Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The 2.0 Å structure of bovine interferon-γ; assessment of the structural differences between species

Michael Randal[a]† and Anthony A. Kossiakoff[b]*‡

[a]Graduate Group in Biophysics, University of California, San Francisco, California 94000, USA, and [b]Genentech Inc., 1 DNA Way, South San Francisco, California 94080, USA

† Present Address: Sunesis Pharmaceuticals Inc., 3696 Haven Avenue Suite C, Redwood City, California 94062, USA.
‡ Present Address: University of Chicago, Department of Biochemistry and Molecular Biology, 920 East 58th Street, Cummings Life Sciences Center, Chicago, Illinois 60637, USA.

Correspondence e-mail: koss@cummings.uchicago.edu

Received 29 June 1999
Accepted 4 November 1999

PDB Reference: bovine interferon-γ, 1d9c.

The structure of bovine interferon-γ (IFN-γ) was determined by multiple isomorphous replacement at 2.0 Å resolution. Bovine IFN-γ crystallizes in two related crystal forms. Crystal form 1 diffracts to 2.9 Å resolution and is reproducible and stable to derivatization. Crystal form 2 diffracts to 2.0 Å resolution, but shows significant non-isomorphism from crystal to crystal. The previously determined structures of several different species of INF-γ were either at too low a resolution [human, 1hig; Ealick *et al.* (1991), *Science*, **252**, 698–702] or were too inaccurate [bovine, 1rfb; Samudzi & Rubin (1993), *Acta Cryst.* D**49**(6), 505–512; rabbit, 2rig; Samudzi *et al.* (1991), *J. Biol. Chem.* **266**(32), 21791–21797] for the structure to be solved by molecular replacement. The structure was solved in crystal form 1 using two derivatives produced by chemically modifying two free cysteine residues that were introduced by site-directed mutagenesis (Ser30Cys, Asn59Cys). After model building and refinement, the final R value was 21.8% ($R_{\rm free}$ = 30.9%) for all data in the resolution range 8.0–2.9 Å. The crystal form 1 structure was then used as a molecular-replacement model for crystal form 2 data collected from a flash-cooled crystal. Subsequent model building and refinement, using all data in the resolution range 15.0–2.0 Å, gave an R value of 19.7% and an $R_{\rm free}$ of 27.5%. Pairwise comparison of C^{α} positions of bovine IFN-γ (BOV) and the previously determined 1rfb and 2rig structures indicated some significant differences in the models (r.m.s.d. values for BOV to 1rfb, 4.3 Å; BOV to 2rig, 4.0 Å). An assessment of the quality of the structures was made using the 3D–1D algorithm [Eisenberg *et al.* (1992), *Faraday Discuss.* **93**, 25–34]. The resulting statistical scoring indicated that BOV was consistent with expected criteria for a 2.0 Å structure, whereas both 1rfb and 2rig fell below acceptable criteria.

1. Introduction

Interferons (IFNs) are a family of cytokines that are related by their ability to protect cells against viral infection. Based on several criteria, including biological activity, cellular origin and tertiary structure, the IFNs have been further divided into two subtypes: type I and type II IFNs. Type I IFNs, including IFN-α and IFN-β, are expressed as a direct result of viral infection (Stewart, 1979; Pestka *et al.*, 1987), while expression of IFN-γ, a type II IFN, is modulated by immune and inflammatory stimuli rather than by viral infection (Farrar & Schreiber, 1993).

Native IFN-γ exists as a homodimer. The monomeric chains of most species are expressed as a 166 amino-acid polypeptide, including a 23 amino-acid signal sequence. Enzymatic removal of the signal peptide produces the mature 143-residue protein,

though proteolysis at the C-terminus can result in further truncation (Rinderknecht *et al.*, 1984). In addition to this postranslational proteolysis, the amino terminus is converted to pyroglutamate (Rinderknecht *et al.*, 1984) and the polypeptide is variably glycosylated at two sites (Hemmi *et al.*, 1989). Two of these polypeptide chains associate to form an obligate non-covalent homodimer (Nagata *et al.*, 1987).

The dimeric nature of IFN-γ was confirmed by the 3.5 Å resolution X-ray crystal structure of human IFN-γ (1 hig; Ealick *et al.*, 1991). This structure also showed the molecule to be primarily helical, with each monomer consisting of six α-helices connected by short linkers. IFN-γ is a two-domain protein; each domain consists of the first four helices of one monomer and the two terminal helices of the other monomer. A similar interdigitation of monomers was first described for *trp* repressor (Schevitz *et al.*, 1985) and was later seen in the crystal structure of IL-10 (Walter & Nagabhushan, 1995; Zdanov *et al.*, 1995).

The existence of IFN-γ as a homodimer has important biological implications. It is generally recognized that hormone-induced receptor activation of the IFN family of cytokines is triggered by an aggregation event similar to that seen for other cytokine families (Sprang & Bazan, 1993). It has been determined that the active IFN-γ complex has a 1:2:2 stoichiometry, represented by one IFN-γ homodimer binding two copies each of two different receptors: the receptor α-chain (IFN-γRα) and the receptor β-chain (IFN-γRβ; Bach *et al.*, 1997). There has been an intense and, recently, successful effort to obtain structures of the extracellular portions of the complex between the hormone and IFN-γRα (Walter *et al.*, 1995; 2.0 Å resolution, Randal & Kossiakoff, unpublished results).

An important element of the hIFN-γ–IFN-γRα analyses is the determination of conformational changes within the IFN-γ structure that occur as the result of its binding to the receptor. Unfortunately, good models at an appropriate resolution are not available. The hIFN-γ structure (1hig) has not been fully refined because of poor crystal quality and only the C^{α} chain trace has been published. Subsequent to that study, the structures of rabbit (2rig) and bovine (1rfb) IFN-γ were reported at 2.7 Å (R = 0.262) and 3.0 Å (R = 0.192), respectively. We note that a cross comparison of these models indicated the presence of some substantial structural differences, even in ordered segments of the main chains. Although IFN-γ is highly species specific in triggering its biological activities (Farrar & Schreiber, 1993), differences in the registering of some of the helices was unexpected based on the high degree of sequence similarity (Fig. 1).

The rabbit and bovine structures are also very different to human models obtained from the analyses of the hIFN-γ–Ra complexes. Within the error limits associated with using a low-resolution C^{α} chain trace for comparison, the structures of the bound and free forms of hIFN-γ appear to be similar; certainly they are more similar than free hIFN-γ is to either the free rabbit or bovine models. It is therefore important to determine whether the structural differences among the different molecules are attributable to species specificity or conformational variation on binding, or whether they are merely a reflection of experimental and interpretive issues in the structure-solution process. The latter issue remains a real possibility because of both the presence of stereochemical inconsistencies and the sparseness of the data used in the refinements of the rabbit and bovine structures. Sorting out these important issues requires a much higher resolution and refined model. Efforts to overcome the diffraction limitations of hIFN-γ crystals were not successful, so we therefore chose to concentrate on obtaining high-quality crystals of bovine IFN-γ.

We report here the structure of bovine IFN-γ (BOV) at 2.0 Å resolution. A structure solution was initially attempted by molecular replacement using both the human and bovine models. However, given the desire to circumvent model bias, structure solution was carried out using the multiple isomorphous replacement method. The structure was first determined in a lower resolution (2.9 Å) crystal form (form 1), with the resulting model being transferred to the higher resolution crystal form (form 2). These two crystal forms have the same internal symmetry ($P2_12_12_1$), but they differ significantly in the lattice packing along one of the principal axes. This change was brought about by a combination of modifying the crystallization conditions and flash-cooling, leading to a more tightly packed lattice and a dramatic improvement in resolution. Comparison of the BOV model with the three available hormone models indicates a close similarity to the human structure. However, there are significant differences between this structure and the previously determined bovine and rabbit structures. In addition, the BOV model is very similar to the 2.0 Å structure of hIFN-γ bound to the extracellular domain of its high-affinity receptor (Randal & Kossiakoff, unpublished results). We feel that there are

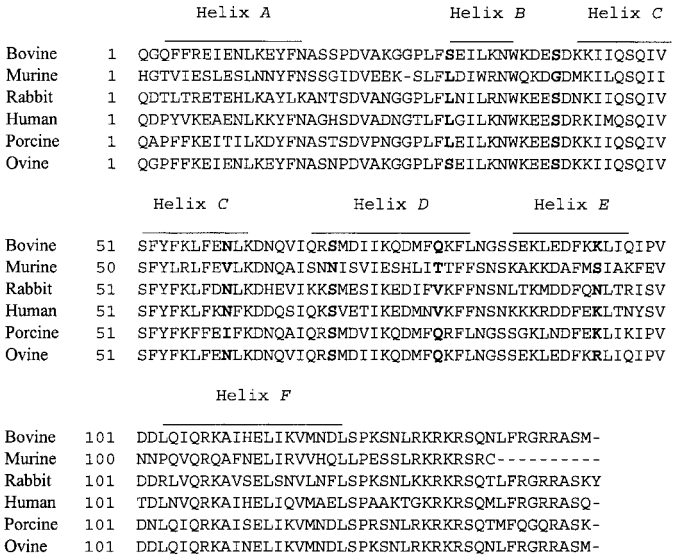

Figure 1
Sequence alignment of IFN-γs from multiple species. Sites chosen for cysteine incorporation are indicated in bold. Amino-acid identities to bovine: murine, 44%; rabbit, 65%; human, 62%; porcine, 79%; ovine, 97%.

interpretive errors in the published rabbit and bovine structures and we discuss the methods of evaluation on which we base our conclusion.

2. Experimental

2.1. Expression and purification

A plasmid that directed the intracellular expression of bovine IFN-γ (residues 1–143) under the control of the T7 promoter was used. This plasmid contained the f1 origin of replication to allow mutagenesis by the method of Kunkel (1985). Following transformation of the wild-type or mutant plasmid into *Escherichia coli* strain BL21, cells were grown in Luria broth until late log phase (OD_{600} = 0.6–0.9) and induced with 1 m*M* IPTG. Following induction, cells were grown for an additional 3–5 h, after which they were pelleted by centrifugation. A purification procedure was developed based on the high isoelectric point of IFN-γ (pI = 10) and its ease of refolding. Pelleted cells were resuspended in 2 *M* urea, 25 m*M* Tris pH 7.5, 25 m*M* NaCl, 5 m*M* EDTA, 0.1% β-mercaptoethanol and were disrupted by sonication. Cell debris was pelleted by centrifugation. The resulting supernatant was loaded onto S-Sepharose fast flow (Pharmacia), followed by elution with a linear NaCl gradient. Fractions containing IFN-γ were pooled and refolded by dialysis into 25 m*M* Tris pH 7.5, 25 m*M* NaCl. This refolding step resulted in quantitative recovery of the IFN-γ. It also provided further purification by the precipitation of contaminant proteins that did not refold by this procedure. Although at this stage the preparation was essentially free of all *E. coli* proteins, it existed as multiple bands on an SDS–PAGE gel owing to truncations at the C-terminus. This heterogeneity was similar to that seen for native IFN-γ *in vivo* (Rinderknecht *et al.*, 1984) and is caused by proteolytic cleavage at one of the eight positively charged amino acids in the last 17 residues of the full-length molecule. To create a homogeneous C-terminus, the dialyzed pool was subjected to limited proteolysis with trypsin [1:200(*w*/*w*), 30 min, room temperature]. This digestion resulted in the coalescence of the multiple bands seen by SDS–PAGE into a single band. Mass spectrometry showed that the digested pool consisted of approximately 80% material cleaved after residue 127, with the remainder cleaved after residue 128 (data not shown). This material was further purified on a Pharmacia FPLC Mono-S column using a linear NaCl gradient in the above buffer. To ensure that all the material was dimeric, a final size-exclusion chromatography step was performed using Sephadex S-100 (Pharmacia) equilibrated with 50 m*M* MES pH 6.0, 0.5 *M* NaCl. The pooled material was concentrated to 4 mg ml^{-1} for crystallization using Centriprep and Centricon 10 concentrators (Amicon).

2.2. Crystallization

Crystals were grown by standard hanging-drop vapor-diffusion techniques. 5 µl each of protein and reservoir were mixed and equilibrated against 1 ml reservoir in 24-well Linbro plates. Crystallization trials produced two related crystal forms. Each form contained one hormone dimer per asymmetric unit. Crystal form 1 was grown using a reservoir solution of 40% saturated ammonium sulfate. These crystals belong to space group $P2_12_12_1$, with unit-cell dimensions a = 42.6, b = 79.6, c = 82.4 Å. This crystal form has been previously reported (Samudzi & Rubin, 1993). Assuming one hormone dimer per asymmetric unit, V_m = 2.3 Å^3 Da^{-1} (Matthews, 1968), which corresponds to a solvent content of ~47%. Prior to data collection, crystals were equilibrated in a mother liquor containing 45% saturated ammonium sulfate. For data collection, crystals were mounted in wax-sealed 0.7 mm glass capillaries. Crystal form 1 diffracts to approximately 2.9 Å using a rotating-anode source (45 kV, 120 mA).

Crystal form 2 was grown from a reservoir solution of 50% saturated sodium citrate and also belongs to space group $P2_12_12_1$. The unit-cell dimensions are a = 36.2, b = 81.1, c = 83.9 Å, indicating a close similarity to form 1 crystals in the b and c axes and a 15% reduction of the a axis. Assuming one homodimer per asymmetric unit, these crystals have V_m = 2.0 Å^3 Da^{-1}, corresponding to a solvent content of 39%. Prior to data collection, these crystals were transferred to a stabilizing solution (40% saturated sodium citrate, 20% saturated ammonium sulfate, 10% saturated sodium succinate, 0.5 *M* sodium chloride) followed by 5 min soaks in mother liquors containing 8 and 15% glycerol. Treated crystals were then scooped into a small nylon loop (Hampton) and flash-cooled by manually plunging the loop into liquid nitrogen. This crystal form diffracts to approximately 2.6 Å before freezing and 2.0 Å after freezing using a rotating-anode source (see below). This rather dramatic increase in resolution results from both the ordering of the crystal lattice and the elimination of decay, allowing longer exposure times.

2.3. Data collection

All data were collected using a 300 mm MAR Research image-plate system mounted on a Rigaku RU2000 rotating-anode generator operated at 45 kV and 120 mA. Data for the form 1 crystals were collected at room temperature using 1° rotations at a crystal-to-detector distance of 250 mm. Crystals were manually aligned with one crystallographic axis slightly skewed (~15°) to the rotation axis and a second axis coincident with the X-ray beam, minimizing the rotation range needed to collect a complete data set. Form 2 crystals were mounted randomly and maintained at 100 K during data collection using an Oxford Cryosystems Cryostream. At this temperature, crystal decay was essentially eliminated, so shorter rotations were used (0.25°) to increase the signal-to-noise ratio of the collected data.

Data were collected from three different IFN-γ variants in crystal form 1. These included the mutation Asn59→Cys (N59C), N59C reacted with mercuric acetate (N59C-Hg) and the mutation Ser30→Cys reacted with mercuric acetate (S30C-Hg). Data were collected from one native form 2 crystal. All data were reduced using *MARXDS* and scaled using *MARSCALE* (Kabsch, 1988). A summary of the data reduction is presented in Table 1.

Table 1
Summary of the X-ray data-collection statistics.

Variant	Distance (mm)	Rotation (°)	Resolution (Å)	Number of observations	Number of reflections	Completeness (%)	R_{sym} (I)† (%)
Crystal form 1							
S30C-Hg	250	88	20–3.15	17174	5214	98	10.0
N59C	250	98	20–3.2	17431	4875	98	8.9
N59C-Hg	250	89	20–2.9	24286	6484	89	6.2
Crystal form 2							
Native	150	131	27–2.0	71885	16274	97	3.0

† $R_{sym} = \sum |I - \langle I \rangle| / \sum I$.

2.4. Phasing and refinement of bovine IFN-γ

Attempts to solve the bovine IFN-γ structure solution by molecular replacement (MR) were based on the available C^{α} coordinates of human IFN-γ (1hig; Ealick *et al.*, 1991). A polyalanine model of the helical regions of 1hig was constructed manually using stretches of idealized helix within the program *MIDASPLUS* (Ferrin *et al.*, 1988). To better model the non-ideality of the actual IFN-γ helices, C^{α} positions were restrained to those found in 1hig and the model energy was minimized in *X-PLOR* (Brünger *et al.*, 1987). Side chains were added in the rotamer most often associated with helical secondary structure (L. Presta, personal communication) and then energy minimized to ensure steric compatibility. The resulting model was used in a standard four-step *X-PLOR* molecular-replacement protocol consisting of a rotation search, Patterson correlation refinement, translation search and rigid-body refinement (Brünger, 1992*b*). Molecular replacement was attempted with data from both crystal forms. More interpretable results were obtained from crystal form 1 (see below). This is probably a consequence of the higher solvent content of crystal form 1, resulting in a smaller contribution from intermolecular vectors to the experimental Patterson map. The rotation search was carried out using all data in the resolution range 15.0–3.5 Å and a maximum vector length of 25 Å (corresponding approximately to the two shortest dimensions of the molecule). The top 115 peaks from the rotation search were used for Patterson correlation (PC) refinement with the model treated as a single rigid body (Brünger *et al.*, 1987). This procedure identified two clear maxima (Patterson correlation = 0.194, with a highest background peak of 0.06) corresponding to the two possible orientations of the symmetric molecule. The correctly rotated molecule was used in a translation search against 10.0–3.5 Å data on a 1 Å search grid. This resulted in a clear solution with a peak height of 0.58 (compared with the highest background peak of 0.39) and an R value of 39.5%. Rigid-body refinement reduced the R value to 37.5% for all data in the resolution range 10.0–3.0 Å.

Despite this clear molecular-replacement solution, this phasing model was not accurate enough to allow corrections or extensions to the model. The inaccuracy probably arises from a combination of two factors: model completeness and coordinate error. The model included about two-thirds of the amino acids (176 of 256 residues) and has a r.m.s.d. for main-chain atoms (including C^{β}) of 1.2 Å from the final refined model. These differences were distributed randomly throughout the molecule and result in a 91° mean phase difference between the molecular-replacement model and the final refined model. Despite the magnitude of this error, these phases did prove to be useful at later stages of the structure solution (see below).

2.5. Heavy-atom incorporation *via* cysteine substitution

The structures in either of the crystal forms could not be solved completely by molecular replacement using any of the available structures. A partial solution was found using a model constructed by adding side chains on the C^{α} trace of hIFN-γ; however, it was not accurate enough to initiate refinement (see above). The successful approach to phase determination involved the incorporation of cysteine mutations into the protein as sites for covalent attachment of mercury. This technique was shown to be successful in several other instances where traditional heavy-atom screening had failed to produce isomorphous derivatives (Sun *et al.*, 1987; Hatfull *et al.*, 1989; Martinez *et al.*, 1993). The absence of cysteine residues in native IFN-γ made this approach especially attractive.

Previous studies have shown empirically that the most effective sites for cysteine derivatization have been at locations that are neither completely exposed nor completely buried. To achieve these criteria, the cysteine sites picked had the following characteristics: (i) they were at positions of defined secondary structure in the human model, (ii) they showed sequence variability among IFN-γs from different species and (iii) this sequence variability included both hydrophilic and hydrophobic residues.

Six sites were chosen for cysteine substitution based upon sequence alignment (Fig. 1). Mutants, constructed by standard Kunkel mutagenesis (Kunkel, 1985), were expressed and purified by the same procedure as the native protein. During purification, it was observed that the Ser40→Cys (S40C) mutation produced covalently bound dimers. This mutant was eliminated from further analysis. Type 1 crystals were grown as previously described, except that 1 m*M* dithiothreitol (DTT) was included to prevent cysteine oxidation. This crystal form was chosen because crystal growth was more reproducible and the higher solvent content would enhance phase improvement by solvent flattening. Following growth, crystals were transferred to a mother liquor consisting of 45% saturated ammonium sulfate. Derivatization was achieved by a 24 h soak in mother liquor containing 1 m*M* mercuric acetate without DTT.

Data were collected and reduced as previously described and are summarized in Table 1. The first two mutants tested (S30C and N59C) produced isomorphous crystals and proved to be adequate for structure determination. Soaking both

native and derivatized crystals in the identical mother liquor prior to data collection was an important factor in maintaining isomorphism. Using the approximation of Crick & Magdoff (1956), the occupancies of the mercury sites were estimated to be over 90%. Heavy-atom sites for these derivatives were located in *PROTEIN* (Steigemann, 1974) difference Fourier maps using phases calculated from the molecular-replacement model of hIFN-γ. Peak heights of 9.0 and 8.7 with a highest background peak of 4.6 were observed for the S30C derivative and peak heights of 9.4 and 8.1 with a highest background peak of 4.4 were observed for the N59C derivative. These peaks were located at positions consistent with the site of mutation in the model.

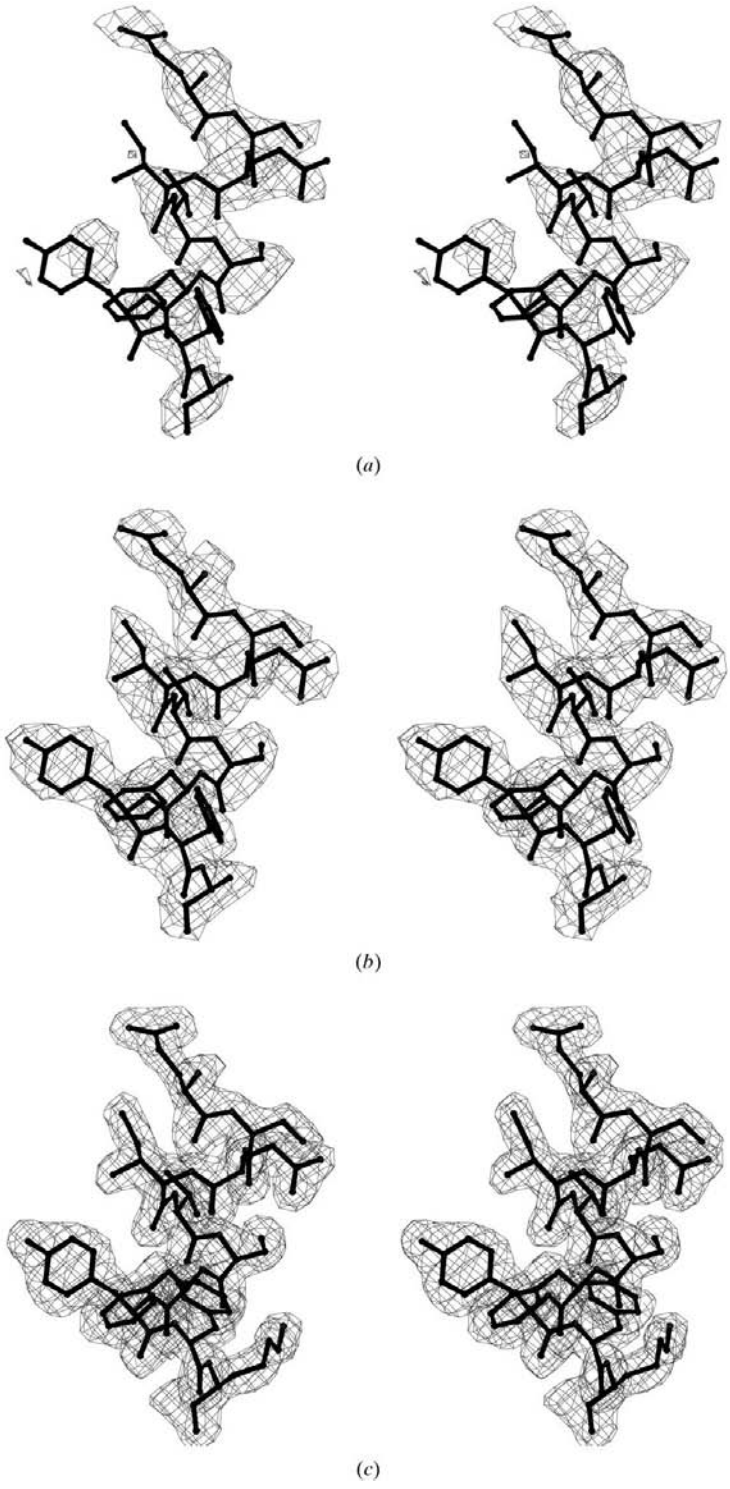

Figure 2
Stereoview of a portion of the electron-density map superimposed on residues 36–55 for (*a*) the NCS-averaged solvent-flattened MIR map, (*b*) the crystal form 1 $2F_o - F_c$ map (1σ, 2.9 Å) and (*c*) the crystal form 2 $2F_o - F_c$ map (1σ, 2.0 Å).

All heavy-atom site refinement and density modifications were performed using the *PHASES* suite of programs (Furey & Swaminathan, 1998). Heavy-atom positions were refined using only the isomorphous differences to a resolution of 3.2 Å, with the *B* factor arbitrarily set to 20 Å^2. The final $R_{\rm Cullis}$ ($\sum\varepsilon/\sum\Delta F$, with the sum taken over all centric reflections; ε is the lack of closure) was 0.55 and 0.69 for the S30C and N59C derivatives, respectively. The mean figure of merit was 0.51. The phasing power ($\sum F_H/\sum\varepsilon$) was 1.77 and 1.26, respectively, for the two derivatives.

Phases were subjected to five cycles of solvent flattening. This increased the figure of merit to 0.80, though the average phase error compared with the final refined model was unchanged (70.4 and 70.5° before and after solvent flattening, respectively). Finally, phases were extended to 2.9 Å in eight resolution steps (3.20, 3.15, 3.10, 3.05, 3.00, 2.96, 2.93, 2.90 Å) using cycles of iterative twofold non-crystallographic symmetry averaging and solvent flattening. The final figure of merit was 0.62 to 2.9 Å resolution. The average phase error compared with the final model was 72.6°.

2.6. Model building and refinement

2.6.1. Crystal form 1. The initial MIR map (Fig. 2*a*) was of sufficient quality to build most of the molecule using the model derived from 1hig as a guide (*FRODO*; Jones, 1978, 1985). Model improvement proceeded *via* iterative cycles of map fitting ($2F_o - F_c$ and $F_o - F_c$) and *X-PLOR* positional and temperature-factor refinement (Brünger, 1992*b*). 10% of the data were excluded from all stages of refinement for calculation of a free *R* value (Brünger, 1992*a*). During the initial stages of refinement, phases were tightly restrained to their experimental value to minimize the effects of model bias. After each cycle of model building, these restraints were relaxed based upon the behavior of the free *R* value. This resulted in a gradual decrease in the magnitude of the phase restraint as the model completeness and accuracy improved.

Non-crystallographic symmetry restraints (reflecting the twofold symmetry of the molecule) were applied to individual helices at all stages of refinement. Appropriate weighting was determined empirically by the behavior of the free *R* value. Grouped temperature-factor (*B*) refinement was used and the data were corrected for anisotropy (B = −8.6, 5.9 and 2.7 Å^2 for the *a*, *b* and *c* axes, respectively). Refinement progressed

until difference maps ($F_o - F_c$) no longer contained features that could be addressed by adjustments to the model. The final R value is 21.8% (R_{free} = 30.9%) for all data in the resolution range 8–2.9 Å. A sample of the electron density is shown in Fig. 2(*b*).

2.6.2. Crystal form 2. While crystal form 1 was more tractable for MIR structure solution (see above), crystal form 2 provided higher resolution diffraction. The crystal form 1 model was used in a standard *X-PLOR* molecular-replacement protocol against data from crystal form 2. As may be inferred from the similarity of the two unit cells, this solution consisted almost entirely of a shift along the crystallographic *a* axis ($\theta_1 = -2.2$, $\theta_2 = -3.7$, $\theta_3 = 2.2^\circ$; −10.5, 1.5 and 2.0 Å shifts along the *a, b* and *c* axes, respectively). Rigid-body refinement reduced the R value from 49.5 to 46.3% for the resolution range 8.0−2.9 Å. Positional and temperature-factor refinement resulted in a model with R = 30.9% and R_{free} = 41.9% for the resolution range 8.0–2.5 Å.

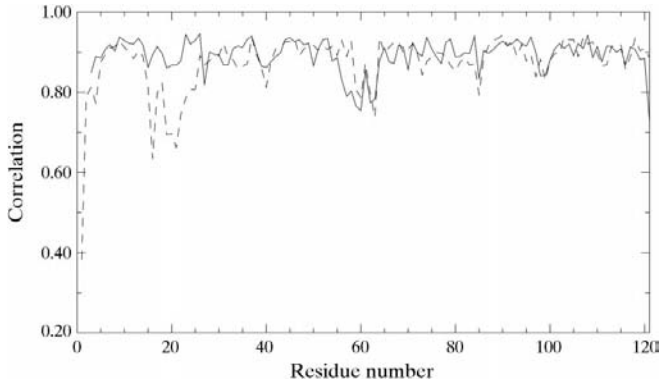

Figure 3
Average main-chain real-space density correlation for the final $2F_o - F_c$ map. Solid line, monomer 1; dashed line, monomer 2.

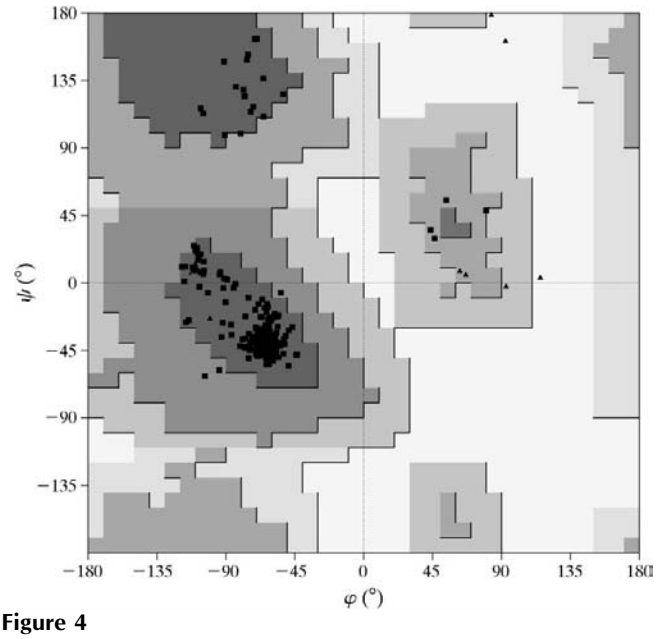

Figure 4
Ramachandran plot for the final crystal form 2 model. Glycines are represented by triangles, all other residues by squares.

Refinement proceeded *via* iterative cycles of map fitting (*O*; Jones *et al.*, 1991) and positional and individual isotropic temperature-factor refinement (*X-PLOR*; Brünger, 1992*b*) with 10% of the data (1612 reflections) omitted for R_{free} calculation (Brünger, 1992*a*). Water molecules were placed at peak positions of $F_o - F_c$ electron density greater than 3σ and appropriate hydrogen-bonding stereochemistry. They were removed from the model if their density dropped below 1σ in a $2F_o - F_c$ electron-density map (129 waters included in the final model). Non-crystallographic symmetry restraints were applied to individual helices at all stages of refinement. The data were corrected for anisotropy (B = −6.4, 7.9 and −1.5 Å^2 for the *a, b* and *c* axes, respectively) and bulk-solvent scattering (solrad = 1.3 Å, k = 0.35 e Å^3, B = 58 Å^2; Jiang & Brünger, 1994). Refinement proceeded until no interpretable features remained in difference maps ($F_o - F_c$). The final R value for all data in the resolution range 15.0–2.0 Å is 19.7%; R_{free} = 27.5%. Fig. 2(*c*) shows an example of the final electron-density map.

2.7. Model accuracy

The refined model consists of residues 1–121 of monomer *A*, residues 3–121 of monomer *B* (*A* and *B* referring to the two

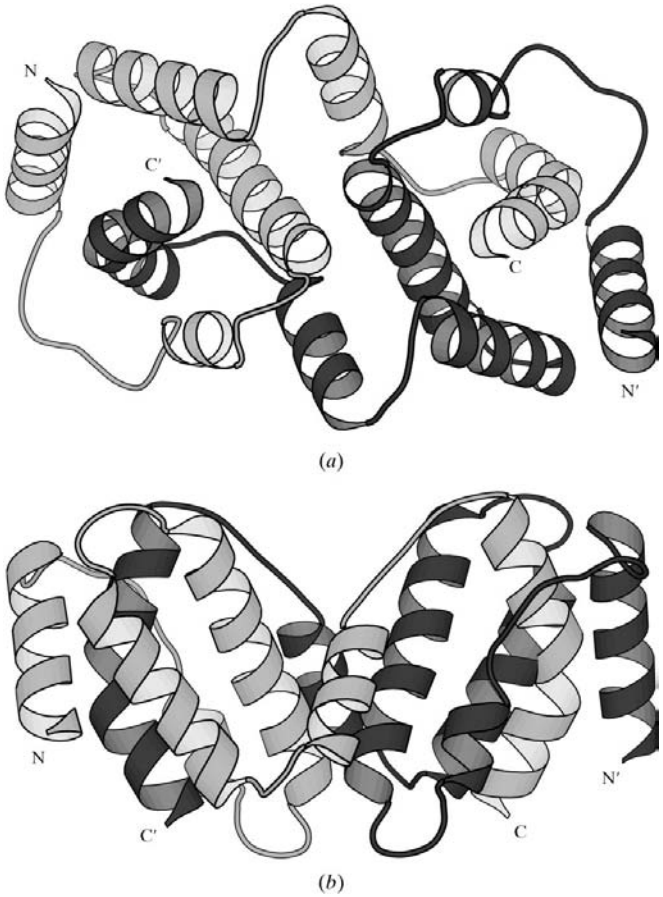

Figure 5
Ribbon representation of the bovine IFN-γ dimer. (*a*) View approximately parallel to the dimer twofold axis; (*b*) view approximately perpendicular to the dimer twofold axis.

independent polypeptides of the homodimer) and 129 water molecules (Table 2). The six C-terminal residues could not be placed in the electron density and static disorder at residues 4*A* and 16*A* was modeled by two alternate side-chain conformations. Side-chain density was not observed for residues 24, 34 and 37 of monomer *A* and residues 3, 24, 34, 37, 61, 64, 68, 87 and 97 of monomer *B*. The average temperature factor of the model is 29.8 Å^2, with r.m.s.d.s for bonded main-chain and side-chain atoms of 2.6 and 4.9 Å^2, respectively. The model has good stereochemistry; r.m.s.d.s for bond lengths and angles (Engh & Huber, 1992) are 0.013 Å and 1.47°, respectively. The estimated coordinate error as calculated by the method of Luzzati (1952) is 0.25 Å. The average real-space density correlation for main-chain atoms is 0.88 (Jones *et al.*, 1991). The density correlation drops below 0.7 at the N-terminus and for the *AB* loop of monomer *A* (Fig. 3), suggesting disorder in these regions. The Ramachandran plot (Ramachandran *et al.*, 1963) of main-chain torsion angles is shown in Fig. 4. 95.5% of the residues exist in the most favored regions as defined by *PROCHECK* (Laskowski *et al.*, 1993), with the remaining 4.5% falling in additional allowed regions.

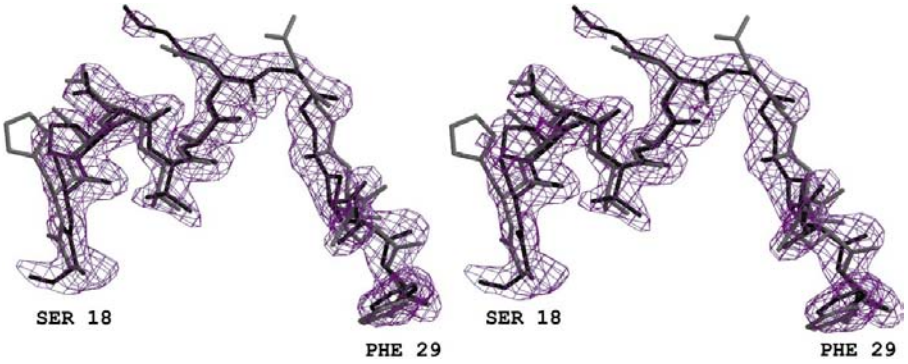

Figure 6
Stereoview of the *AB* loop for bovine (black bonds) and human (grey bonds) IFN-γ superimposed on the bovine electron density ($2F_o - F_c$, 1σ).

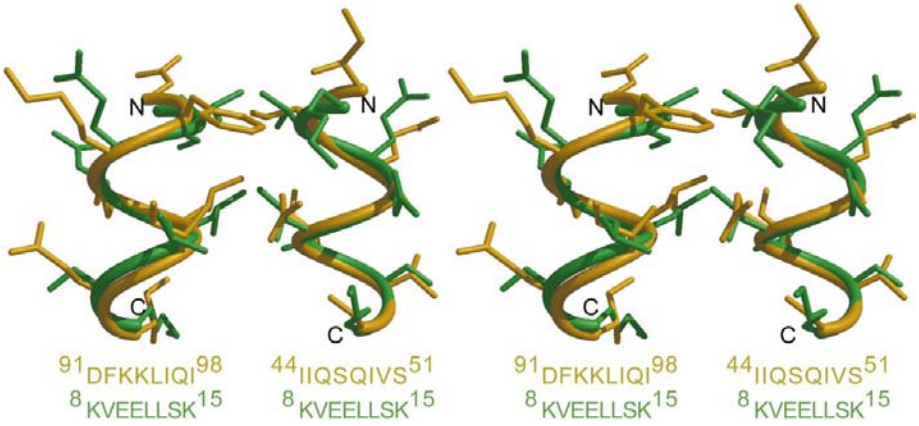

Figure 7
Stereoview of residues 44–51 and 91–98 of the IFN-γ homodimeric interface (yellow) superimposed on residues 8–15 of the GCN4 leucine zipper (green).

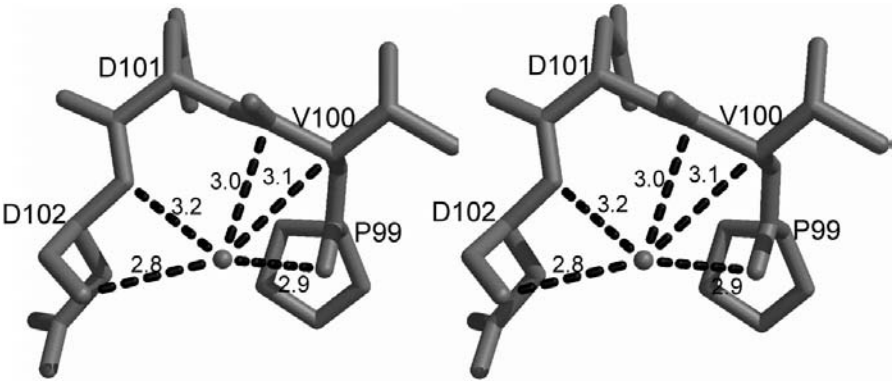

Figure 8
Stereoview of the *EF* loop showing a conserved water molecule. This water makes three hydrogen bonds (Pro99 O, Asp102 N, Asp102 O) and two close contacts (Val100 C$^{\alpha}$, Val100 C). Distances are in Å.

3. Results

3.1. Overall structure

A description of the structure of human IFN-γ has been reported previously, both in a free state (3.5 Å resolution; Ealick *et al.*, 1991) and in complex with two copies of its high-affinity receptor (IFN-γRα; Walter *et al.*, 1995). The polypeptide chain consists of six α-helices (*A*–*F*) related by a twofold non-crystallographic symmetry axis (Fig. 5). The protein consists of two domains, each domain including the four N-terminal helices of one monomer and the two C-terminal helices of the second monomer. These helices range in length from six to 18 residues. The intervening loops are between five and seven residues in length, with the exception of the 15-residue loop connecting helices *A* and *B*. Helix *F* is characterized by a 51° bend (*HELIXANG*; Collaborative Computational Project, Number 4, 1994) at residue Glu112, facilitating its packing into a groove created by symmetry-related helices *A*–*D*. Previous X-ray and NMR (Grzesiek *et al.*, 1992) studies have shown the long *AB* loop is flexible in free human IFN-γ. Upon binding to IFN-γRα, this loop becomes ordered, including the formation of a 3_{10} helix over residues 20–23 (Walter *et al.*, 1995). In the

Table 2
Refinement statistics.

Model	
No. of residues	240
No. of solvent molecules	129
No. of non-H atoms	2082
Missing residues	
Chain *A*	121–127
Chain *B*	1–2; 121–127
Alternate conformations	4*A*, 16*A*
Refinement	
Resolution (Å)	15.0–2.0
Completeness (%)	97
R value ($F > 0$) (%)	19.7
No. of reflections	14481
Free *R* value ($F > 0$) (%)	27.5
No. of reflections	1612
Anisotropic correction (Å^2)	12.5
Stereochemistry	
R.m.s.d. in bonds (Å)	0.013
R.m.s.d. in angles (°)	1.47
R.m.s.d. *B* for bonded atoms	
Main chain (Å^2)	2.6
Side chain (Å^2)	4.9

present bovine model, this loop is essentially identical to that of the bound human hormone; the r.m.s.d. for C^{α} atoms 18–29 is 0.5 Å between the free bovine structure and the 2.0 Å crystal structure of the human hormone bound to IFN-γRα (Fig. 6; Randal & Kossiakoff, unpublished results).

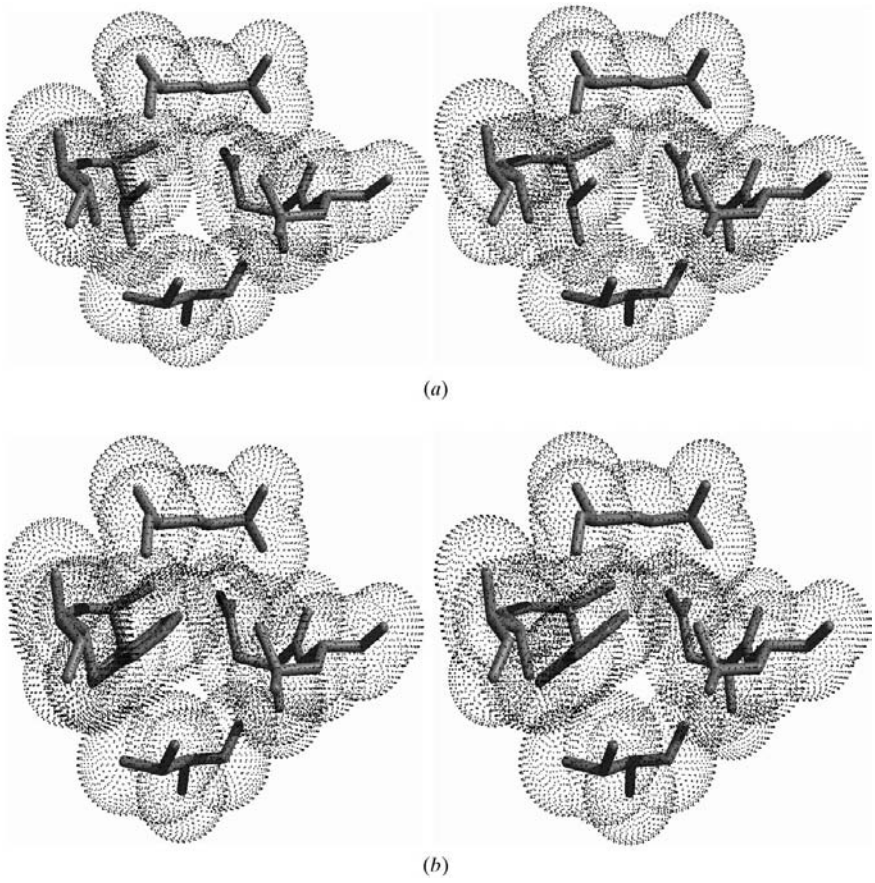

Figure 9
(*a*) van der Waals surface representation of the residues lining the buried cavity in bovine IFN-γ. (*b*) Substitution of a low-energy conformation of Trp at residue 33 results in the elimination of this cavity.

Table 3
Conserved inter-chain hydrogen bonds (<3.5 Å).

The two distances refer to the two independent monomers.

Atom 1	Atom 2	Distance 1 (Å)	Distance 2 (Å)
Phe15 O	Arg107 N^{ε}	3.2	3.2
Pro27 O	Lys94 N^{ζ}	3.4	3.4
Leu28 N	Glu112 $O^{\varepsilon 2}$	3.0	3.2
Leu28 O	Lys94 N^{ζ}	3.1	3.1
Phe29 N	Glu112 $O^{\varepsilon 1}$	2.6	2.7
Ser30 N	Glu112 $O^{\varepsilon 1}$	3.1	3.2
Asn35 $N^{\delta 2}$	Asp91 $O^{\delta 2}$	3.4	3.0
Trp36 $N^{\varepsilon 1}$	Asp91 $O^{\delta 2}$	2.7	2.8
Glu38 $O^{\varepsilon 1}$	Lys88 N^{ζ}	3.0	3.2
Lys42 N^{ζ}	Asp119 $O^{\delta 1}$	2.7	3.3
Ile44 O	Ser47 O^{γ}	2.7	2.5
Gln46 $N^{\varepsilon 2}$	Asp119 $O^{\delta 2}$	2.8	2.6
Ser47 O	Gln48 $N^{\varepsilon 2}$	3.1	3.2
Tyr53 O^{η}	Ala109 O	2.6	2.7

3.2. Dimer interaction

The interdigitating helices of the IFN-γ homodimer result in extensive interactions between the two chains. The surface areas buried by this association is 3670 and 3690 Å^2 for the two chains (*AREAIMOL*; Collaborative Computational Project, Number 4, 1994). This buried surface is mostly hydrophobic and is made up of the ratio 66:20:13:1 carbon:oxygen:nitrogen:sulfur. A portion of this dimeric interface is made up of helical contacts reminiscent of those seen in leucine zipper coiled coils. Residues 44–51 of one monomer and residues 91–98 of the second monomer superimpose on residues 8–15 of GCN4 (2zta; O'Shea *et al.*, 1991) with an r.m.s.d. of 0.6 Å for all main-chain atoms (Fig. 7).

There are 14 hydrogen bonds which are found in both monomers (Table 3). Of these, nine are side-chain–main-chain hydrogen bonds, while five are side-chain–side-chain hydrogen bonds. 11 of the hydrogen bonds exist at or near the surface of the molecule, while three are buried. Two of the buried hydrogen bonds (Ile44 O$\cdots$Ser47 O^{γ}; Tyr53 O^{η} $\cdots$Ala109 O) are at the domain–domain interface. The other is within a domain between the hydroxyl group of Tyr53 and the carbonyl group of residue 109. This tyrosine is part of a conserved structural core similar to that in IL-10.

3.3. Solvent model

Of the 129 water molecules included in the final model, 50 exist as pairs related by the molecular symmetry (defined as being within 1.2 Å of each other after applying the non-crystal-

lographic symmetry). 13 of the 25 related pairs hydrogen bond to hydrophilic groups on the surface of the protein. Additionally, 11 pairs of waters satisfy unmet hydrogen bonds at the ends of helices (seven carbonyl and four amide). One pair of waters exists at the center of the *EF* loop, making hydrogen bonds to the carbonyl O atom of residue 99 (2.9 Å) and both the carbonyl O atom (2.8 Å) and amide NH (3.2 Å) of residue 102. This water also makes close contacts with the carbonyl C atom (3.0 Å) and the C^{α} atom (3.1 Å) of residue 100 (Fig. 8).

While the average *B* factor of these conserved waters does not differ from that of the remainder of the waters (35 *versus* 37 Å^2), they are generally more buried (15 Å^2 exposed surface area *versus* 42 Å^2). Using this calculation, a fully exposed water molecule has 113 Å^2 of surface area. There are five buried water molecules (<5 Å^2 solvent-accessible area) that are not found as symmetry-related pairs. Four of these are in the vicinity of side chains that are ordered in only one of the two monomers. It is not apparent why the remaining water has no symmetry pair. Although the symmetry-related environment is somewhat smaller owing to an ∼0.7 Å movement of helix *A* towards the water position, there appears to be both the room and the requisite hydrophilic environment to satisfy a water molecule. Additionally, the temperature factors of these hydrogen-bonding partners are essentially identical in both halves of the homodimer.

There are two cavities (related by the molecular symmetry) that do not contain ordered solvent. These cavities are approximately 5 Å from the surface of the molecule and have a volume of 30 Å^3 as calculated using *GRASP* (Nicholls *et al.*, 1991). The cavities are essentially apolar, lined by the hydrophobic side chains of residues Leu33, Ile49, Val116, Leu120 and the methylenes of the side chain of Gln46 (Fig. 9*a*). Three O atoms contribute to the walls of this cavity, but they are already associated in hydrogen bonds: two carbonyl groups (residues 46 and 116) in main-chain hydrogen bonds and the carboxylate of residue Asp19, which is hydrogen bonded to the side chain of residue Gln46.

Based on the volume of the cavity, there is sufficient space for a water molecule. However, volume alone is not a sufficient condition; the shape of the cavity is also a determining feature. In this case, although the shape of the pocket is somewhat asymmetric, a water molecule can be modeled to fit into it. The electron density in the Fourier difference map in this area has low-level negative density ($F_o - F_c = -1.0\sigma$, $2F_o - F_c = -0.45\sigma$) suggesting that it is 'empty'. Using the estimation of Eriksson *et al.* (1992), a cavity of this volume would destabilize the protein by about 0.7 kcal mol^{-1} [$\Delta G = 0.024\ (\Delta V)$ kcal mol^{-1} where 1 kcal mol^{-1} = 4.184 kJ mol^{-1}]. Examination of IFN-γ sequences from six species (Fig. 1) revealed only a single mutation in the five residues lining the cavity, Leu33→Trp in mouse. Modeling of the bovine structure indicates that the substitution of a low-energy rotamer of Trp at position 33 results in elimination of this cavity (Fig. 9*b*).

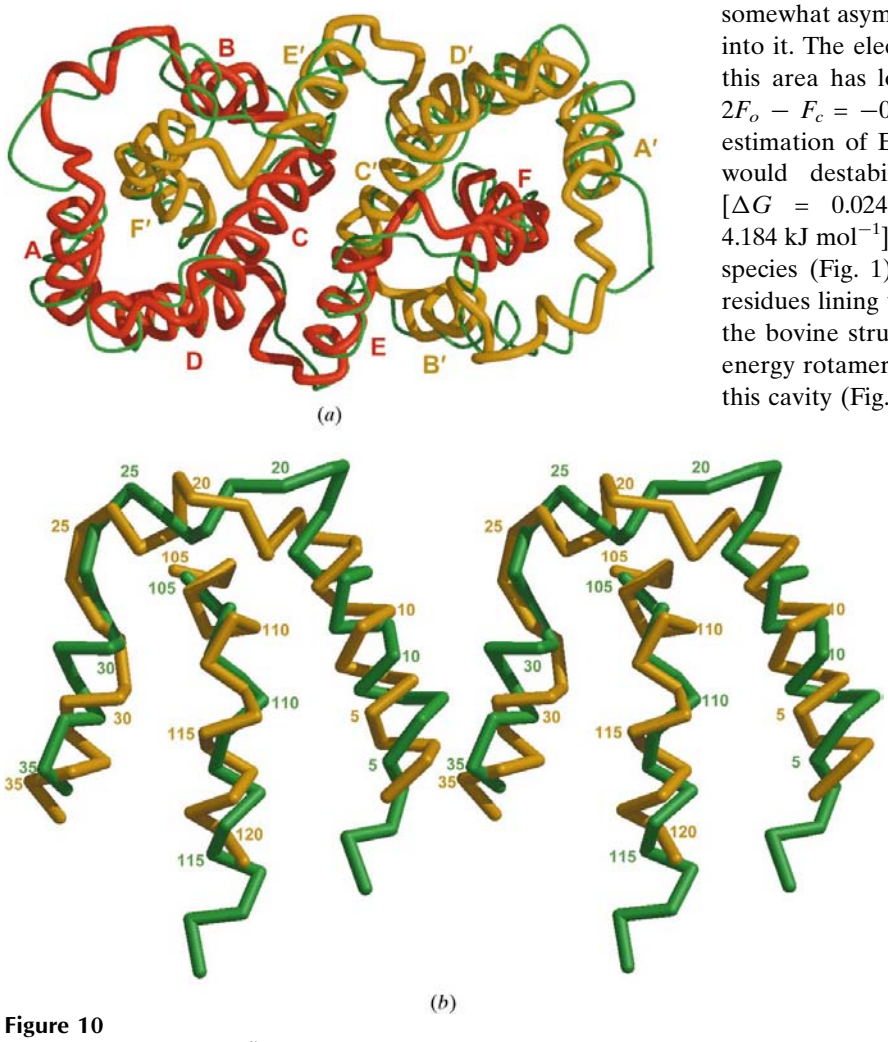

Figure 10
(*a*) Superposition of the C^{α} traces for the refined crystal form 2 model (yellow and orange thick lines) and 1rfb (green thin lines). (*b*) Stereoview of the exposed helices (*A, B* and *F*) of bovine IFN-γ after a superposition was made based on the buried helices (*C, D* and *E*). Yellow, the refined crystal form 2 model (BOV); green, the 1rfb model.

4. Discussion

4.1. Comparison with other IFN-γs

Structural models of IFN-γ from several species now exist. Comparisons of these models show they share the same general topology; however, there are some significant structural differences. Table 4 lists the C^{α} r.m.s.d.s for the pairwise superposition of the four available structures: the present high-resolution bovine model (2.0 Å; hereafter referred to as BOV), the human model (3.5 Å; 1hig; Ealick *et al.*, 1991), a lower resolution (3.0 Å) bovine model (1rfb; Samudzi & Rubin, 1993) and the rabbit model (2.7 Å; 2rig; Samudzi *et al.*, 1991). Based on this comparison, the structures separate into two groups: one containing 1hig and BOV, the other containing the 1rfb and 2rig. The 1rfb and 2rig structures are

reported from the same group; the BOV and 1hig structures were determined in different laboratories.

1hig and BOV have an r.m.s.d. of 1.2 Å, while 1rfb and 2rig have an r.m.s.d. of 2.5 Å (helical C^{α} atoms). Pairwise comparisons between these groups give r.m.s.d.s of 4.0–4.3 Å. These two groups are similar in the core of the molecule (helices *C, D* and *E*), but differ extensively in the register of residues along the more exposed helices (*A, B* and *F*). Fig. 10 shows a structural overlay of BOV and 1rfb after a superposition was made based on the C^{α} atoms of the central helices (*C, D* and *E*; 90 C^{α} atoms). The fact that these molecules were crystallized from similar conditions and have essentially identical unit cells suggests that the large differences between these two models arises from experimental and interpretive errors.

4.2. 3D–1D evaluation

To compare the quality of the two bovine structures, an analysis was made using the three-dimensional profile method of Eisenberg *et al.* (1992). This procedure is useful in assessing the correctness of structural models (Luthy *et al.*, 1992). The *z*-score calculated from this analysis gives an estimate of the overall compatibility of the proteins sequence to the structure, while plots averaged along the amino-acid sequence can point to discrete regions of suspect structure.

The *z*-score for the BOV model is 112, a value determined empirically to correspond to a correctly determined structure of this size (Luthy *et al.*, 1992). The score for the previous bovine structure (1rfb) is 60, a value indicative of a partially incorrect structure. The 'three-dimensional' plot, averaged using a ten-residue sliding window, is shown in Fig. 11. Note that the high-resolution structure does not drop below a value of 0.2, indicating the'correctness' of this structure. In contrast, much of the plot for 1rfb (47%) is below 0.2, confirming that regions of this model are problematic.

The errors in 1rfb apparently resulted from misinterpretation of the electron density and emphasize the possible effects of the poor data-to-parameter ratio inherent to low-resolution structures. In this refinement, 2744 reflections ($I > 2\sigma$, 49% completeness) were used to refine four variables (three positional and one thermal) for 2024 atoms (including 43 waters), an observation-to-parameter ratio of 1:3. This refinement also shows the limitation of using a model's *R* value as the ultimate measure of its correctness. Under conditions of such significant underdetermination, cautious refinement strategies are essential to prevent refining into a false low-energy minima.

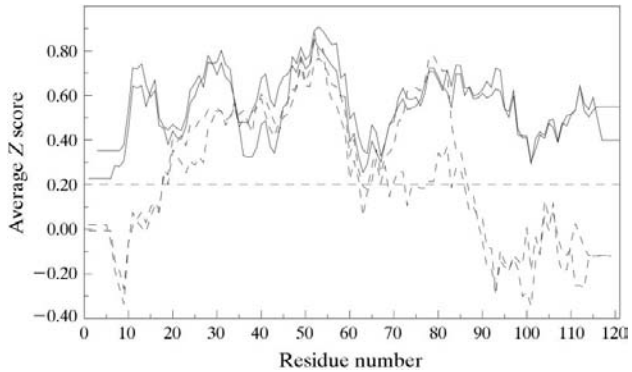

Figure 11
3D–1D plot calculated using a ten-residue sliding window. The horizontal line at 0.2 serves to delineate sequence compatible with its tertiary environment (>0.2) from that inconsistent with its tertiary environment (<0.2). Solid line, BOV; dashed line, 1rfb.

Table 4
R.m.s.d.s for the pairwise comparison of the four available IFN-γ structures (Å).

Numbers above the diagonal were generated using C^{α} atoms 3–119. Numbers below the diagonal represent only the helical C^{α} atoms (3–15, 30–36, 42–60, 67–81, 86–96, 103–119).

	BOV	1hig	1rfb	2rig
BOV	—	2.2	4.5	4.2
1hig	1.2	—	5.1	4.8
1rfb	4.3	4.5	—	3.0
2rig	3.8	4.0	2.5	—

This investigation was supported by NIH Research Service Award GM08388 (Biotechnology Training Grant). We wish to thank Dr Mark Zoller for providing the plasmid expressing bovine IFN-γ and Mark Ultsch for developing the crystallization conditions.

References

Bach, E. A., Aguet, M. & Schreiber, R. D. (1997). *Annu. Rev. Immunol.* **15**, 563–591.
Brünger, A. T. (1992*a*). *Nature (London)*, **355**, 472–475.
Brünger, A. T. (1992*b*). *X-PLOR. A System for X-ray Crystallography and NMR.* New Haven, Connecticut: Yale University Press.
Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Crick, F. H. C. & Magdoff, B. S. (1956). *Acta Cryst.* **9**, 901–908.
Ealick, S. E., Cook, W. J., Vijay-Kumar, S., Carson, M., Nagabhushan, T. L., Trotta, P. P. & Bugg, C. E. (1991). *Science*, **252**, 698–702.
Eisenberg, D., Bowie, J. U., Luthy, R. & Choe, S. (1992). *Faraday Discuss.* **93**, 25–34.
Engh, R. A. & Huber, R. (1992). *Acta Cryst.* A**47**, 392–400.
Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P. & Matthews, B. M. (1992). *Science*, **255**, 178–183.
Farrar, M. A. & Schreiber, R. D. (1993). *Annu. Rev. Immunol.*, **11**, 571–611.
Ferrin, T. E., Huang, C. C., Jarvis, L. E. & Langridge, R. (1988). *J. Mol. Graph.* **6**, 13–27.
Furey, W. & Swaminathan, S. (1998). *Methods Enzymol.* **277**, 590–620.
Grzesiek, S., Dobeli, H., Gentz, R., Garotta, G., Labhardt, A. M. & Bax, A. (1992). *Biochemistry*, **31**(35), 8180–8190.
Hatfull, G. F., Sanderson, M. R., Freemont, P. S., Raccuia, P. R., Grindley, N. D. & Steitz, T. A. (1989). *J. Mol. Biol.* **208**(4), 661–667.
Hemmi, S., Peghini, P., Metzler, M., Merlin, G., Dembic, Z. & Aguet, M. (1989). *Proc. Natl Acad. Sci. USA*, **86**(24), 9901–9905.
Jiang, J. S. & Brünger, A. T. (1994). *J. Mol. Biol.* **243**(1), 100–115.
Jones, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.
Jones, T. A. (1985). *Methods Enzymol.* **115**, 157–171.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–934.
Kunkel, T. A. (1985). *Proc. Natl Acad. Sci. USA*, **82**(2), 488–492.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Luthy, R., Bowie, J. U. & Eisenberg, D. (1992). *Nature (London)*, **356**, 83–85.
Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.
Martinez, C., de Geus, P., Stanssens, P., Lauwereys, M. & Cambillau, C. (1993). *Protein Eng.* **6**(2), 157–165.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Nagata, K., Izumi, T., Kitagawa, T. & Yoshida, N. (1987). *J. Interferon Res.* **7**(3), 313–320.
Nicholls, A., Sharp, K. A. & Honig, B. (1991). *Proteins*, **11**(4), 281–296.
O'Shea, E. K., Klemm, J., Kim, P. S. & Alber, T. (1991). *Science*, **254**, 539–544.
Pestka, S., Langer, J. A., Zoon, K. C. & Samuel, C. E. (1987). *Annu. Rev. Biochem.* **56**, 727–777.
Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. (1963). *J. Mol. Biol.* **7**, 95–99.
Rinderknecht, E., O'Connor, B. H. & Rodriguez, H. (1984). *J. Biol. Chem.* **259**(11), 6790–6797.
Samudzi, C. T., Burton, L. E. & Rubin, J. R. (1991). *J. Biol. Chem.* **266**(32), 21791–21797.
Samudzi, C. T. & Rubin, J. R. (1993). *Acta Cryst.* D**49**(6), 505–512.
Schevitz, R. W., Otwinowski, Z., Joachimiak, A., Lawson, C. L. & Sigler, P. B. (1985). *Nature (London)*, **317**, 782–786.
Sprang, S. R. & Bazan, J. F. (1993). *Curr. Opin. Struct. Biol.* **3**, 815–827.
Steigemann, W. (1974). PhD thesis, Technical University Munich, Germany.
Stewart, W. E. (1979). *The Interferons.* New York: Springer-Verlag.
Sun, D. P., Alber, T., Bell, J. A., Weaver, L. H. & Matthews, B. W. (1987). *Protein Eng.* **1**(2), 115–123.
Walter, M. R. &. Nagabhushan, T. L (1995). *Biochemistry*, **34**(38), 12118–12125.
Walter, M. R., Windsor, W. T., Nagabhushan, T. L. Lundell, D. J., Lunn, C. A., Zauochy, P. T. & Narula, S. K. (1995). *Nature (London)*, **376**, 230–235.
Zdanov, A., Schalk-Hihi, C., Gustchina, A., Tsang, M., Weatherbee, J. & Wlodawer, A. (1995). *Structure*, **3**(6), 591–601.