

nature Structural biology

march 1996
vol. 3 no. 3

Challenging the predictors II

The business of prediction has always been regarded with a mixture of fascination and scepticism: perhaps as a consequence, its practitioners have received considerable reward as well as ridicule (or worse). The astrologers of the Roman Empire, for example, became so powerful and unpopular through their alleged divinations of the future that they were at one point expelled *en masse* from the whole of Italy. Although a similar fate seems unlikely for those trying to predict the structures of biological macromolecules—and the way in which they interact with one another—objective tests of the various prediction algorithms do much to guard against overenthusiastic claims of success (and therefore the inevitable subsequent brickbats) as well as illuminating their weaknesses and stimulating further research. In this vein, a huge prediction challenge organised by John Moult, at the Centre for Advanced Research in Biotechnology, Rockville, Maryland, and colleagues provides a chance for the predictors to ‘put up or shut up’.

In the face of the deluge of gene sequences being generated by the numerous gene-sequencing projects, the ability to predict accurately the structure of proteins has obvious utility. These are therefore potentially fertile times for the scientist who wishes to set up as a predictor. Of necessity, such scientists develop their tools on molecules where the two- and three-dimensional structures are already known. While this is only reasonable, the true test of a predictive tool is a blind one, but it can be difficult to organise in a way that is both thorough and effective.

The vast prediction ‘experiment’ organised by John Moult and colleagues^{1,2} in 1994 aimed to provide such a test. The message of the first experiment seems to be that structure prediction is still, as yet, an imprecise science. Exactly how imprecise depends on the specific category of prediction; in general, comparative modelling is more accurate than threading/fold identification, which in turn is more accurate than *ab initio* prediction. Of course it is the shortcomings—as much as the successes—of the various prediction algorithms revealed by the challenge that inform and teach. The organisers quote Shneior Lifson, one of the founders of the field: “You don’t learn anything until something goes wrong.”

Moult and colleagues second prediction challenge is now underway: prediction targets are being distributed as they become available, and are being withdrawn as they are solved (the final deadline is 30 September). As before, the success (or not) of the endeavour must depend on the predictors rising to

the challenge: those who decided not to participate in the first experiment may be encouraged to enter this second challenge, now that they have a better idea of what is in store. Certainly, the challenge seems like an ideal opportunity for synergy in the field. (The results of the experiment will be presented at a meeting in Asilomar, California 12–16 December; see <http://iris4.carb.nist.gov> or <http://www.mrc-cpe.cam.ac.uk/casp2>).

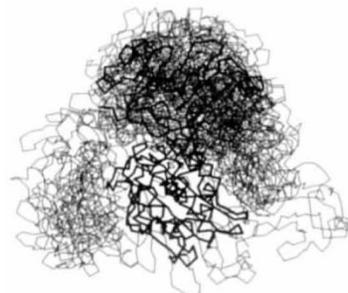
Docking algorithm challenge

As well as comparative modelling, fold recognition, and *ab initio* structure prediction, the organisers are also including a docking challenge as part of the experiment. This addition is an important one—molecular recognition is fundamental to the biological function of all proteins. The ability to predict the mode of interaction of a protein and its ligand(s) also has implications for protein engineering and structure-based drug design. But whereas our knowledge of proteins is ever more extensive—the number of structures being determined is increasing by the day—our comprehension of the underlying forces that govern their form and function is still far from complete.

The problem of understanding molecular recognition (like that of folding) can be tackled by drawing together knowledge of the basic structure of proteins with insights gleaned from the methods of predicting the response of a protein to different conditions (in this case, interaction with another molecule). The two avenues of research should, ideally, feed off each other: indeed, docking prediction may be considered a mature science when it is both reliable and able, in its turn, to provide information about the basic properties of proteins.

Results of a docking challenge organised by Michael James and colleagues now reported on page 233 of this issue of *Nature Structure Biology*³ suggest that docking prediction at last may be becoming a reliable art. Six groups accepted the challenge of predicting the structure of the complex formed between TEM-1 β -lactamase and the inhibitor protein BLIP before completion of the structure determination in the James laboratory. The structure of the individual proteins was known; the lack of supporting biochemical insight made the challenge all the more rigorous. Nonetheless, using a variety of different algorithms, all six groups were able to predict the general mode of interaction between the two proteins with 'best ranked' solutions having root-mean-square deviations for the BLIP protein in the complex ranging from 3.35 to 6.11 Å (ref. 3).

This success is perhaps all the more surprising when one considers that the predictors were unable to determine the local or global conformation changes on formation of the complex or the presence of a number of water molecules buried at the interface⁴. As James and colleagues point out, apparently all that matters is that the complementary surfaces of the two proteins are approximately matched together—this seems to be enough for the prediction algorithms to identify the general configuration of the native complex. Not an earth-shattering mechanistic insight, perhaps. More important may be the spur that these results will give to researchers in their quest for ever more accurate and reliable predictions. Let's hope that many more groups take up the challenges offered by Moult and colleagues.



The 40 predicted binding sites for BLIP (thin lines) are superimposed on the experimentally determined complex (thick lines) of BLIP and TEM-1 (see refs 3,4).

1. Moult, J., Pedersen, J.T., Judson, R. & Fidelis, K. *Proteins; Struct. Funct. Genet.* **23**, ii–iv (1995).
2. Shortle, D. *Nature Struct. Biol.* **2**, 91–93 (1995).
3. Strynaka, et al. *Nature Struct. Biol.* **3**, 233–239 (1996).
4. Strynaka, et al. *Nature Struct. Biol.* **3**, 290–297 (1996).