Correct structure prediction?

SIR — Nearly everyone agrees that one of the 'big' tasks in structural biology is to predict the folded conformation of proteins from the amino-acid sequence. Recently developed prediction methods based on an analysis of aligned identical protein sequences (see ref. 1 for review) have made several predictions that were shown to be remarkably accurate by subsequently determined crystal structures. Does this work represent a "spectacular achievement [that] will come to be recognized as a major breakthrough" (ref. 2)? Or does structure prediction still remain "more a matter for soothsayers than scientists" (ref. 3)?

One need not hold an opinion on this issue to realize that the best way for emerging prediction methods to achieve acceptance is through use. Structures must be predicted and the predictions published *before* crystallographic data are available. This way, knowledge of the structure cannot bias the prediction, the predictions (both correct and incorrect) are visible, and the method is placed 'at risk'. The only obstacle is one of coordination. A prediction published years in advance of a crystal structure is uninteresting. A prediction submitted even days after a crystal structure appears is useless.

We were fortunate that Musacchio *et al.* contacted us a few weeks ago to challenge us to predict the conformation of an SH3 domain (a small domain homologous to various signal-transduction proteins[4]) using prediction methods that we have developed in Zurich[5]. Our prediction has been submitted to the *Journal of Molecular Biology*. Musacchio *et al.* report the crystal structure of an SH3 domain on page 851 of this issue[6]. Thus, readers of *Nature* have another opportunity to compare a crystal structure with a prediction made *de novo*, and can decide for themselves whether there has been a breakthrough, or whether soothsaying prevails.

For those who cannot wait for the paper to appear in *J. molec. Biol.*, we summarize the prediction here. The protein is predicted to be built from β-strands with a single two-turn α-helix lying on one face. The predicted secondary structural elements correspond to the following positions in the sequence of the protein analysed crystallographically (which presumably starts with residues TGKEL): 7–13 (β1); 19–23 (β2); 26–33 (α1); 38–41 (β3); 48–50 (β4); and 53–55 (β5). This prediction contrasts sharply with that made by standard Chou–Fasman and Garnier–Osguthorpe–Robson methods, where the structure is predicted to be largely helical.

Predictions such as these are extremely important to the development of methods for predicting protein conformation. We welcome additional challenges to make predictions using our method, especially if (1) a structure will shortly be solved; (2) no structure is available for any obviously homologous protein; (3) sequences are sent to us by computer mail together with a few literature citations that provide an overview of the chemistry and biology of the protein family; and (4) this material is sent enough in advance to allow coordination of the publication of the prediction and publication of the structure.

Steven A. Benner
Mark A. Cohen
Dietlind Gerloff
Laboratory for Organic Chemistry,
ETH Zurich,
CH-8092 Switzerland
benner@czheth5a

1. Benner, S. A. *Curr. Opin. struct. Biol.* **2**, 402 (1992).
2. Lesk, A. M. & Boswell, D. A. *BioEssays* **14**, 407 (1992).
3. Hunt, T. & Purton, M. *Trends biochem. Sci.* **17**, 273 (1992).
4. Musacchio, A. *et al. FEBS Lett.* **307**, 55–61 (1992).
5. Benner, S. A. & Gerloff, D. *Adv. Enzyme Regul.* **31**, 121 (1991).
6. Musacchio, A. *et al. Nature* **359**, 851–855 (1992).

Speciation events

SIR — Coyne suggests[1] that the adaptive divergence of 'ordinary' genes explains reproductive isolation, in particular Haldane's rule that unisexual sterility or inviability is limited to the heterogametic sex. Others have proposed that 'novel' genetic phenomena make a large contribution to speciation. For example, genomic imprinting and other heritable epigenetic marks[2] or meiotic drive genes[3,4] might be involved. If this alternative view is correct then speciation genes, rather than being a representative sample of all genes, will be a distinctive set of genes with particular genetic properties (for example, imprinting, segregation distortion).

In his review, Coyne[1] did not discuss studies of hybrid sterility in the mouse except to note that they support the view that the X chromosome has a disproportionately large effect. In fact, three of the four known mouse hybrid sterility genes are autosomal, and in accordance with Haldane's rule all four genes have male-specific activity. The first known mouse hybrid sterility gene *Hybrid sterility-1* (*Hst-1*) maps to chromosome 17, as does *Hst-4*. *Hst-2* probably maps to chromosome 9 and only *Hst-3* is X-linked. High resolution mapping of *Hst-1*, *Hst-4* and X-linked sterility supports the involvement of 'novel' genetic events in causing speciation.

Hst-1 has been narrowed down to a region of chromosome 17 that contains at least four interesting genes: *Sod-2*, *tcl-w73*, *Tme* and *Igf2r* (ref. 5). *Sod-2* is a nonimprinted gene which encodes mitochondrial superoxide-dismutase; *tcl-w73* is a recessive lethal factor of a *t*-haplotype. Both the insulin-like growth factor II receptor gene (*Igf2r*) and the T-associated maternal effect gene (*Tme*) are imprinted such that only the maternal copy is expressed in embryos[6]. It had been postulated that *Igf2r* and *Tme* were identical[7]. But when it was found that in interspecies hybrids *Igf2r* and *Tme* assumed different patterns of imprinting from each other[6] it was concluded that they are probably two separate but tightly linked loci. This interference with imprinting in the hybrid strongly suggests that these loci are involved in speciation.

The *Hst-4* locus causes male sterility in crosses between *Mus domesticus* and *M. spretus*. It, too, is located on chromosome 17 and is not separable from the *t* complex distorter locus *Tcd-2*, which is involved in male transmission ratio distortion[8]. This finding is direct evidence that meiotic drive genes are associated with unisexual hybrid sterility.

X-linked male sterility affects crosses between *M. domesticus* and *M. spretus*. Hybrids show a high frequency of X–Y dissociation in the first meiotic metaphase[9]. In backcrosses, both X–Y dissociation and sterility cosegregate with the *Amel* locus, close to the X–Y pairing region. Genetic divergence of the X–Y pairing region seems to cause X-linked hybrid sterility[9]. Again, this implicates meiotic drive, as lack of pairing between the sex chromosomes is associated with segregation distortion[10]. In summary, studies of unisexual sterility in the mouse suggest that the first step towards speciation involves more distinctive genetic events than those suggested by Coyne.

Laurence D. Hurst
Department of Zoology,
University of Oxford,
South Parks Road,
Oxford OX1 3PS, UK
Andrew Pomiankowski
Department of Genetics and Biometry,
University College London,
4 Stephenson Way,
London NW1 2HE, UK

1. Coyne, J. A. *Nature* **355**, 511–515 (1991).
2. Jablonka, E. & Lamb, M. J. *Nature* **356**, 752 (1992).
3. Hurst, L. D. & Pomiankowski, A. *Genetics* **128**, 841–858 (1991).
4. Frank, S. *Evolution* **45**, 262–267 (1991).
5. Forejt, J. *et al. Mammal. Genome* **1**, 84–91 (1991).
6. Forejt, J. & Gregorová, S. *Cell* **70**, 443–450 (1992).
7. Barlow, D. P. *et al. Nature* **349**, 84–87 (1991).
8. Pilder, S. H., Hammer, M. F. & Silver, L. M. *Genetics* **129**, 237–246 (1991).
9. Matsuda, Y., Hirobe, T. & Chapman, V. M. *Proc. natn. Acad. Sci. U.S.A.* **88**, 4850–4854 (1991).
10. McKee, B. D. *Am. Nat.* **137**, 332–339 (1991).