PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

1. Interferon: A Tertiary Structure Predicted from Amino Acid Sequences

M. J. E. Sternberg and F. E. Cohen

Phil. Trans. R. Soc. Lond. B 1982 **299**, 125-127
doi: 10.1098/rstb.1982.0113

Phil. Trans. R. Soc. Lond. B **299**, 125–144 (1982) [125]
Printed in Great Britain

Synopses from the poster exhibition organized by I. M. Kerr

1. Interferon: a tertiary structure predicted from amino acid sequences

By M. J. E. Sternberg and F. E. Cohen†
Laboratory of Molecular Biophysics, Department of Zoology, South Parks Road, Oxford OX1 3PS, U.K.

Introduction

Functional studies on interferon would be helped by a three-dimensional structure for the molecule. However, it may be several years before the structure of the protein is determined by X-ray crystallography. We have therefore used available methods for predicting the secondary – and the tertiary – structure of a protein from its amino acid sequence to propose a tertiary model involving the packing of four α-helices. Details of this work have been published elsewhere (Sternberg & Cohen 1982).

Amino acid sequences

Four homologous amino acid sequences for interferon were considered: from human fibroblast (Derynck *et al.* 1980), from human leucocyte I and II (Taniguchi *et al.* 1980; Streuli *et al.* 1980), and peptide fragments from human lymphoblastoid interferon (Zoon *et al.* 1980). When the sequences are aligned the strong homology between them suggests that they will have similar tertiary structures. For convenience the numbering of the fibroblast sequence is used.

Prediction of secondary structure

Three methods of predicting the secondary structure from amino acid sequence were used: those of Lim (1974), Chou & Fasman (1978), and Robson and coworkers (Garnier *et al.* 1978). There was considerable variation in the results of secondary structure prediction both between different methods on the same sequence and between the different sequences. However, one likely interpretation is that there are four α-helical segments (denoted A, B, C and D) that will be important in the tertiary fold. In the fibroblast sequence these are A (Phe 15 to Arg 27), B (Leu 57 to Gln 72), C (Glu 81 to Leu 106), and D (Trp 143 to Phe 154). In addition, short β-strands could occur around Leu 32 to Met 36 and Tyr 125 to Leu 133.

Prediction of tertiary structure

If the conclusion from the secondary structure prediction that there are four α-helices that are important in the tertiary fold is correct, then a probable tertiary structure is an α-helical bundle (Weber & Salemme 1980), which is a motif found in several globular proteins. In this

† Present address: Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143, U.S.A.

motif (see figure 1) sequential α-helices pack antiparallel and the topology of four α-helices is a right-handed bundle. This motif was shown to be stereochemically favourable by the use of an α-helix docking algorithm (Richmond & Richards 1978; Cohen *et al.* 1979; Cohen & Sternberg 1980). The basic notions of the algorithm are the prediction of hydrophobic docking sites on the surfaces of the α-helices and the generation of a fold in which the α-helices pack with a

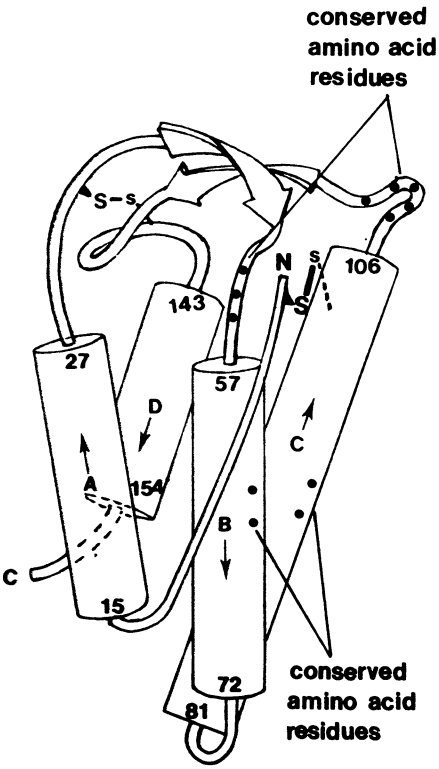

FIGURE 1. The proposed interferon structure. The figure (drawn by Dr W. R. Taylor) shows schematically the four α-helices (cylinders) and the two possible β-strands (arrows). The helix termini are numbered according to the fibroblast sequence and the disulphide bridges in leucocyte-I interferon are indicated (–S–S–). (From Sternberg & Cohen (1982).)

standard geometry. Furthermore it was shown that in the proposed fold for the α-helices no polar side chains were totally buried. In the proposed fold the possible β-strands could form a small β-sheet that lies at one end of the bundle.

DISULPHIDE BRIDGES

The disulphide bridges in leucocyte-I interferon have been assigned (Wetzel 1980). With the numbering of fibroblast interferon these linkages are in leucocyte interferon Cys 31 – Cys 141 and Cys 4 – Cys 101. The distance between Cys 31 and Cys 141 in the proposed model is consistent with this linkage (see figure 1). For the Cys 4 – Cys 101 bridge to be formed, the amino-terminal segment before the start of α-helix A would have to be roughly antiparallel to the direction of α-helix A.

Implications

Subjective judgements had to be made at several stages during the prediction but we consider that the proposed structure for interferon has better than a 50% likelihood of being correct.

From the predicted structure possible sites for function can be proposed based on residues that are spatially close (figure 1) and are conserved in the different sequences. The two regions with the highest conservation of nucleotides and high sequence homology between fibroblast and leucocyte-I interferons (see fig. 4 of Taniguchi *et al.* 1980) are residues 37 to 48 and 105 to 119. These regions would be spatially close at the same edge of the bundle as the Cys 31–Cys 141 disulphide bridge and could form a functionally important site. Similarly Glu 61 and Gln 64 on helix B and Gln 94 and Asn 96 on helix C are conserved in all sequences and, being exposed to solvent and spatially close, might have a functional role.

It is intended that the suggestions of a stereochemically reasonable model will help detailed studies about the structure and function of interferon.

We thank Sir David Phillips, Sec.R.S., for encouragement and advice. M.J.E.S. holds a Stothert Fellowship from the Royal Society and F.E.C. was a Rhodes Scholar.

References

Chou, P. Y. & Fasman, G. D. 1978 Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* **47**, 45–148.

Cohen, F. E., Richmond, T. J. & Richards, F. M. 1979 Protein folding: evaluation of some simple rules for the assembly of helices into tertiary structures with myoglobin as an example. *J. molec. Biol.* **132**, 275–228

Cohen, F. E. & Sternberg, M. J. E. 1980 On the use of chemically derived distance constraints in the prediction of protein structure with myoglobin as an example. *J. molec. Biol.* **137**, 9–22.

Derynck, R., Content, J., De Clercq, E., Volckaert, G., Tavernier, J., Devos, R. & Fiers, W. 1980 Isolation and structure of a human fibroblast interferon gene. *Nature, Lond.* **285**, 542–547.

Garnier, J., Osguthorpe, D. J. & Robson, B. 1978 Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. molec. Biol.* **120**, 97–120.

Lim, V. I. 1974 Algorithms for prediction of α-helical and β-structural regions in globular proteins. *J. molec. Biol.* **88**, 873-894.

Richmond, T. J. & Richards, F. M. 1978 Packing of α-helices: geometrical constraints and contact areas. *J. molec. Biol.* **119**, 537–555.

Sternberg, M. J. E. & Cohen, F. E. 1982 The prediction of the secondary and tertiary structures of interferon from four homologous amino acid sequences. *Int. J. biol. Macromol.* **4**, 137–144.

Streuli, M., Nagata, S. & Weissmann, C. 1980 At least three human type α interferons: structure of α2. *Science, Wash.* **209**, 1343–1347.

Taniguchi, T., Mantei, N., Schwarzstein, M., Nagata, S., Maramatsu, M. & Weissmann, C. 1980 Human leukocyte and fibroblast interferons are structurally related. *Nature, Lond.* **285**, 547–549.

Weber, P. C. & Salemme, F. R. 1980 Structural and functional diversity in 4-α-helical proteins. *Nature, Lond.* **287**, 82–84.

Wetzel, R. 1980 Assignment of the disulphide bonds of leukocyte interferon. *Nature Lond.*, **289**, 606–607.

Zoon, K. C., Smith, M. E., Bridgen, P. J., Anfinsen, C. B., Hunkapiller, M. W. & Hood, L. E. 1980 Human fibroblast interferon: amino acid analysis and amino terminal amino acid sequence. *Science, Wash.* **207**, 525–527.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES