

REFERENCES AND NOTES

1. M. Matsuoka *et al.*, *Science* **254**, 81 (1991).
2. F. Nagawa *et al.*, unpublished observation.
3. J. J. M. Chun, D. G. Schatz, M. A. Oettinger, R. Jaenisch, D. Baltimore, *Cell* **64**, 189 (1991).
4. M. Matsuoka, K. Yoshida, T. Maeda, S. Usuda, H. Sakano, *Cell* **62**, 135 (1991).
5. We thank J. A. Winer and D. Larue for their help in cytochemical analyses of transgenic mice.

1 June 1992; accepted 18 June 1992

Unraveling the Structure of IL-2

Interleukin-2 (IL-2) is the prototype member of an emerging class of pleiotropic cytokines that have extraordinary immunologic and therapeutic promise (1). An influential guide for much of the research (2) on IL-2 structure and function has been a paper by Brandhuber *et al.* (3). However, their x-ray structure of human IL-2 at 3 Å resolution appears to be at odds with experimental results garnered from deletion, mutagenic, and binding studies of IL-2. An alternative protein folding for IL-2, presented here, is closely similar to that of the recently elucidated structures of granulocyte-macrophage colony-stimulating factor (GM-CSF) (4) and interleukin-4 (IL-4) (5).

Schrader *et al.* have suggested that IL-2 and GM-CSF are structurally homologous proteins that share a low but significant degree of sequence identity (6). However, the x-ray structures appear topologically unrelated. Although both IL-2 and GM-CSF show a compact core bundle of four antiparallel α helices, the IL-2 x-ray fold features short hairpin loops between helices (3) (Fig. 1A) and is classified as a left-handed, type 0 bundle (7). The GM-CSF structure has two long crossover connections between the first and second and third and fourth helices (4) (Fig. 1B). Aside from IL-4 (5), the latter folding topology [a left-handed, type 2 bundle (7)] is observed only in two other cytokine structures, growth hormone and interferon- β (8).

I have compared available GM-CSF and IL-4 sequences and structures with IL-2 (Fig. 2) to extend the findings of Schrader *et al.* (6). Except for the COOH-terminal helices, my findings show a significant lack of spatial correspondence between sequential helices in the IL-2 x-ray fold proposed by Brandhuber *et al.* and those of GM-CSF and IL-4. I set aside the x-ray helix assignment for IL-2 (3) and used six IL-2 sequences to derive a consensus secondary structure of five α helices and two short β strands (9) (Fig. 2). Four of the predicted amphipathic α helices and both β strands in the IL-2 model are naturally aligned with similar elements in GM-CSF and IL-4 (Fig. 2). [The IL-2 helix predictions for the most part match the segments located by Cohen *et al.* (10) in their early model of human IL-2 based on a right-handed, type 0 bundle.] The largely helical nature of IL-2 is

supported by circular dichroism spectra that also distinguish the presence of some β structure (10), in further analogy with GM-CSF and IL-4 (11). The structure-based alignment of the three cytokines also matches the critical Cys⁵⁸-Cys¹⁰⁵ disulfide bridge in human IL-2 (3), with analogous Cys⁵⁴-Cys⁹⁶ and Cys⁴⁶-Cys⁹⁹ links in the human GM-CSF and IL-4 structures (4, 5), respectively (Fig. 2A). Similar helices and loops are also encoded by like exons in IL-2, GM-CSF, and IL-4 genes (Fig. 2) (1).

A tertiary model of the human IL-2 structure, similar to GM-CSF/IL-4, that incorporates the predicted α and β segments can be parsimoniously accommodated by the existing x-ray scaffold if a few changes are made (compare Fig. 1, A and B): (i) the four core helices of the x-ray bundle—2-2', 3, 4, and 6—are replaced by model helices B, C, A, and D while retaining only the 2'-3 (now B-C) loop; (ii) the long A-B crossover loop of the model includes a short helical segment (spatially equivalent to x-ray helix 5) and a β strand which now forms a link to the "downward" chain (Fig. 1B) of a meandering loop that connects extracore x-ray helix 1 to 2. There is no counterpart to x-ray helix 1; (iii) the long C-D crossover loop of the model traces a chain across the space occupied by x-ray helix 1, through the "upward" part of the loose x-ray 1-2 loop to form a second short β strand that leads to COOH-terminal model helix D; (iv) the model requires a

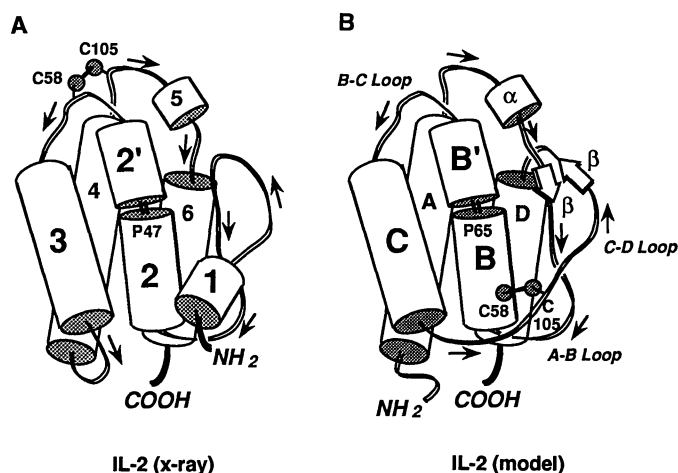
relocation of the Cys⁵⁸-Cys¹⁰⁵ disulfide from the "top" of the helix bundle (Fig. 1A) to form a new link between model helix B and the C-D loop (Fig. 1B).

By implication, the major helical features of the IL-2 x-ray fold appear to be correct; however, approximately 87% of the protein chain may have to be retraced and several loops rebuilt in order to link the core helices in the new topology. Difficulty in interpreting the electron density map at medium resolution is the likely cause of these errors (12). The backbone course of the proposed long loops was perhaps ambiguous because density was poor, a salt bridge or bulky side chain could have mimicked the disulfide link, and a correct sequence tracing through the core helices might have been impeded by unresolved side chains. Indeed, of the several "tether" points located by the x-ray analysis to guide the chain tracing (3), only the heavy atom-aided positioning of Cys¹²⁵ in COOH-terminal helix 6 (D in Fig. 1B) remains fully consistent with the new IL-2 topology. Similar mistracings of the structures of *ras* p21 and HIV-1 protease (13) have occurred recently.

Among structure modification techniques, fine deletion mapping analyses of GM-CSF have proved effective in outlining structurally important regions of the chain before the determination of the GM-CSF x-ray fold (Fig. 2) (4, 14). An analogous, exhaustive structure mapping of mouse IL-2 (15) should prove equally diagnostic of the helix boundaries in IL-2. These results are more supportive of the model IL-2 secondary structure (Fig. 2): the x-ray determined helices appear too short (helix 1), lie in apparently nonessential (loop-like) regions (helices 2-2' and 5), or are out of register with the critical zones (helices 3 and 4).

Another line of inquiry that is distinctly at odds with the x-ray findings concerns the mapping of receptor-binding epitopes in the folded IL-2 molecule. The cellular IL-2

Fig. 1. Comparison of x-ray-derived and model folds of IL-2. (A) Schematic drawing of the IL-2 x-ray helix bundle (3). Cylindrical helices are marked 1 to 6. Loops are drawn as loose ribbons; Pro⁴⁷ (P47) is marked. The disulfide is noted by linked spheres. In the GM-CSF/IL-4-like IL-2 model (B), the chain through the core helices is retraced and reconnected, the disulfide bridge is relocated, the existence of a small β sheet is proposed (4, 5); Pro⁶⁵ (P65) is marked. Only helix D remains fully equivalent in sequence to x-ray helix 6.



receptor complex likely consists of at least three receptor subunits, two of which—the low-affinity p55 α chain and intermediate-affinity p75 β chain—have been completely elucidated (2). The p55 binding site has been localized to the x-ray helices 2 and 5, while p75 likely interacts with the con-

served Asp²⁰ after x-ray helix 1 (16). A third, γ -chain receptor binding epitope centers instead on Gln¹²⁶ in the COOH-terminal helix (17). Unlike the x-ray structure, the model places the two p55 epitopes in close proximity in paired A-B and C-D loops at the “top” end of the bundle (Fig.

1B). Also, mutagenic perturbation of either Asp²⁰ or Gln¹²⁶ produces a similar functional effect, which suggests that they are exposed on a common face for recognition by closely interacting β - and γ -chain IL-2 receptors (17). In the model IL-2 structure these two critical amino acids are clustered

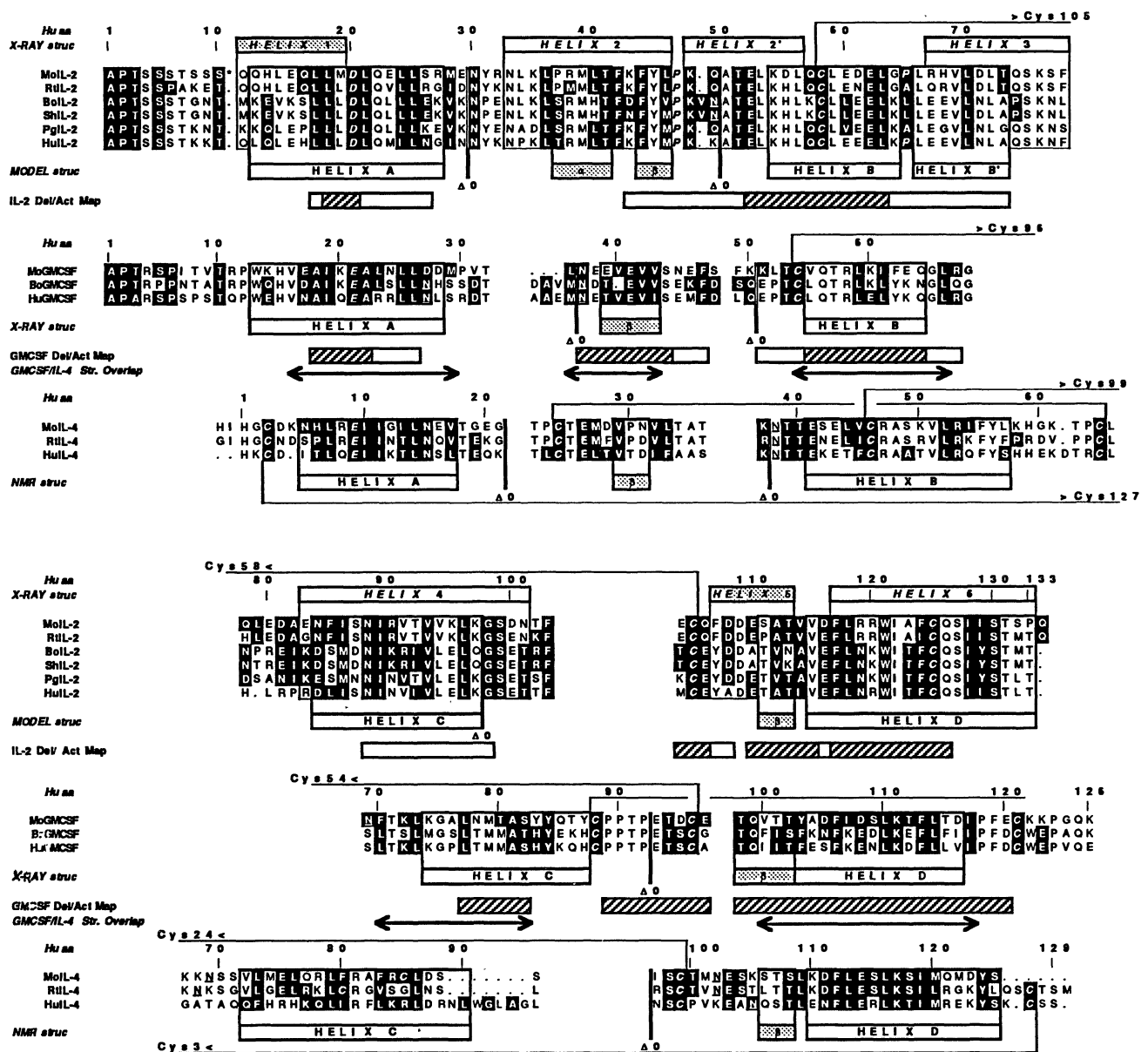


Fig. 2. Structure-based alignment of IL-2, GM-CSF, and IL-4 sequences and summary of deletion studies. Available IL-2, GM-CSF, and IL-4 sequences are from mouse (Mo), rat (Rt), bovine (Bo), sheep (Sh), pig (Pg), and human (Hu) species (1, 23). An asterisk indicates the insertion of sequence TAEA(Q)₁₀ in mouse IL-2. Exon boundaries in the cognate genes feature a vertical marker; $\Delta 0$ indicates phase 0 introns. Positional similarity between the three cytokines is marked by reverse lettering [criteria of (18)]; functionally important NH₂- and COOH-terminal helices share a greater identity (6, 18). Regions of structural overlap between GM-CSF and IL-4 are noted by arrows; 112 C α positions in the two-helix bundle cores can be superimposed with a root-mean-square (rms) deviation of 1.48 Å; 16 additional C α atoms in loops increase the rms deviation to ~1.58 Å. A like degree of similarity is predicted between IL-2 and either GM-CSF or IL-4. Cysteine residues that form disulfide bridges (shown) are italicized in reverse-letter. Possible N-glycosylated aspar-

agines are underlined. Model helices in IL-2 (9) and the determined helices in GM-CSF (4) and IL-4 (5) are boxed and bottom-labeled A to D, and x-ray helices for IL-2 are top-labeled 1 to 6. Extra-core helices and β strands (and x-ray helices 1 and 5) are noted by shaded boxes. Asp²⁰ in IL-2, and analogous Glu²¹ and Glu⁹ residues in GM-CSF and IL-4, respectively [important in receptor binding (16–18)] are distinguished by italics. Deletion maps for mouse IL-2 (15) indicate structurally critical regions of the chain (deletions in hatched and white boxes are the $\leq 0.05\%$ and $\leq 10\%$ of wild-type activity, respectively). For human and mouse GM-CSF (14), hatched and white boxes mark sensitive regions of $\leq 0.01\%$ and 1.0% of wild-type activity, respectively. Blank areas between critical regions can be deleted with only moderate loss of activity (14, 15). Abbreviations for the amino acid residues are as follows. A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

on adjacent helices A and D (Fig. 1B). Homologous residues in a similar structural setting are proposed for GM-CSF and other cytokines that also bind to receptors of the hematopoietic supergroup (18).

There remain other computational approaches to gauging the correctness of the structure with existing coordinates (12, 19). However, there now appears to be sufficient evidence to merit a reappraisal of the IL-2 x-ray fold (3). As in other cases (20), the utility of the new model lies in simulating further experiments and a full refinement of the IL-2 structure.

J. Fernando Bazan

Department of Biochemistry and Biophysics,
University of California,
San Francisco, CA 94143-0448

REFERENCES AND NOTES

1. K. I. Arai *et al.*, *Annu. Rev. Biochem.* **59**, 783 (1990).
2. K. A. Smith, *Science* **240**, 1169 (1988); T. A. Waldmann, *J. Biol. Chem.* **266**, 2681 (1991).
3. B. J. Brandhuber, T. Boone, W. C. Kenney, D. B. McKay, *Science* **238**, 1707 (1987); coordinates for this structure have not been deposited in the Brookhaven Protein Databank.
4. K. Diederichs, T. Boone, P. A. Karplus, *ibid.* **254**, 1779 (1991).
5. C. Redfield *et al.*, *Biochemistry* **30**, 11029 (1991); L. J. Smith *et al.*, *J. Mol. Biol.* **224**, 899 (1992).
6. J. W. Schrader, H. J. Ziltener, K. B. Leslie, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 2458 (1986).
7. S. R. Presnell and F. E. Cohen, *ibid.* **86**, 6592 (1989).
8. S. S. Abdel-Meguid *et al.*, *ibid.* **84**, 6434 (1987); T. Senda *et al.*, *Proc. Jpn. Acad.* **66B**, 77 (1990).
9. J. F. Gibrat *et al.*, *J. Mol. Biol.* **198**, 425 (1987); T. Niermann and K. Kirschner, *Methods Enzymol.* **202**, 45 (1991); F. E. Cohen *et al.*, *Biochemistry* **25**, 266 (1986); J. S. Richardson and D. C. Richardson, *Science* **240**, 1648 (1988).
10. F. E. Cohen *et al.*, *Science* **234**, 349 (1986).
11. P. Wingfield *et al.*, *Eur. J. Biochem.* **173**, 65 (1988); W. T. Windsor *et al.*, *Biochemistry* **30**, 1259 (1991).
12. J. S. Richardson and D. C. Richardson, *Methods Enzymol.* **115**, 189 (1985); C.-I. Branden and T. A. Jones, *Nature* **343**, 687 (1990); J. Janin, *Biochimie* **72**, 705 (1990).
13. A. M. de Vos *et al.*, *Science* **239**, 888 (1988); E. F. Pai *et al.*, *Nature* **341**, 209 (1989); M. A. Navia *et al.*, *ibid.* **337**, 615 (1989); A. Wlodawer *et al.*, *Science* **245**, 616 (1989).
14. A. B. Shanafelt and R. A. Kastelein, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 4872 (1989); A. B. Shanafelt *et al.*, *J. Biol. Chem.* **266**, 13804 (1991).
15. S. M. Zurawski and G. Zurawski, *EMBO J.* **7**, 1061 (1988).
16. L. Collins *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 7709 (1988); U. Weigel, M. Meyer, W. Sebald, *Eur. J. Biochem.* **180**, 295 (1989); S. M. Zurawski and G. Zurawski, *EMBO J.* **8**, 2583 (1989); K. Saueve *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 4636 (1991).
17. S. M. Zurawski, J. L. Imler, G. Zurawski, *EMBO J.* **9**, 3899 (1990).
18. A. B. Shanafelt *et al.*, *ibid.* **10**, 4105 (1991); J. F. Bazan, *Neuron* **7**, 197 (1991).
19. J. Novotny *et al.*, *Proteins* **4**, 19 (1988); L. M. Gregoret and F. E. Cohen, *J. Mol. Biol.* **211**, 959 (1990); T. A. Jones *et al.*, *Acta Crystallogr.* **A47**, 110 (1991); L. Holm and C. Sander, *J. Mol. Biol.* **218**, 183 (1991); R. Luthy *et al.*, *Nature* **356**, 83 (1992).
20. J. S. Richardson, *Biochem. Biophys. Res. Commun.* **90**, 285 (1979); I. T. Weber *et al.*, *Science* **243**, 928 (1989).
21. A. J. McKnight *et al.*, *Immunogenetics* **30**, 145

(1989); J. C. Goodall *et al.*, *Biochim. Biophys. Acta* **1089**, 257 (1991); C. R. Maliszewski *et al.*, *Mol. Immunol.* **25**, 843 (1988); A. J. McKnight *et al.*, *Eur. J. Immunol.* **21**, 1187 (1991).

27 January 1992, revised 6 May 1992; accepted 7 May 1992

Response: In answer to the comment by Bazan, we have completed a refinement of a molecular model of human recombinant interleukin-2 (IL-2) in which Cys¹²⁵ is replaced with alanine ([Ala¹²⁵]IL-2) (1). The x-ray crystallographic data have a resolution of 2.5 Å and were collected by diffractometer on triclinic crystals, as previously reported (2). We used the program FRA-GLE (DuPont version 6.6dup) (3), which allows one to replace manually built segments of a model with molecular fragments of refined protein structures, to build models of the two independent molecules in the unit cell. We used the program XPLOR (4) to refine molecular models with simulated annealing and restrained least-squares positional refinement. Phase combination between phases computed from a molecular model and "solvent-flattened" phases (that is, model-independent phases computed from a solvent-flattened multiple isomor-

phous replacement map) was done as previously described (5).

The following evidence supports the correctness of our current model. (i) The refinement statistics are consistent with a correct structure. (ii) The folding topology is consistent with that of GM-CSF (6) and IL-4 (7). (iii) Residues that are widely separated in the primary sequence of mouse IL-2 have been identified by exhaustive substitution mutagenesis with binding to either the p55 or the p70 receptor subunit. In our current model, these residues map to two contiguous patches on the surface of the human [Ala¹²⁵]IL-2 structure (8).

Historically, methods such as chemical modification in crystals have been used to identify specific amino acids or amino acid types, which serve as "brass tacks" to pin down the sequence of a protein model into a relatively noisy map (9). In the case of IL-2, Cys¹²⁵ was identified correctly (2) as a site at which mercurials bind in crystals of IL-2, but do not bind in crystals of [Ala¹²⁵]IL-2. Also, crystals were iodinated in an attempt to identify tyrosines and place the remainder of the sequence in register in the maps. The major site of iodination in the crystal was assumed incorrectly to correspond to the

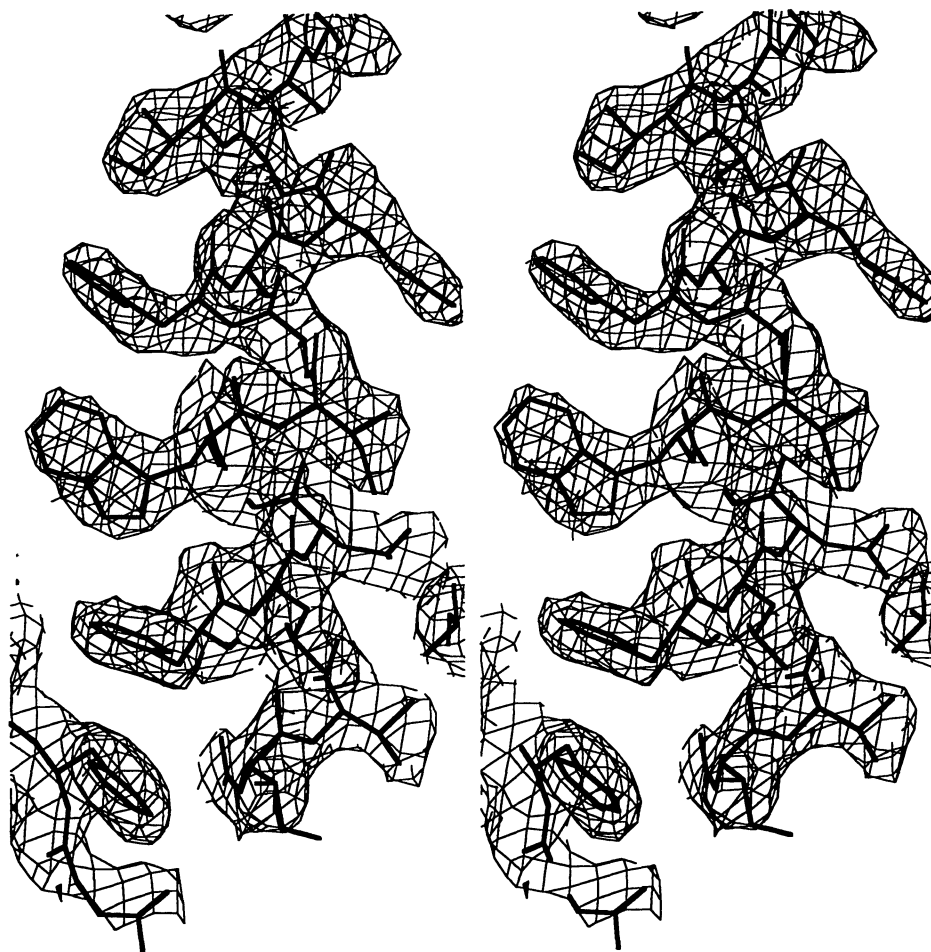


Fig. 1. Stereogram of a portion of the final $(2F_o - F_c)\exp(i\alpha)$ map; phases computed as described in (1).

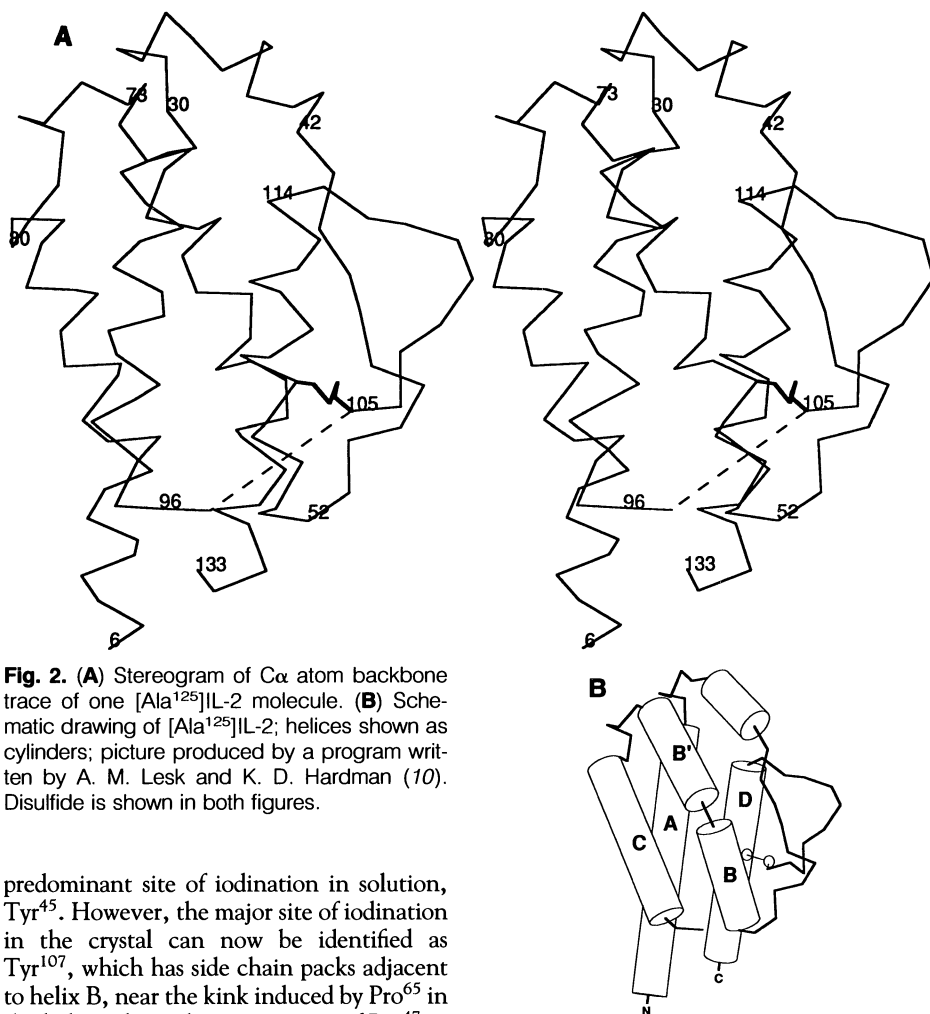


Fig. 2. (A) Stereogram of C α atom backbone trace of one [Ala¹²⁵]IL-2 molecule. (B) Schematic drawing of [Ala¹²⁵]IL-2; helices shown as cylinders; picture produced by a program written by A. M. Lesk and K. D. Hardman (10). Disulfide is shown in both figures.

predominant site of iodination in solution, Tyr⁴⁵. However, the major site of iodination in the crystal can now be identified as Tyr¹⁰⁷, which has side chain packs adjacent to helix B, near the kink induced by Pro⁶⁵ in the helix. The earlier assignment of Pro⁴⁷ to the proline in helix B, and Tyr⁴⁵ to the residue whose side chain was adjacent to the iodine site, led to our misinterpretation of the connectivity of the IL-2 molecule.

David B. McKay

Department of Cell Biology,
Stanford University Medical Center,
Stanford, CA 94305-5400

REFERENCES AND NOTES

1. D. B. McKay *et al.*, in preparation; Refinement was initiated with a minimal molecular model. In the first cycle, the model included the α -helices of the four-helix bundle for each molecule, as well as an additional short helical segment for one of the molecules, for a total of 189 amino acid residues

out of a possible 266; no connections were built between helices. The human IL-2 sequence was incorporated for the 15-residue COOH-terminal helix of each molecule, in which the position of Ala¹²⁵/Cys¹²⁵ had been unambiguously identified in (2). For the remainder of the model, side chains that appeared to fill the density of the solvent-flattened map were used without reference to the IL-2 sequence. After one cycle of refinement and phase combination, it became clear in the resulting phase-combined $2F_o - F_c$ maps that the connection leading into the COOH-terminal helix had been incorrectly traced in (2). Correcting this error allowed us to build a model of an additional 26 residues. After the second cycle of refinement and phase combination, the placement of the IL-2 sequence in the model became apparent. The first two cycles of refinement included 5982 reflections between 8.0 and 2.7 Å resolution. For the

map calculation, solvent-flattened phases were used between 40.0 and 6.0 Å, combined phases between 6.0 and 3.0 Å, and model phases between 3.0 and 2.7 Å. Once the placement of the sequence was clear, the resolution was extended to 2.5 Å and included 7275 reflections. Solvent-flattened phases were used between 40.0 and 6.0 Å, combined phases between 6.0 and 3.5 Å, and model phases between 3.5 and 2.5 Å (Fig. 1). Subsequent cycles of refinement have yielded models that include residues 6 to 97 and 105 to 133 for each of the two molecules in the unit cell. The final *R* factor was 0.202, with root-mean-square (rms) deviations from ideal geometry of 0.21 Å for bonds and 4.0° for angles. Essentially all of the peptide dihedral angles lie within allowed regions on a Ramachandran diagram. The first five amino acids could not be resolved. Also, although there is density in the maps that generally appears to connect residue 97 to residue 105 in both molecules, we could not trace a peptide with an unambiguous conformation for this connection at any stage of the refinement; hence residues 98 to 104 have been omitted in the final model. No solvent molecules have been incorporated. Throughout the model-building and refinement, the two molecules in the unit cell were built independently. After refinement, when the atoms of the polypeptide backbone of one molecule were superimposed on the second, the rms distance between corresponding atomic positions was 0.51 Å. Coordinates have been deposited in the Brookhaven Protein Data Bank.

2. B. J. Brandhuber, T. Boone, W. C. Kenney, D. B. McKay, *Science* **238**, 1707 (1987); —, *J. Biol. Chem.* **262**, 12306 (1987).
3. B. C. Finzel *et al.*, in *Crystallographic and Modeling Methods in Molecular Design*, S. Ealick and C. Bugg, Eds. (Springer-Verlag, New York, 1990), p. 175.
4. A. T. Brunger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1985).
5. V. S. Allured, R. J. Collier, S. F. Carroll, D. B. McKay, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 1320 (1986); S. J. Remington, B. Wiegand, R. Huber, *J. Mol. Biol.* **158**, 111 (1982).
6. K. Diederichs, T. Boone, P. A. Karplus, *Science* **254**, 1779 (1991).
7. C. Redfield *et al.*, *Biochemistry* **30**, 11029 (1991).
8. S. M. Zurawski, E. L. Doyle, F. A. Vega, B. Huyghe, G. Zurawski, in preparation.
9. P. B. Sigler, *Biochemistry* **18**, 3609 (1970).
10. A. M. Lesk and K. D. Hardman, *Science* **216**, 539 (1982); —, in *Methods in Enzymology*, H. W. Wyckoff, C. H. W. Hirs, and S. N. Timasheff, Eds. (Academic Press, Orlando, FL, 1985), pp. 381–390.
11. I thank K. Flaherty for assistance in the computations during refinement and G. Zurawski *et al.* (8) for permission to reference their data before publication. IL-2 and [Ala¹²⁵]IL-2 protein on which crystallographic data were collected was originally provided by Amgen Biologicals (Thousand Oaks, CA), as acknowledged previously (2).

28 February 1992; revised 26 May 1992; accepted 15 June 1992