van Zundert, A.S.J. Melquiond and A.M.J.J. Bonvin. Integrative modeling of biomolecular complexes: HADDOCKing with Cryo-EM data. *Structure*. **23**, 949-960 (2015).
7. J.P.G.L.M. Rodrigues, A.S.J. Melquiond, E. Karaca, M. Trellet, M. Van Dijk, G.C.P. Van Zundert, C. Schmitz, S.J. de Vries, A. Bordogna, L. Bonati, P.L. Kastritis and A.M.J.J. Bonvin Defining the limits of homology modelling in information-driven protein docking *Proteins: Struc. Funct. & Bioinformatics*, **81**, 2119-2128 (2013).
8. S.J. de Vries, A.S.J. Melquiond, P.L. Kastritis, E. Karaca, A. Bordogna, M. van Dijk, J.P.G.L.M. Rodrigues and A.M.J.J. Bonvin Strengths and weaknesses of data-driven docking in CAPRI *Proteins: Struc. Funct. & Bioinformatic*, **78**, 3242-3249 (2010).
9. Jorgensen, W. L. & Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **110**, 1657–1666 (1988).
10. Fernández-Recio, J., Totrov, M. & Abagyan, R. Identification of protein-protein interaction sites from docking energy landscapes. *J. Mol. Biol.* 335, 843–65 (2004).
11. A Vangone and A.M.J.J. Bonvin. Contacts-based prediction of binding affinity in protein-protein complexes. *eLife*, **4**, e07454 (2015).