Structure of human chorionic gonadotropin at 2.6 Å resolution from MAD analysis of the selenomethionyl protein

Hao Wu[1], Joyce W Lustbader[2], Yee Liu[1], Robert E Canfield[2] and Wayne A Hendrickson[1,3]*

[1]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032, USA, [2]Department of Medicine, Columbia University, New York, NY 10032, USA and [3]The Howard Hughes Medical Institute, Columbia University, New York, NY 10032, USA

Background: Human chorionic gonadotropin (hCG) is a placental hormone that stimulates secretion of the pregnancy-sustaining steroid progesterone. It is a member of a family of glycoprotein hormones that are disulfide-rich heterodimers, with a common α-chain and distinctive β-chains specific to their particular G-protein linked receptors.

Results: We have produced recombinant hCG in mammalian cells as the selenomethionyl protein, and have determined its structure (after partial deglycosylation) at 2.6Å resolution from multiwavelength anomalous diffraction (MAD) measurements. Despite only limited sequence similarity (10% identity), the α- and β-subunits of hCG have similar tertiary folds. Each subunit has a cystine-knot motif at its core of extended hairpin loops. There is a very extensive subunit interface featuring two inter-chain β-sheets and a unique, disulfide-tethered 'arm' from the β-subunit which 'embraces' the α-subunit. The carboxy-terminal peptide of the β-subunit, which is rich in O-linked sugars, is disordered.

Conclusions: Structural and sequence comparisons indicate an evolutionary homology, albeit remote, between the glycoprotein hormone chains and other cystine-knot proteins, notably platelet-derived growth factor. Segments of the α- and β-chains that have been convincingly implicated in receptor binding by hCG are juxtaposed on one side of the molecule. A glycosylation site implicated in signal transduction but not in binding is also close to the presumed binding site suggesting a possible coupling between ligand binding and signaling. This study with selenomethionyl protein produced in mammalian cells extends the realm of MAD phasing.

Structure 15 June 1994, 2:545–558

Key words: anomalous scattering, cystine-knot, glycoprotein hormones, pregnancy, selenomethionine

Introduction

Human chorionic gonadotropin (hCG) is a glycoprotein hormone secreted by the placenta during the early weeks of pregnancy [1]. It stimulates the ovarian *corpus luteum* to produce progesterone until the placenta itself acquires the ability to produce this pregnancy-sustaining steroid. Excess hCG is excreted, and measurement of the level of hCG in urine is the basis of most pregnancy tests [2]. Because of the central role played by hCG in reproduction, agonists of hCG have the potential for treatment of infertility (or stimulation of fertility in endangered species) and antagonists are candidates for contraceptives.

The glycoprotein hormones are a family that includes the pituitary hormones lutropin (luteinizing hormone; LH), follitropin (follicle-stimulating hormone; FSH) and thyrotropin (thyroid-stimulating hormone; TSH), as well as hCG. Members of the family are all heterodimers formed by the non-covalent association ($K_d \approx 10^{-7}$ M under physiological conditions) [3] of an α-subunit, which is common to all the members of the family within a given species, with distinct but homologous β-subunits which confer hormone specificities [1]. The glycoprotein hormones exert their activities by binding to distinct cell surface receptors and activating adenylyl cyclase [1]. LH binds to the same receptor as hCG and produces the same cellular response, but its secretion is under different physiological regulation. The receptors for these hormones belong to the family of G-protein-coupled receptors that have seven transmembrane helices [4,5]; but unlike the other members of the family that respond to light or to small ligands, these hormone receptors have large extracellular domains that bind the glycoprotein ligands.

The α- and β-subunits of hCG contain 92 and 145 residues, respectively. Of relevance to this study, four of these are methionines. There are five disulfide bridges in the α-subunit and six in the β-subunit, and these are conserved across the family. Each subunit has two N-linked glycosylation sites and there are four O-linked glycosylation sites on the carboxy-terminal extension of the hCG β-subunit, a segment that does not exist in other glycoprotein hormones [1]. The carboxy-terminal peptide is not essential for *in vitro* hCG activity, and it has been suggested that this glycosylated extension may impart extra solubility and *in vivo* circulation life-

*Corresponding author.

 ISSN 0969-2126

time to hCG [6]. Carbohydrates account for 30–35% of the total mass of the hormone [7]. Various attempts have been made to construct models of hCG [8–11], but these prove to have been based in part on false homologies or incorrect disulfide assignments.

Here we report the structure of hCG as determined from the diffraction analysis of crystals of the selenomethionyl (Se-met) protein. We and others have previously described crystals of partially deglycosylated urinary hCG [12,13]. In order to incorporate selenomethionine for analysis by the method of multiwavelength anomalous diffraction (MAD) [14,15], we needed to express the recombinant protein; because of the heavy glycosylation and intricate disulfide structure, this was done in mammalian cells rather than in bacteria. The recombinant Se-met hCG, after partial deglycosylation, crystallized similarly to urinary hCG (Lustbader, J.W., *et al.*, & Canfield, R.E., unpublished data) and our MAD analysis of synchrotron data from a frozen crystal gave the structure which is presented here at 2.6 Å resolution. The structure of hCG has also been solved independently by another method (NW Isaacs and co-workers, personal communication).

Results and discussion

Structure determination

Se-met hCG was expressed in Chinese hamster ovary (CHO) cells, purified and partially deglycosylated with hydrogen fluoride (HF), as described elsewhere (Lustbader, J.W., *et al.*, & Canfield, R.E., unpublished data). Amino acid analyses showed that 92% of the methionines were replaced by selenomethionines. The HF treatment left roughly 50% of the carbohydrate intact. We grew crystals of HF-treated Se-met hCG in an anaerobic environment by micro-seeding with native crystals, under a condition similar to that used for native hCG [12,13] except for the addition of the reductant dithiothreitol (DTT) and the sealing of each crystallization well in an anaerobic glove box. Both an anearobic environment and reductant were required for useful crystals, perhaps due to the oxidation of selenomethionines during the two month period of crystal growth. The sensitivity of crystal growth to oxygen implies that oxidation of the selenomethionine residues may destroy the crystal lattice. The crystal structure shows that three of the four methionines in the hCG structure are at the surface and in lattice contacts. This is in keeping with chromatographic properties which show Se-met hCG to be more hydrophobic than native hCG (Lustbader, J.W., *et al.*, & Canfield, R.E., unpublished data).

The diffraction data were measured from four wavelengths near and at the selenium K-shell edge. Phase determinations from these data were performed by algebraic [16] and probabilistic [17] procedures developed in other applications of MAD phasing [14]. Selenium sites were located by Patterson search methods [18] and in difference Fourier maps, and after refinement these were used to produce MAD phases. A Fourier synthesis computed from merged MAD phases in space group $P6_522$ had a clear solvent boundary, thus determining the hand. Solvent flattening was subsequently applied to the electron density map [19]. A retrospective analysis based on phases calculated from the refined model showed average phase discrepancies of 53.1° against the initial MAD phases and 50.5° against phases after solvent flattening.

Using both the direct MAD-phased and solvent-flattened electron density maps, we were able to make an unambiguous tracing of the molecule (Fig. 1). The knowledge of the methionine sites aided and provided additional confidence in the tracing. After several iterations of restrained refinement and manual rebuilding, our current model has 1554 non-hydrogen atoms including residues 5–89 of the α-subunit, residues 2–111 of the β-subunit, two N-linked pyranose units attached to the glycosylation sites at Asn52 and Asn78 of the α-subunit, and 64 solvent molecules. This model gives a 19.9% crystallographic R-value for the 8876 reflections between 5 Å and 2.6 Å spacings having $|F| > 3\sigma_{|F|}$, and it has root mean square (rms) deviations from ideality in bond lengths of 0.017 Å and in bond angles of 2.3°.

Overall structure and tertiary folding

The ordered portion of the hCG molecule seen in the crystal structure has an elongated shape with dimensions of roughly 75 Å × 35 Å × 30 Å. The subunits are also elongated, ~60 Å × 25 Å × 15 Å for the α-subunit and ~65 Å × 25 Å × 20 Å for the β-subunit, and they are highly intertwined. Each chain is folded into three looped-out pairs of extended strands that are characteristically bridged by disulfide bonds (Fig. 2). Unexpectedly, as sequence similarity had not been detected, the α- and β-subunits of hCG share similar polypeptide folds. Each is organized about a common set of disulfide bridges (10–60, 28–82 and 34–84 in the α-subunit; 9–57, 34–88 and 38–90 in the β-subunit) from which the three hairpin-like loops emanate. This core structure is a cystine-knot motif [20,21] whereby one disulfide bridge between two of the strands (β1 and β3) passes through an eight-residue circle formed by a pair of bridges joining two other strands (β2 and β4). Other disulfide bonds tie the termini of the α-subunit to the core (7–31 and 59–87) and bridge more peripheral elements of the β-subunit (23–72, 26–110 and 93–100). Some disulfides seem to have been rearranged in earlier experiments [22,23]. The overall topology of the hCG subunits in association with one another is shown in Fig. 3.

Except for loop $\mathcal{L}2$ of the β-chain, each loop emanating from the cystine-knots is a β-hairpin with an extended ladder of hydrogen bonds between the strands (Fig. 4). Loops $\mathcal{L}1$ and $\mathcal{L}3$ from one end of each knot are highly twisted, whereas the $\mathcal{L}2$ loops from the other end are rather flat. Loop $\mathcal{L}2$ of the β-chain is an antiparallel extension despite the fact that it lacks hydrogen bonds between the β3 and β2′ strands. The mode of

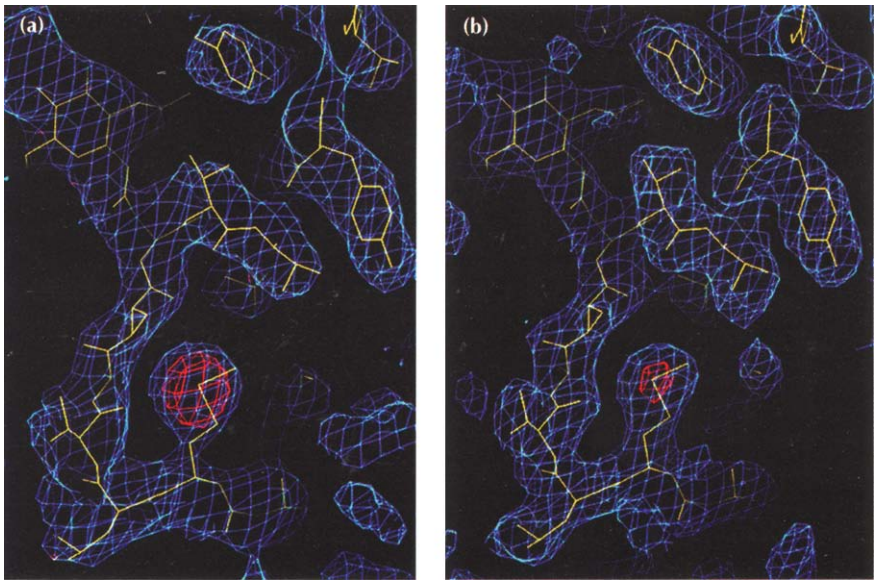

Fig. 1. (a) A portion of the MAD-phased map and **(b)** the current $2|F_o|-|F_c|$ map superimposed with the model refined at 2.6 Å resolution. Electron density is contoured at 1.0σ (blue) and 5.5σ (red). Two prominent features are the *N*-acetylglucosamine residue attached to Asn52, and Met47 in the hCG α-subunit.

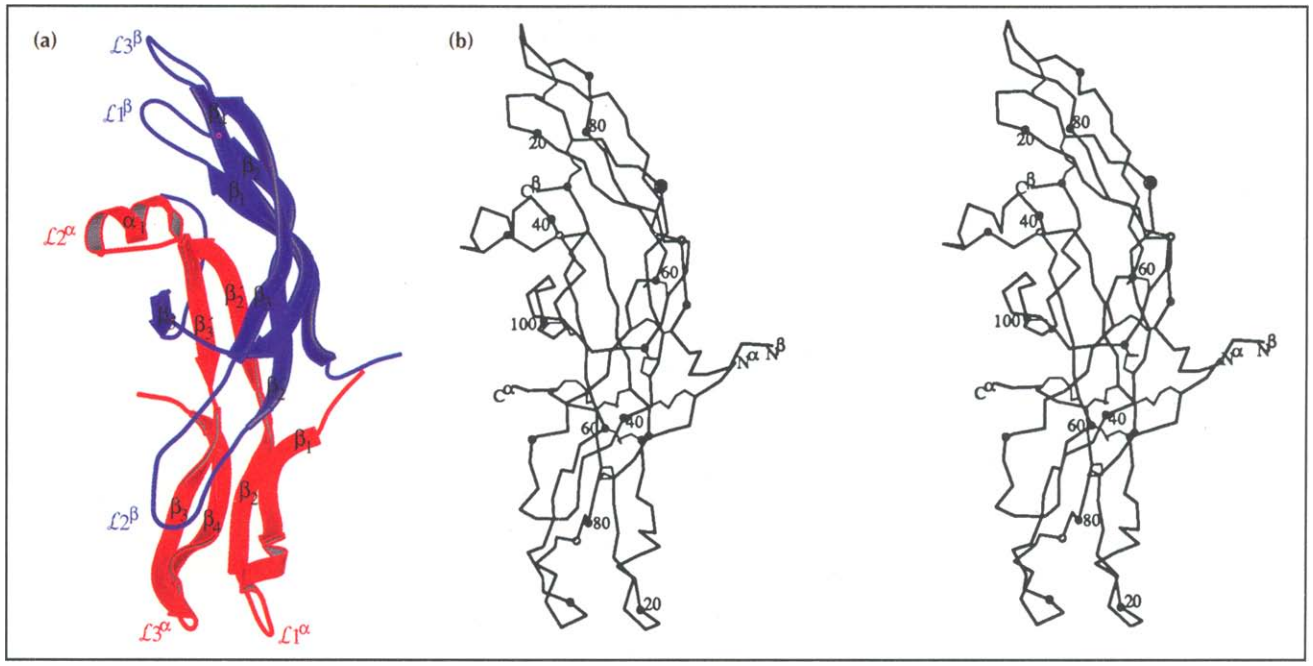

Fig. 2. (a) A ribbon diagram of hCG. The α-subunit is shown in red and the β-subunit in blue. The secondary structural elements are labeled. **(b)** A stereo C_α trace of hCG in the same orientation as (a). Every tenth residue is marked by a filled circle on the C_α position and glycosylation sites (Asn52 and Asn78 of the α-subunit, Asn13 and Asn30 of the β-subunit) are denoted by open circles. The amino and carboxyl termini and every twentieth residue are labeled. (Figure produced by MOLSCRIPT [57].)

twisting in loops $\mathcal{L}1$ and $\mathcal{L}3$ is similar; a quasi-dyad axis of symmetry passes roughly through the two disulfide bridges at the double looped-out end of the knot, relating the core portions of $\mathcal{L}1$ and $\mathcal{L}3$ of each chain. The least-squares superposition [24] of 10 C_α and 2 S_δ pairs gives rms deviations of 1.14 Å and 1.02 Å for the α- and β-chains, respectively, for screw transformations of 180° and 0.69 Å in the α-subunit and 165.8° and 0.42 Å in the β-subunit. These intra-subunit superpositions (Fig. 5) are interrupted by the β-bulges in strands β4, at residues 80 in the α-subunit and 86 in the β-subunit. These β-bulges have also been noted in other cystine-knot proteins [21].

The lengths of the loops and the conformations at their tips are variable. The $\mathcal{L}1$ and $\mathcal{L}3$ loops of the β-chain are several residues longer than their α-chain counterparts. While loops $\mathcal{L}1$ and $\mathcal{L}3$ of the α- and β-subunits terminate in classical β-turns, the remaining loop tips are more open. Although hCG is primarily a β-structure protein, in accord with circular dichroism measurements [25] and secondary structure predictions [26], there are three helical segments, two in the α-subunit and one in the β-subunit. In the α-subunit, one is a single turn (3_{10}) after strand β1, and the other is a two-turn α-helix at the tip of $\mathcal{L}2$ perpendicular to the strands. In the β-subunit, there is a short α-helical-like

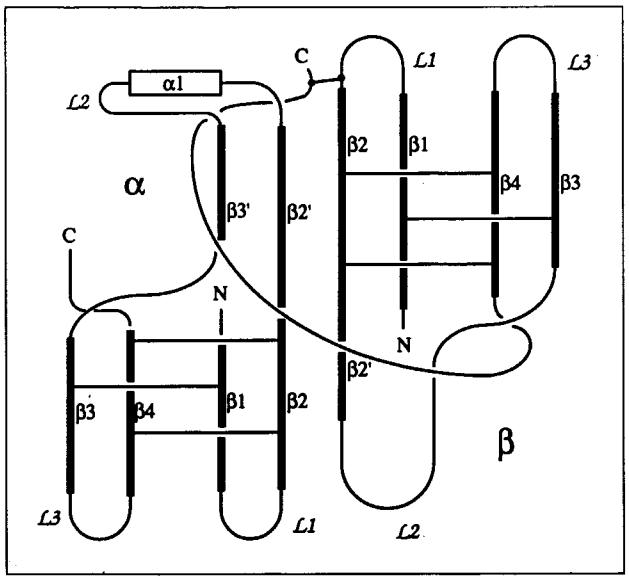

Fig. 3. A schematic drawing of the hCG dimer topology.

segment between β4 and β5. The amino termini of both chains are perpendicular to the strand axes of the cystine-knots, and are disordered before residue 5 in the α-chain and at residue 1 of the β-chain. Both chains turn sharply away from the core strands immediately and into carboxy-terminal extensions after the last cysteine residue of the cystine-knot. For the α-subunit, this extension has eight residues, the last three of which are disordered. For the β-subunit, the ordered part of the extension is 21 residues long. It forms an 'arm' that 'embraces' L2 of the α-chain and which is tethered to the tip of L1 of the β-chain by the disulfide bridge 26–110. The carboxyl terminus of the β-subunit projects into a solvent expanse and there is no defined electron density for the last 34 residues.

Dimeric association

The subunits of hCG are intimately associated in the heterodimer, fitting somewhat like clasped hands. A total of 3860 Å^2 of the surface area of the non-glycosylated dimer (25% of the sum of the surface areas for the isolated subunits) is inaccessible to a solvent probe of 1.4 Å radius. Inter-chain hydrogen bonding is at the heart of this interface. Strand β2 of each subunit forms main chain hydrogen bonds with the β2′ strand of the other subunit, thereby extending intra-chain β-hairpins into antiparallel inter-chain β-sheets. In addition, a short strand (β5) in the carboxy-terminal extension of the β-subunit is hydrogen bonded in parallel with β3′ of the α-subunit; a stretch at the start of β4 from the α-subunit is antiparallel to a segment near the end of β2′ of the β-subunit, and there are two hydrogen bonds between the amino-terminal extensions (Fig. 4). Of the two inter-chain β-sheets, one is five-stranded [β1β2(of β):β2′β3′(of α):β5(of β)] and the other is three-stranded [β1β2(of α):β2′(of β)]. Owing to the

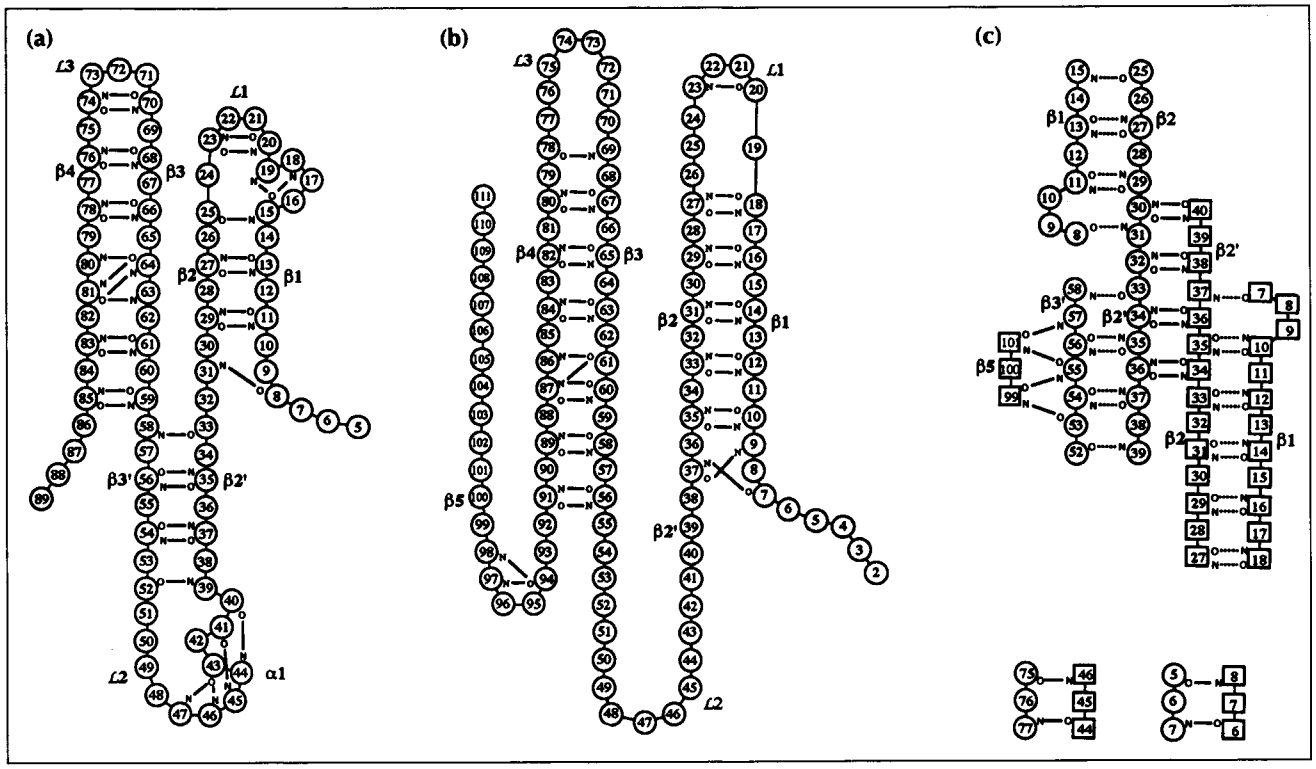

Fig. 4. Hydrogen-bonding diagram of **(a)** the α-subunit of hCG, **(b)** β-subunit and **(c)** between the α- and β-subunits, with residues of the β-subunit shown as squares in (c). Hydrogen bonds have been included for amide to carbonyl interactions having O–N distances less than 3.5 Å and carbonyl to amide angles greater than 120°.

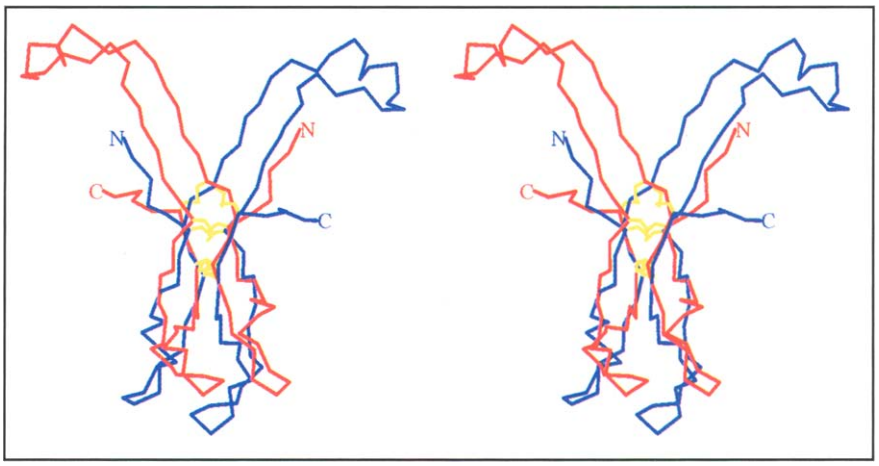

Fig. 5. Inter-chain superposition of the cystine knot of the α-subunit of hCG. C_α positions 9–13 and 26–30 were aligned with 59–63 and 80–84, respectively. In addition, S_δ atoms of residues 10 and 28 were aligned with residues 60 and 82, respectively. The quasi-dyad-related copies are distinguished in red and blue with disulfide bonds in yellow. (Figure produced by MOLSCRIPT [57].)

peculiar twist of cystine-knots, these two sheets are essentially perpendicular to one another.

The reciprocal character of the dimer interface enforces a quasi-dyad symmetry despite the negligible sequence similarity between the α- and β-subunits (Fig. 6). The 39 C_α positions from the α-subunit that could be superimposed within 2 Å of their counterparts in the β-subunit were related, with 0.90 Å rms deviation, by a 178° rotation and −0.46 Å translation. A marked asymmetry in the heterodimer occurs at the side of the molecule opposite from the amino termini. Here the carboxy-terminal extension of the β-subunit from the cystine-knot core wraps around the helix-tipped *L*2 loop of the α-subunit (Fig. 7).

It is an open question whether heterodimer assembly involves threading of the *L*2 loop of the α-subunit through the disulfide-tethered carboxy-terminal extension of the β-subunit or, alternatively, whether an open arm from the β-subunit wraps around the α-subunit before being tied down to the *L*1 loop of the β-subunit. The trapping of *in vivo* folding intermediates of hCG suggested that the 93–100 disulfide bridge of the β-subunit forms roughly concurrently with heterodimer assembly, but that the 26–110 bond forms later [27]. This observation is consistent with an open embrace followed by closure. However, the *in vitro* reconstitution of hCG from the α- and β-subunits does not require the presence of added reducing agents [28,29], which indicates that threading through a tethered arm is possible. On the other hand, addition of protein disulfide isomerase during reconstitution accelerates the *in vitro* folding and assembly of hCG [30], supporting the suggestion that natural dimerization may indeed involve opening of the 26–110 disulfide bond. Moreover, when the cysteine residues of the 93–100 or the 26–110 disulfide bridge of the β-subunit were mutated to alanines, no heterodimer assembly was detected although disulfide bonds that form earlier along the folding pathway were unaffected [31].

Solvent exposure and folding implications

The hCG structure is remarkably open. Surface to volume ratios for the isolated chains are exceptionally high at 0.62 $Å^{-1}$ and 0.61 $Å^{-1}$ for the α- and β-chains, respectively; and, despite the burial of 25 % of this sur-

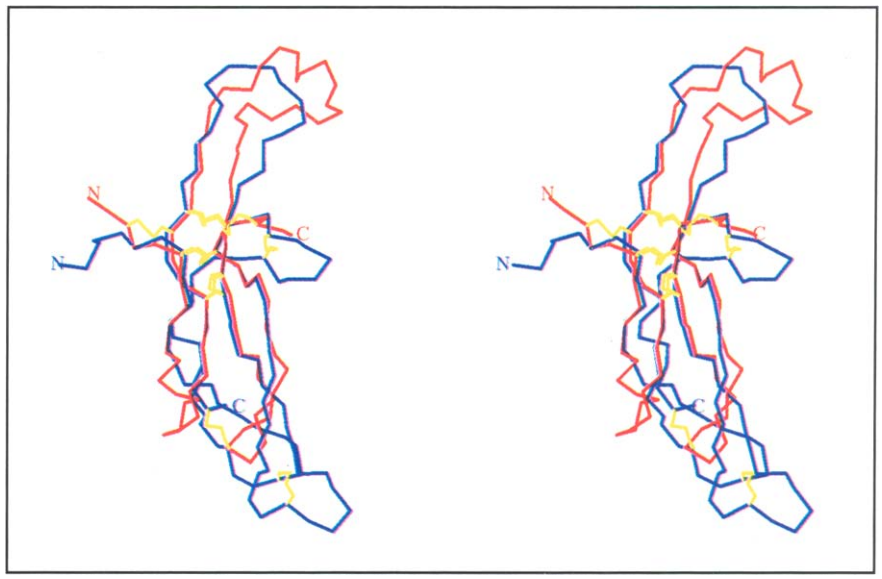

Fig. 6. Superposition of hCG α- and β-subunits by superimposing 39 C_α positions (residues 7–15, 25–36, 58–65 and 77–86 of the α-chain) with their counterparts in the β-chain. The C_α traces for the α-subunit are shown in red and those for the β-subunit are in blue. Disulphide bridges are in yellow. (Figure produced by MOLSCRIPT [57].)

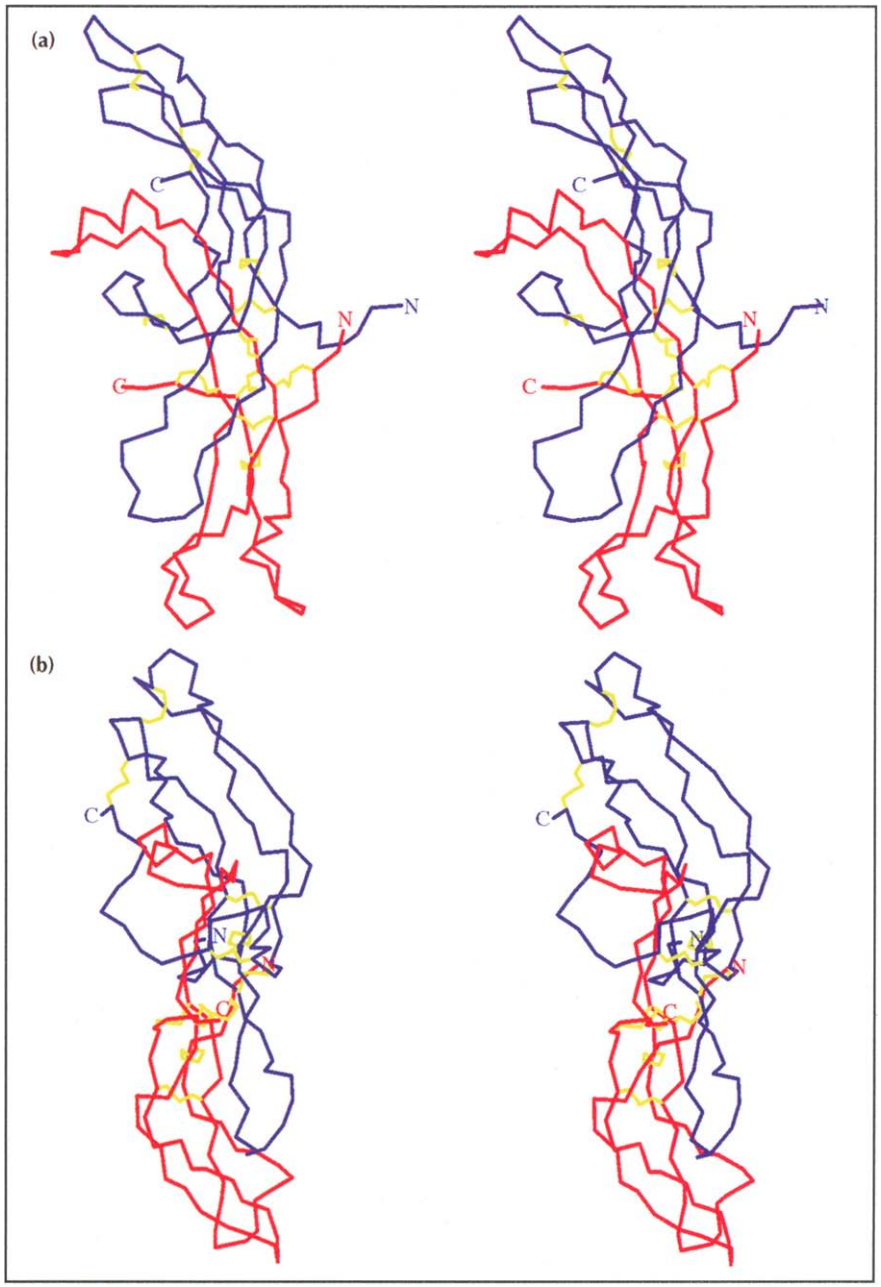

Fig. 7. Two views of the hCG heterodimer oriented by the quasi-dyad axis of symmetry: **(a)** with the dyad horizontal, **(b)** viewed along the dyad axis. The C_α traces for the α-subunit are shown in red and those for the β-subunit are in blue. Disulfide bridges are in yellow. (Figure produced by MOLSCRIPT [57].)

face in the dimer interface, even for the dimer this ratio remains 2.8 times greater than for a sphere of equivalent volume. For the most part, hCG is a one-layer structure with the consequence that neither the individual subunits, as separated out from the hCG model, nor the intact heterodimer has much of a hydrophobic core. Only four residues from the α-subunit and seven from the β-subunit are more than 90% buried in the isolated molecules and only two of these, a leucine and an isoleucine, have large hydrophobic side chains. While many more residues are buried in the heterodimer (22 in the α-subunit and 18 in the β-subunit are less than 10% exposed), only six of these (two leucines, two isoleucines, and two tyrosines) are large hydrophobic residues. In fact, 31 of these buried residues have three or fewer side-chain atoms. Tyrosine 37 (of α), Ile33 (of β) and Tyr59 (of β) do, however, constitute a small hydrophobic core in the dimer interface.

Despite the apparent lack of hydrophobic cores in the separate subunits, both isolated α- and β-subunits do retain significant secondary structure [25], and the excess uncombined β-chains that are secreted from choriocarcinoma cells seem to be folded with appropriate disulfide bonds [27]. Given the extended nature of loops from the cystine knots of these chains, the isolated subunits might be quite flexible. This may explain the relatively modest dimer affinity, which is at first sight surprising given the extraordinarily large area of interface. Surface energy calculations [32] would suggest an association energy of $\sim 48\,\text{kcal mol}^{-1}$, whereas

the observed subunit affinity actually corresponds to about ~ 9 kcal mol^{-1}. Entropic losses on assembly of the flexible subunits into the hCG heterodimer could be substantial.

Comparisons with other cystine-knot proteins

The cystine-knot motif that is at the heart of both hCG subunits was first observed in nerve growth factor (NGF) [33], and has since also been found in transforming growth factor-β (TGF-β2) [34,35] and platelet-derived growth factor (PDGF-BB) [36]. As noted previously, while the disulfide knots from these structures superimpose closely, the loop extensions vary widely [20,21]. We have compared the hCG subunits with the other known cystine-knot units by superimposing a common core of 21 C_α positions and looking for the number of aligned C_α positions brought to within 2 Å by the transformation (Table 1). While there is great divergence outside this core, the hCG subunits are strikingly similar to PDGF-BB within this core, and to a lesser extent they are also similar to TGF-β2. The superposition of NGF with any of the other cystine-knot proteins results in deviations two to three times greater than among the others. This correlates with an insertion between cysteine residues in the β4 strand of NGF (and related neurotrophic factors), whereas all of the others have the sequence Cys-X-Gly-X-Cys with a positive φ torsion angle at the glycine position.

There is very little amino acid sequence similarity among the cystine-knot family members apart from the defining cysteine residues. This is in keeping with the structural variability outside the knot region and the lack of a well-defined hydrophobic core. We have aligned the sequences by using the cysteine residues of the knot as fixed points of reference while permitting a single gap at each loop (Fig. 8, Table 1). Despite the excellent superposition of core residues in the α- and β-subunits, there is only 10 % identity between the two sequences. This is at the low level found between NGF and the others. On the other hand, the comparison between the α-subunit and PDGF-BB (22 % identity) is at a level normally indicative of evolutionary homology and that between the α-subunit and TGF-β2 (19 %) may be significant.

All cystine-knot proteins that have been identified are dimers, and except for hCG, they are symmetric homodimers. All have the long axes of their subunits in parallel (or antiparallel) and have relatively large areas of buried surface. (We calculate that the buried surface areas are ~ 1160 Å^2 per subunit for NGF, ~ 1330 Å^2 for TGF-β2 and ~ 1120 Å^2 for PDGF-BB as compared with ~ 1930 Å^2 for hCG.) Amazingly, however, the modes of association are completely different in each case. Indeed, the surfaces of subunit interaction have almost no overlap (Fig. 9). This variability is reflected in the low sequence similarity.

The similarity among β-subunits of glycoprotein hormones is pronounced (Table 1) and well known [3].

Table 1. Comparison of cystine-knot protein structures and sequences.

	NGF	TGF-β2	PDGF-BB	hCG α	hCG β
NGF		1.81 Å (17)	1.97 Å (15)	1.85 Å (16)	1.77 Å (16)
TGF-β2	8 %		0.91 Å (30)	0.81 Å (29)	0.75 Å (30)
PDGF-BB	14 %	14 %		0.58 Å (37)	0.75 Å (39)
hCG α	10 %	19 %	22 %		0.59 Å (37)
hCG β	10 %	16 %	13 %	10 %	
hLH β	10 %	15 %	13 %	11 %	80 %
hFSH β	11 %	15 %	14 %	19 %	34 %
hTSH β	10 %	18 %	17 %	18 %	37 %

Structural comparisons and sequence comparisons between nerve growth factor (NGF), transforming growth factor β (TGF-β2), platelet-derived growth factor (PDGF-BB) and the α- and β-chains of human chorionic gonadotropin (hCG α and hCG β) are given in the upper and lower triangles, respectively. Additional sequence comparisons with the β-chains of luteinizing hormone (LH), follicle stimulating hormone (FSH) and thyroid stimulating hormone (TSH) are shown below. For the structural comparison, a core of C_α positions associated with the cystine-knot (residues 9–14, 26–38, 32, 58–62 and 80–85 for hCG α and their counterparts) were superimposed by the TOSS procedure [24]. The root mean square (rms) deviations from the superpositioning of these 21 atoms are reported together with the number of C_α positions brought within 2Å of one another by the superposition (in parentheses). Coordinates sets of TGF-β2 (entry name 1TGF) and PDGF-BB (entry name 1PDG) from the Protein Data Bank were used and coordinates of NGF (entry name 1BET, will be available) were obtained from NQ McDonald. The sequences were aligned to bring the cystine-knot residues into correspondence and a single gap was introduced where needed at each loop and also between the cysteine residues on strand β2 where NGF differs from all of the others. The percentage of identical residues is given as the ratio of paired identities to the number of paired residues plus one equivalent for each gap and for each terminal extension.

Sequence similarities between the α- and β-chains had gone undetected, but there does appear to be a significant similarity between the common α-chain and the β-chains of hFSH and hTSH when aligned by the structure (Table 1). As a common α-chain associates with four distinctive β-chains to produce the different hormones, one might expect greater conservation at the interface if they share a common mode of association. Residues 33–46 and 54–59 of the β-subunit of hCG interact closely with the α-subunit, and these segments are appreciably more conserved than for the polypeptide chain overall (90 %, 45 % and 55 % identity for hCG *versus* hLH, hFSH and hTSH, respectively). Moreover, large hydrophobic side chains that are predominantly buried in the hCG interface are conserved at β-chain positions Leu16, Ile33, Tyr59 and Tyr82.

Receptor interactions

Intact hCG exerts its effect through binding to the LH/CG receptor and activating adenylyl cyclase to produce cAMP and progesterone (in appropriate cells). The affinity of hCG for its receptor is at a sub-nanomolar level [37]. Many efforts have been made to determine sites on hCG that are responsible for binding. These include the testing of chimeric or mutant hormones for impairment of activity, chemical modifica-

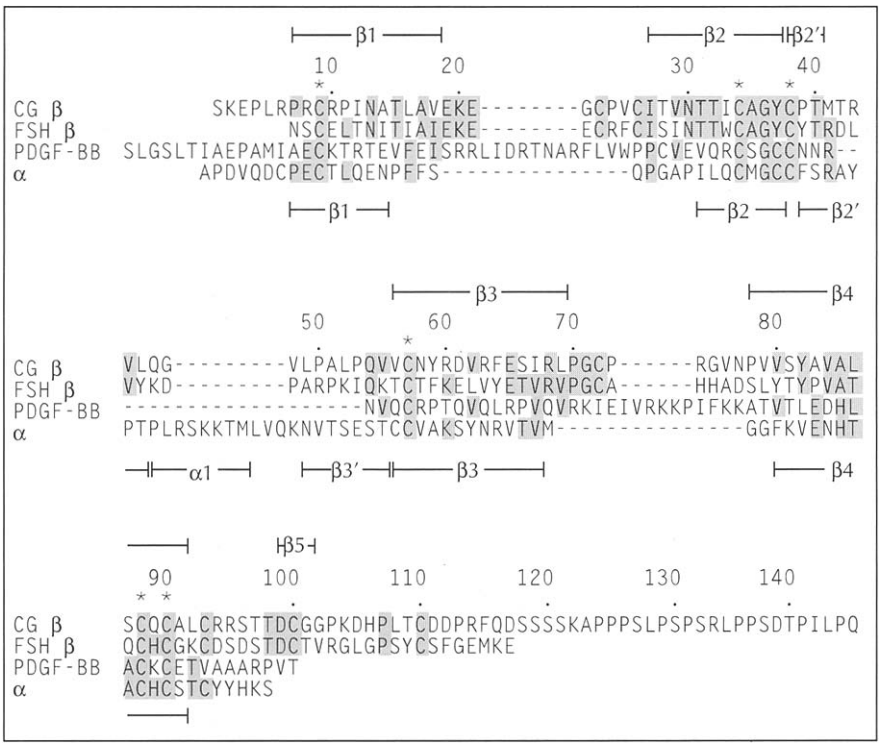

Fig. 8. Representative cystine-knot protein sequences. The sequences of the β-subunits of chorionic gonadotropin (CG β), follicle-stimulating hormone (FSH β), platelet-derived growth factor B (PDGF-BB) and the α-subunit of hCG (α) are aligned as described for Table 1. Residue numbers are marked for the hCG β-subunit. Conserved cysteine residues in the cystine knots are marked by asterisks. Secondary structural elements of hCG α- and β-subunits are marked below and above the alignments, respectively. Residues that are in common between any two sequences are shaded.

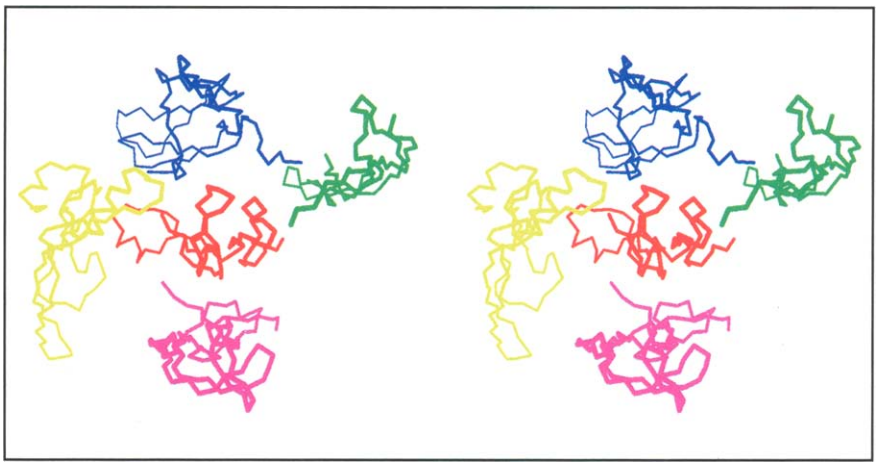

Fig. 9. Stereoview illustrating the different dimer interfaces used by hCG, NGF, TGF-β2 and PDGF-BB. One protomer from each of the dimers of NGF, TGF-β2 and PDGF-BB was superimposed with the hCG α-subunit by aligning the C_α positions of the six conserved cysteine residues in the cystine knots. The protomers not used in these superpositions were then transformed according to these superpositions and are shown along with the hCG heterodimer. The associated subunits have been translated out from the central subunit for clarity. The hCG α-subunit (red) is shown surrounded by the β-subunits of hCG (blue), NGF (magenta), TGF-β2 (yellow) and PDGF-BB (green). (Figure produced by MOLSCRIPT [57].)

tion, competition with peptides from hCG and epitope mapping of blocking antibodies. While the molecule has not been exhaustively examined in reported experiments and the segments tested were sometimes large, thereby defining the sites with limited resolution, a picture consistent with the crystal structure is emerging.

Regions from both subunits of hCG have been implicated in receptor binding. Removal of residues 88–92 of the α-subunit by carboxypeptidase treatment or site-directed mutagenesis greatly reduced the affinity of hCG for its receptor [38]. On the β-subunit, a region lying between residues 94 and 114 has been identified through the engineering of chimeric hormones. Transfer of this portion of hCG to hFSH resulted in a hybrid molecule endowed with binding affinity for the LH/CG receptor [39], and the substitution of this portion of hCG with the homologous region in hFSH abolished binding of hCG to the LH/CG receptor [40]. These portions of the α- (residues 88–92) and β-subunits (residues 94–114) are close together on the surface of hCG (Fig. 10a). Studies on other segments have given conflicting results. Although residues 38–57 of the β-subunit had been identified by antibody blocking [8,9] and stimulation by a synthetic peptide [41] as being involved in receptor interactions, substitution of this segment from hFSH in hCG had no effect on binding affinity [39].

Segments of hCG in the vicinity of the carboxy-terminal regions of the α- and β-subunits include prominences (Fig. 10b) that may also be candidates for an extended receptor-binding site. These include helix α1 on the *L*2 loop of the α-subunit. There is some indication of involvement of this region in receptor binding from the inhibition of hormone binding by peptides corresponding to residues 25–45 of the α-subunit [42,43]. The structural results also suggest an electro-

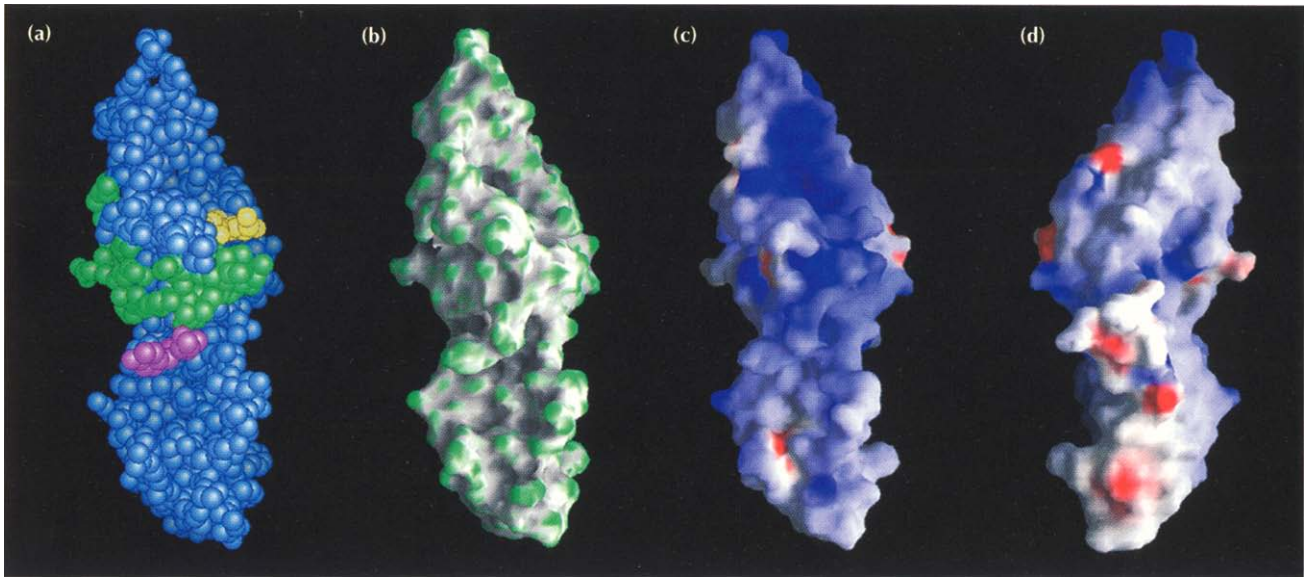

Fig. 10. (a) Space-filling representation of the hCG heterodimer showing the positions of residues involved in receptor binding and signal transduction. Residues in the α-subunit implicated in receptor binding are shown in magenta with those in the β-subunit in green. The glycosylation site, Asn52, of the α-subunit and its carbohydrate are shown in yellow. The glycosylation has been shown to be essential in receptor signal transduction. **(b)** Surface curvature of hCG. The most convex parts of the molecular surface are coded green while the most concave and planar are coded gray and white, respectively. **(c), (d)** Two views of the surface electrostatic potential of hCG [with (d) rotated by 180° about the vertical axis relative to (c)]. The surface is colored according to the local electrostatic potential, ranging from dark blue (most positive region) to deep red (most negative). The views in (a), (b) and (c) are the same (60° rotation about the vertical axis compared with the orientation in Fig. 2). These figures were drawn with GRASP [56].

static contribution to hCG activity. The electrostatic potential of hCG is markedly positive on the side implicated in receptor binding (Fig. 10c), in striking contrast with the opposite side (Fig. 10d). Mutagenesis has shown that at least one positively charged residue should be present between residues 93–100 for tight binding to the LH/CG receptor. Loss of positive charge reduced binding affinity and the introduction of negatively charged residues gave the lowest affinity for the LH/CG receptor, accounting, at least in part, for the low affinity of native hFSH for the LH/CG receptor [40].

The extracellular domain of the hCG receptor is rich in acidic residues and could thus be complementary to the positive face of hCG. This portion of the receptor contains 14 imperfect leucine-rich repeats (LRRs) [4,5]. The crystal structure of ribonuclease inhibitor [44] which has a similar number of LRRs, has a horseshoe shape (~30 Å inner diameter) with the ribonuclease-binding site on its inner surface. Such a structure could accommodate the shorter dimensions of hCG (see above).

Role of glycosylation

The abundance of carbohydrate moieties on the glycoprotein hormones has prompted the study of their function. The O-linked sites, which are disordered in the crystal, are expendable for hCG function [6]. Several observations indicate that N-linked glycosylations of hCG, although not affecting receptor binding affinity, play an essential role in signal transduction [45]. Deglycosylated hCG forms a tight complex with the LH/CG receptor, but such a complex fails to elicit effective second messenger production. Site-directed mutagenesis showed that glycosylation at Asn52 of the α-subunit is indispensable for signaling [46,47], whereas other N-linked sugars have a lesser role. Whether Asn52 has a direct effect on signal transduction through interaction with the LH/CG receptor or whether it affects the conformation of the hormone is not clear. From the structure, both options are possible since Asn52 is located close to the receptor-binding site as presently identified (Fig. 10a) and it is near a depression on the hCG surface (Fig. 10b). A site with sequence similarity to soybean lectin has been found in the LH/CG receptor [4] which might be a candidate for direct interaction with the Asn52-linked glycosylation in hCG. On the other hand, it has been shown that polyclonal antibodies against the hCG β-subunit could restore the signaling activity of deglycosylated hCG, suggesting a conformational role for the carbohydrates [48]. The fact that desialylated hCG, which retains most of the hormonal activity, could be crystallized in the same way as the HF-treated hCG suggests that the HF-treated hCG in the crystals also has a conformation that is competent for signal transduction.

Structural conclusions

Although the sequences of the α- and β-chains of hCG are only remotely related, their cystine-knot cores are very similar in tertiary structure. They are also similar to other cystine-knot proteins — NGF, TGF and PDGF — in this core, but are highly divergent beyond the core. Moreover, although all these proteins are dimers, their modes of dimer association are completely different.

The hCG heterodimer has an extraordinarily large area of interface with only a small hydrophobic core. The interface features quasi-dyad-related inter-chain β-sheets and an embracing arm of the β-subunit that surrounds a loop of the α-subunit and is tethered by a disulfide bridge to a loop on the β-subunit. The isolated subunits have no hydrophobic interiors, but they are known to fold with defined secondary structure and appropriate disulfide bridges in isolation from one another. Segments of the α- and β-subunits that have been implicated in receptor binding map close to one another on the hCG surface, as does a site of glycosylation implicated in signal transduction.

Biological implications

Glycoprotein hormones are central to the regulation of reproductive physiology. They stimulate testicular and ovarian function by inducing gametogenesis and steroid hormone synthesis in the gonads. The related pituitary hormone, thyroid-stimulating hormone, regulates a variety of processes in the thyroid that result in the secretion of steroid hormones. The production of human chorionic gonadotrophin (hCG), which is limited to the placenta, results in the secretion of progesterone to sustain the early stages of pregnancy. Both hCG and luteinizing hormone (LH) bind to the same receptor, but in males and non-pregnant females, the LH/CG receptor is usually only exposed to LH secreted by the pituitary.

Because of the controlling role played by hCG in pregnancy, there is interest in antagonists to serve as contraceptives and in agonists that might stimulate fertility. Knowing the structure of hCG should assist in the design of such compounds. The interaction of hCG with its receptor is rather complex in that binding and signaling functions can be segregated. This probably reflects the unusual receptor structure which comprises an extracellular binding domain integrated with a seven-transmembrane-helix domain that couples to a G-protein, which ultimately stimulates adenylyl cyclase. Residues implicated in receptor binding are juxtaposed on the hCG molecular surface, and they are surrounded by other prominences and by markedly positive electrostatic potential. These features may also be involved in receptor binding. An N-linked glycosylation site that seems to be involved in signal transduction but not in receptor binding is nearby. Other seven-transmembrane-helix receptors bind small ligands within the helix bundle, which suggests a direct conformational transduction of the signal. It is not yet possible to envisage how signals are transduced through the LH/CG receptor, but the hCG structure and the receptor sequence limit the possibilities.

The structure of the hCG heterodimer with its extended and highly-exposed subunits presents an interesting problem in protein folding. While lacking hydrophobic cores, these subunits do nevertheless fold. *In vivo* assembly into the mature heterodimer appears to involve the formation of a disulfide bridge tethering the arm of the β-chain which embraces α-chain. The structure provides an explicit model for the further exploration of folding and assembly mechanisms.

Our analysis of the hCG structure has exploited the selenomethionyl protein produced in Chinese hamster ovary (CHO) cells [16]. Thus, incidentally, these studies demonstrate, perhaps surprisingly, that mammalian cells can be highly tolerant of the systematic replacement of methionine by selenomethionine. We, and others, have shown previously that bacteria can survive indefinitely in a selenomethionine medium [15], leading to complete substitution in all proteins. While the level of incorporation achieved in CHO cells (92% here and 87% for CD4, [15]) has not been complete, clearly, even in a more complex eukaryotic system, most if not all proteins function satisfactorily as their selenomethionyl analogs.

Materials and methods

Production of selenomethionyl hCG

The expression, purification and deglycosylation of Se-met hCG will be described elsewhere (Lustbader J.W.,*et al.*, & Canfield R.E., unpublished data) and are merely summarized below. CHO cells that had been transfected with the DNA corresponding to the hCG α- and β-subunits were grown for three days in a serum-free selenomethionine medium. Secreted Se-met hCG was purified from the medium by immuno-affinity chromatography and partially deglycosylated with HF treatment and further purified by gel filtration.

Crystal growth and freezing

Crystals of the HF-treated Se-met hCG were grown in an anaerobic environment by micro-seeding with native crystals [49], under conditions similar to those used for native hCG [12,13] except for the addition of 10 mM DTT (Lustbader, J.W., *et al.*, & Canfield R.E., unpublished data). Lyophilized Se-met hCG was dissolved in buffer (30 mM sodium acetate at pH 4.4) to a concentration of 30–40 mg ml^{-1}. A seed stock was made by transferring a native hCG crystal to a 40% saturated ammonium sulfate (SAS) solution (containing 10 mM DTT) in the acetate buffer and crushing it with a needle. The equilibrating reservoir contained 1 ml of 29–31% SAS and 10 mM DTT in the same acetate buffer. Each hanging droplet contained 1.5 μl of serially-diluted seed solution and 1.5 μl protein solution. The oxygen sealed

inside each crystallization well was exchanged to a mixture of 95% nitrogen and 5% hydrogen using an anaerobic glove box. It is essential to exchange the air for anaerobic gases with DTT added to react with any remaining oxygen in the solutions. Experiments lacking either an anaerobic environment or DTT gave only small crystals and experiments lacking both gave no crystals at all.

The oxidizing requirements for disulfide bridges and the reducing environment needed to maintain selenomethionine residues seem contradictory, and initially we were reluctant to use any reducing agent during crystallization. Surprisingly, trial experiments with DTT added gave good crystals that were isomorphous with those from native hCG, suggesting that the conformation of hCG in the crystal was not affected by this reductant. The structure has ultimately confirmed the presence of all 11 disulfide bridges. Lower concentrations of DTT (for example, 1 mM) did give crystals (in contrast to no crystals without DTT), but these crystals failed to grow to a suitable size. The molar ratio of DTT to hCG in these crystallizations was somewhat lower than in disulfide reduction experiments on hCG [23,24], and since DTT is unprotonated in its reactive state its potency will be much reduced in the crystallization medium at pH 4.4.

These crystals belong to space group $P6_{5(1)}22$ with cell dimensions roughly isomorphous with crystals from other sources (Table 2) at room temperature. For data collection at 110 K with cooling from the Oxford cryostream system, crystals were transfered to 40% SAS plus 20% ethylene glycol (EG) with 30 mM sodium acetate at pH 4.4 as buffer. The transfer was done in four steps starting from 5% EG, and increasing 5% at a time, with a 5–10 minute soak between steps. The frozen crystals have anisotropic changes in cell dimensions (Table 2).

Table 2. Cell dimensions of different human chorionic gonadotropin crystals.

	T (K)	a, b (Å)	c (Å)	Method
u-hCG, asialic	~293	88.6	177.2	precession photo [12]
u-hCG, HF-treated	~293	88.7	177.1	precession photo [12]
r-hCG, HF-treated	~277	87.6	177.2	X4A IP data
Se-hCG, HF-treated	~293	88.8	176.8	SDMW detector
Se-hCG, HF-treated	~110	85.1	177.7	X4A IP data

u-hCG = urinary human chorionic gonadotropin; r-hCG = recombinant hCG; Se-hCG = selenomethionyl hCG; HF = hydrogen fluoride.

Diffraction measurements

The MAD data were collected at the HHMI X4A beam line of the National Synchrotron Light Source (NSLS). A 2 mrad swathe of radiation was sagittally focused through a double Si 111 crystal monochromator and vertically focused by a Rh-coated spherical mirror. A vertical slit before the monochromator was set at about 1 mm and a 0.3 mm collimator was used. The NSLS source was operated at 2.5 GeV with currents starting at approximatley 240 mA.

The X-ray absorption spectrum near the K-shell edge of selenium of a frozen crystal was measured by a scintillation counter and was used to select monochromator settings for the peak (maximum f″; 0.9792 Å), the inflection point (minimum f′; 0.9793 Å) and two remote energies at 50 eV below (0.9832 Å) and 150 eV above (0.9679 Å) the absorption edge. This spectrum was transformed to obtain approximate f′ and f″ values at each wavelength (Fig. 11). The choice of the remote wavelength was limited by the non-uniformity of X-ray intensities and the non-ideal focusing across a large energy span. The crystal was aligned so that the six-fold axis of point group 622 was along the horizontal spindle direction and a two-fold axis was vertical. A sweep of approximately 30° was collected starting from this orientation and the missing cone region was collected by rotating the two-fold axis to the spindle direction. In both cases, Bijvoet pairs were measured on the left and right sides of the imaging plates. The detector plates were set at 300 mm on a carousel mechanism.

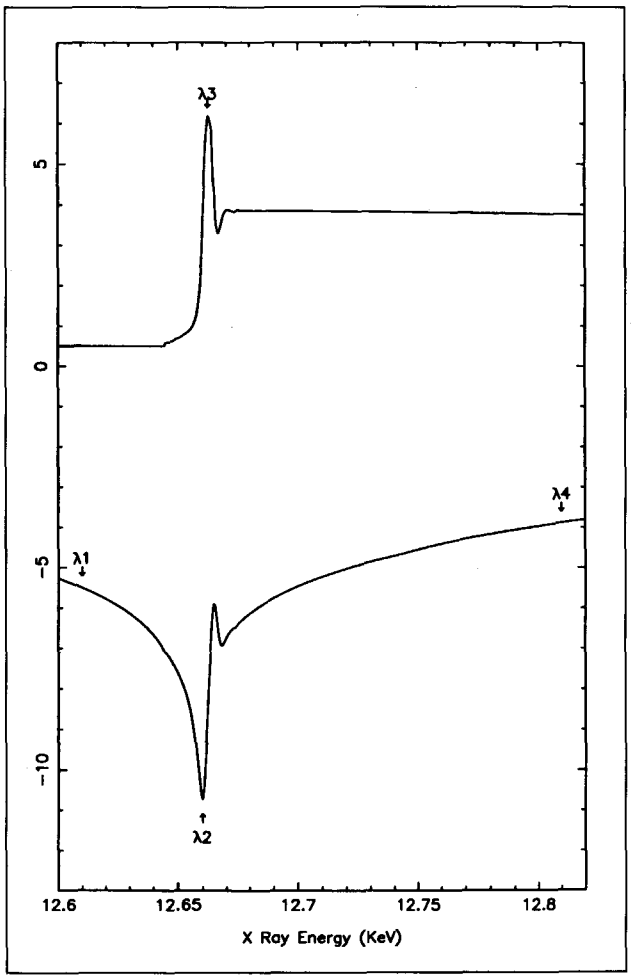

Fig. 11. Anomalous scattering factors f′ and f″ of the hCG crystal as a function of X-ray energy, as transformed from the crystal absoption spectrum.

Oscillation diffraction patterns were recorded on Fuji phosphorous imaging plates and read out with a Fuji imaging plate scanner. These images were indexed and integrated through profile fitting by DENZO [50]. They were then scaled within each wavelength by ROTAVATA and AGROANO (modified from AGROVATA of the CCP4 package by W Weis) [SERC (UK) Collaborative Computing Project 4, Daresbury Laboratory, Warrington, UK, 1979] where symmetry-related reflections or redundant measurements were not merged. Data statistics are given in Table 3.

Phase determination

The deterministic approach of MAD phasing [16] was carried out by using the MADSYS programs developed in our laboratory. Parametrized local scaling was done within each wavelength by ANOSCL to reduce noise in the Bijvoet signal and among the

Table 3. Data statistics for MAD phasing.

Wavelength (Å)	No. of unique reflections	Redundancy	Completeness (%)	R_{sym}[a](%)
0.9832 (remote λ_1)	17 477 (8490)	4.2 (4.0)	97.5 (96.6)	5.5 (11.2)
0.9793 (inflection λ_2)	17 405 (8444)	3.9 (4.0)	96.9 (96.0)	5.7 (11.2)
0.9792 (peak λ_3)	17 419 (8468)	4.0 (3.8)	97.0 (96.3)	6.0 (11.8)
0.9679 (remote λ_4)	17 413 (8459)	3.9 (3.7)	97.0 (96.2)	5.8 (11.8)

Unique reflections are determined by point group 622 (not 6/mmm) to distinquish Bijvoet-related reflections. Values shown are for reflections at 15–2.8 Å and 3.5–2.8 Å (in parentheses) spacing intervals. [a]$R_{sym} = \Sigma\Sigma|I_i(h) - \langle I(h)\rangle|/\Sigma\Sigma I_i(h)$ where for each reflection h, $I_i(h)$ is the i^{th} observation and $\langle I(h)\rangle$ is the weighted mean of all observations.

four wavelengths by WVLSCL to reduce noise in the dispersive signal. Properly scaled data were entered into MADLSQ to obtain the solution to the MAD phasing equations. Three quantities, $|^{\circ}F_A|$, the amplitude from the normal scattering of the anomalous scatterers, $|^{\circ}F_T|$, the amplitude from the scattering of the normal scatterers, and $\Delta(\phi_T - \phi_A)$, the phase difference between these two sets of scatterers, were extracted. This was followed by MERGIT to merge redundant evaluations and produce MAD phasing statistics (Table 4).

Table 4. MAD phasing statistics.

Observed ratios[a] (15.0–2.8 Å)

	λ1	λ2	λ3	λ4	f′ (e)	f″ (e)
λ1	0.044 0.034	0.048	0.044	0.047	−5.4 (−5.5)	0.9 (0.5)
λ2		0.062 (0.037)	0.041	0.056	−11.3 (−11.1)	4.2 (4.0)
λ3			0.073 (0.039)	0.048	−8.5 (−8.7)	6.0 (6.6)
λ4				0.061 (0.038)	−3.8 (−3.9)	3.8 (3.8)

Observed ratios[a]

	15.0–5.0 Å	5.0–3.5 Å	3.5–2.8 Å	
Peak Bijvoet	0.074 (0.031)	0.065 (0.039)	0.113 (0.078)	$R(|^{\circ}F_A(h)|)^b = 0.38$ $R(|^{\circ}F_T(h)|)^b = 0.05$
Max. dispersive	0.053	0.043	0.093	$\langle\Delta(\Delta\phi(h))\rangle^c = 45.6°$
Min. dispersive	0.027	0.032	0.086	$R(|^{\circ}F_A \text{ model}|)^d = 0.40$

Observed ratios are Bijvoet differences ratios (diagonal elements) at each wavelength (values in parentheses are of centric reflections, which indicate the level of noise in the anomalous signals) and dispersive difference ratios (off-diagonal elements) between pairs of wavelengths. Scattering factors were initially derived through the absorptive and dispersive transformations of a composite absorption spectrum and refined iteritively by MADLSQ. $R(|^{\circ}F_A(h)|)$, $R(|^{\circ}F_T(h)|)$ and $\Delta(\Delta\phi(h))$ are merging statistics of 26 833 pairs of independent evaluations for 10 011 unique reflections from the MAD least-squares fitting. [a]$\langle\Delta|F|^2\rangle^{1/2}/\langle|F|^2\rangle^{1/2}$ where $\Delta|F|$ is the absolute value of either the Bijvoet or dispersive difference. [b]$\Sigma\Sigma|F_i(h) - \langle F(h)\rangle|/\Sigma\Sigma\langle F(h)\rangle$ where for each reflection h, $F_i(h)$ is the i^{th} determination from MADLSQ. [c]Average difference between independent determinations of $\Delta\phi(h)$ from MADLSQ. [d]$\Sigma||^{\circ}F_A(h)| - |^{\circ}F_{A,calc}(h)||/\Sigma|^{\circ}F_A(h)|$ where $|^{\circ}F_{A,calc}(h)|$ is the structure factor calculated from the anomalous scattering model.

The positions of three of the selenium sites were determined with PATSOL [18] from the $|^{\circ}F_A|$ Patterson map calculated with reflections of spacings between 10 Å and 3 Å. A fourth, high B-factor selenium site was located in a $|^{\circ}F_A|$ difference-Fourier synthesis computed at 5 Å resolution. The positions and B-factors for these selenium sites were refined and input into MADFAZ to obtain values for ϕ_T and figure of merit to 2.8 Å resolution in both space group enantiomorphs $P6_122$ and $P6_522$. A clear solvent boundary was observed in the electron density map in space group $P6_522$. Phases were also computed with greater completeness with phase probability distributions [17] produced by MADABCD. Solvent flattening was applied to the resulting map [19] and phases were refined by iterative phase combination.

Model building and refinement

Map interpretation and model building were carried out with computer graphics. A C_α model was first constructed in FRODO [51] by tracing the chains in the MAD-phased and solvent-flattened maps at 2.8 Å resolution. Full-atom residues were then built in program 'O' [52] from approximate C_α positions with fragments from the structural data base and fitted in FRODO [51] to the electron densities.

For refinement, the small Bijvoet differences at the first wavelength were ignored and the data were processed to the diffraction limit of 2.6 Å spacings and reduced as if there had been no anomalous scattering. After several iterations of molecular dynamics and restrained refinement with X-PLOR [53] and PROLSQ [54] and manual rebuilding, we have obtained a model of 1554 non-hydrogen atoms including 64 water molecules. This model gives a 19.9 % crystallographic R value for 8876 reflections ($|F| > 3\sigma_{|F|}$) between 5 Å and 2.6 Å spacings. Deviations from stereochemical ideality are typified by the rms values for bond lengths of 0.017 Å and for bond angles of 2.3°. The thermal restraints used give rms deviations of 4.7 Å^2 for the B-factors of bonded atoms. Resolution dependence in the refinement is shown in Table 5.

Table 5. Data statistics for refinement.

	5.0–4.0 Å	4.0–3.0 Å	3.0–2.8 Å	2.8–2.7 Å	2.7–2.6 Å	5.0–2.6 Å
R_{sym}[a] (%)	2.4	5.2	11.7	16.1	18.8	4.5
Completeness[b]	92.9	90.0	80.0	74.8	71.6	84.9
R_{ref}[c] (%)	13.6	20.5	26.7	29.2	31.0	19.9

[a]See Table 3 for definition. [b]Completeness figures (%) are given after applying a 3σ cutoff. [c]$R_{ref} = \Sigma||F_{obs}(h)| - |F_{calc}(h)||/\Sigma|F_{obs}(h)|$

Structural analysis and comparison

Surface areas were computed with DSSP [55]. Electrostatic and surface curvature calculations were done by GRASP [56]. Structural comparisons were performed by the TOSS procedure [24].

The atomic coordinates have been deposited in the Brookhaven Protein Data Bank.

Note added in proof

The report of the hCG structure, as independently solved by AJ Lapthorn, DC Harris, A Littlejohn, JW Lustbader, RE Canfield, KJ Machin, FJ Morgan and NW Isaacs, is now in press with *Nature*.

Acknowledgments: We thank C Ogata and X Yang for their help with the beamline; H-E Aronson, PD Kwong and X Zhao for assistance in data collection; NQ McDonald for NGF coordinates; and L Tong, S Birken and NQ McDonald for useful discussions. This research was supported in part by grants GM34102 (WAH) and HD15454 (REC and JWL) from the NIH, DMB-8917570 (WAH) from the NSF, and H9-181-103 (JWL and REC) from the WHO. HW is a postdoctoral fellow of the Aaron Diamond Foundation. Beamline X4A at the National Synchrotron Light Source, a DOE facility, is supported by the Howard Hughes Medical Institute.

References

1. Pierce, J.G. & Parsons, T.F. (1981). Glycoprotein hormones: structure and function. *Annu. Rev. Biochem.* **50**, 465–495.
2. Hussa, R.O. (1981). Human chorionic gonadotropin; a clinical marker: review of its biosynthesis. *Ligand Rev.* **3 (suppl 2)**, 6–43.
3. Forastieri, H. & Ingham K.C. (1982). Thermal stability of human chorionic gonadotropin. *J. Biol. Chem.* **257**, 7976–7981.
4. McFarland, K.C., *et al.*, & Seeburg, P.H. (1989). Lutropin-choriogonadotropin receptor: an unusual member of the G protein-coupled receptor family. *Science* **245**, 494–499.
5. Minegishi, T., *et al.*, & Igarashi, M. (1990). Cloning and sequencing of human LH/hCG receptor cDNA. *Biochem. Biophys. Res. Commun.* **172**, 1049–1054.
6. LaPolt, P.S., Nishimori, K., Fares, F.A., Perlas, E., Boime, I. & Hsueh, A.J. (1992). Enhanced stimulation of follicle maturation and ovulatory potential by long acting follicle-stimulating hormone agonists with extended carboxy-terminal peptides. *Endocrinology* **131**, 2514–2520.
7. Kessler, M.J., Reddy, M.S., Shah, R.H. & Bahl, O.P. (1979). Structure of N-glycosidic carbohydrate units of human chorionic gonadotropin. *J. Biol. Chem.* **254**, 7901–7908.
8. Lustbader, J.W., Yarmush, D.L., Birken, S., Puett, D. & Canfield, R.E. (1993). The application of chemical studies of human chorionic gonadotropin to visualize its three-dimensional structure. *Endocr. Rev.* **14**, 291–311.
9. Rao, S.N.V. & Moyle, W.R. (1994). Modeling human chorionic gonadotropin using distance geometry and immunological constraints. In *Techniques in Protein Chemistry V.* pp. 413–420, Academic Press, San Diego.
10. Wiley, K.P. & Leidenberger, F. (1989). Functionally distinct agonist and receptor-binding regions in human chorionic gonadotropin. *J. Biol. Chem.* **264**, 19716–19729.
11. Noort, M.H., *et al.*, & Meloen, R.H. (1992). Synthetic peptides based upon a three-dimensional model for the receptor recognition site of follicle-stimulating hormone exhibit antagonistic or agonistic activity at low concentrations. *Proc. Natl. Acad. Sci. USA* **89**, 3922–3926.
12. Lustbader, J.W., *et al.*, & Canfield, R.E. (1989). Crystallization and characterization of human chorionic gonadotropin in chemically deglycosylated and enzymatically desialylated states. *Biochemistry* **28**, 9239–9243.
13. Harris, D.C., Machin, K.J., Evin, G.M., Morgan, F.J. & Isaacs, N.W. (1989). Preliminary X-ray diffraction analysis of human chorionic gonadotropin. *J. Biol. Chem.* **264**, 6705–6706.
14. Hendrickson, W.A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* **254**, 51–58.
15. Hendrickson, W.A., Horton, J.R. & LeMaster, D.M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. *EMBO J.* **9**, 1665–1672.
16. Hendrickson, W.A. (1985). Analysis of protein structure from diffraction measurement at multiple wavelengths. *Trans. Am. Crystallogr. Assoc.* **21**, 11–21.
17. Pahler, A., Smith, J.-L. & Hendrickson, W.A. (1990). A probability representation for phase determination from multiwavelength anomalous dispersion. *Acta Crystallogr. A* **46**, 537–540.
18. Tong, L. & Rossmann, M.G. (1993). Patterson-map interpretation with noncrystallographic symmetry. *J. Appl. Crystallogr.* **26**, 15–21.
19. Wang, B.C. (1985). Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.* **115**, 90–112.
20. McDonald, N.Q. & Hendrickson, W.A. (1993). A structural superfamily of growth factors containing a cystine knot motif. *Cell* **73**, 421–424.
21. Murray-Rust, J., *et al.*, & Bradshaw, R.A. (1993). Topological similarities in TGF–β2, PDGF-BB and NGF define a superfamily of polypeptide growth factors. *Structure* **1**, 153–159.
22. Mise, T. & Bahl, O.P. (1981). Assignement of disulfide bonds in the β subunit of human chorionic gonadotropin. *J. Biol. Chem.* **256**, 6587–6592.
23. Mise, T. & Bahl, O.P. (1980). Assignment of disulfide bonds in the α subunit of human chorionic gonadotropin. *J. Biol. Chem.* **255**, 8516–8522.
24. Hendrickson, W.A. (1979). Transformations to optimize the superposition of similar structures. *Acta Crystallogr. A* **35**, 158–163.
25. Garnier, J. (1978). Molecular aspects of the subunit assembly of glycoprotein hormones. In *Structure and Function of Gonadotropins.* (McKerns, K.W., ed), pp. 381–414, Plenum Press, New York.
26. 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. Mol. Biol.* **120**, 97–120.
27. Huth, J.R., Mountjoy, K., Perini, F. & Ruddon, R.W. (1992). Intracellular folding pathway of human chorionic gonadotropin β subunit. *J. Biol. Chem.* **267**, 8870–8879.
28. Aloj, S.M., Edelhoch, H., Ingham, K.C., Morgan, F.J., Canfield, R.E. & Ross, G.T. (1973). The rates of dissociation and reassociation of the subunits of human chorionic gonadotropin. *Arch. Biochem. Biophys.* **159**, 497–504.
29. Morgan, F.J., Canfield, R.E., Vaitukaitis, J.L. & Ross, G.T. (1974). Properties of the subunits of human chorionic gonadotropin. *Endocrinology* **94**, 1601–1606.
30. Huth J.R., Perin F., Lockridge, O., Bedows, E. & Ruddon, R.W. (1993). Protein folding and assembly *in vitro* parallel intracellular folding and assembly. *J. Biol. Chem.* **268**, 16472–16482.
31. Bedows, E., Huth, J.R., Suganuma, N., Bartels, C.F., Boime, I. & Ruddon, R.W. (1993). Disulfide bond mutations affect the folding of the human chorionic gonadotropin-β subunit in transfected Chinese hamster ovary cells. *J. Biol. Chem.* **268**, 11655–11662.
32. Janin, J., Miller, S. & Chothia, C. (1988). Surface, subunit interfaces and interior of oligomeric proteins. *J. Mol. Biol.* **204**, 155–164.
33. McDonald, N.Q., Lapatto, R., Murray-Rust, J., Gunning, J., Wlodawer, A. & Blundell, T.L. (1991). New protein fold revealed by a 2.3Å resolution crystal structure of nerve growth factor. *Nature* **354**, 411–414.
34. Schlunegger, M.P. & Grutter, M.G. (1992). An unusual feature revealed by the crystal structure at 2.2Å resolution of human transforming growth factor-β2. *Nature* **358**, 430–434.
35. Daopin, S., Piez, K.A., Ogawa, Y. & Davies, D.R. (1992). Crystal structure of transforming growth factor-beta 2: an unusual fold for the superfamily. *Science* **257**, 369–373.
36. Oefner, C., D'Arcy, A., Winkler, F.K., Eggimann, B. & Hosang, M. (1992). Crystal structure of human platelet-derived growth factor BB. *EMBO J.* **11**, 3921–3926.
37. Segaloff, D.L. & Ascoli, M. (1993). The lutropin/choriogonadotropin receptor...4 years later. *Endocrinol. Rev.* **14**, 324–347.
38. Chen, F., Wang, Y. & Puett, D. (1992). The carboxy-terminal region of the glycoprotein hormone α-subunit: contributions to receptor binding and signaling in human chorionic gonadotropin. *Mol. Endocrinol.* **6**, 914–919.
39. Campbell, R.K., Dean-Emig, D.M. & Moyle, W.R. (1991). Conversion of human choriogonadotropin into a follitropin by protein engineering. *Proc. Natl. Acad. Sci USA.* **88**, 760–764.
40. Moyle, W.R., Campbell, R.K., Myers, R.V., Bernard, M.P., Han, Y. & Wang, X. (1994). Co-evolution of ligand–receptor pairs. *Nature* **368**, 251–255.
41. Keutmann, H.T., Charlesworth, M.C., Mason, K.A., Ostrea, T., Johnson, L. & Ryan, R.J. (1987). A receptor-binding region in human choriogonadotropin/lutropin β subunit. *Proc. Natl. Acad. Sci. USA* **84**, 2038–2042.

42. Erickson, L.D., Rizza, S.A., Bergert, E.R., Charlesworth, M.C., McCormick, D.J. & Ryan, R.J. (1990). Synthetic alpha-subunit peptides stimulate testosterone production *in vitro* by rat Leydig cells. *Endocrinology* **126**, 2555–2560.
43. Charlesworth, M.C., McCormick, D.J., Madden, B. & Ryan, R.J. (1987). Inhibition of human choriotropin binding to receptor by human choriotropin α peptides. *J. Biol. Chem.* **262**, 13409–13416.
44. Kobe, B. & Deisenhofer, J. (1993). Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature* **366**, 751–756.
45. Moyle, M.R., Bahl, O.P. & Marz, L. (1975). Role of the carbohydrate of human chorionic gonadotropin in the mechanism of hormone action. *J. Biol. Chem.* **250**, 9163–9169.
46. Matzuk, M.M. & Boime, I. (1989). Mutagenesis and gene transfer define site-specific roles of the gonadotropin oligosaccharides. *Biol. Reprod.* **40**, 48–53.
47. Matzuk, M.M., Keene, J.L. & Boime, I. (1989). Site specificity of the chorionic gonadotropin N-linked oligosaccharides in signal transduction. *J. Biol. Chem.* **264**, 2409–2414.
48. Rebois, R.V. & Liss, M.T. (1987). Antibody binding to the β-subunit of deglycosylated chorionic gonadotropin converts the antagonist to an agonist. *J. Biol. Chem.* **262**, 3891–3896.
49. Stura, E.A. & Wilson, I.A. (1990). Analytical and production seeding techniques. *Methods* **1**, 38–49.
50. Otwinowski, Z. (1990). *DENZO Data Processing Package.* Yale University, New Haven, CT.
51. Jones, T.A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Crystallogr.* **11**, 268–272.
52. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building models in electron density maps and the location of errors in those models. *Acta Crystallogr. A* **47**, 110–119.
53. Brünger, A.T. (1992). *X-PLOR Manual, Version 3.1.* Yale University, New Haven, CT.
54. Hendrickson, W.A. & Konnert, J.H. (1980). Incorporation of stereochemical information into crystallographic refinement. In *Computing in Crystallography.* (Diamond, R., Ramaseshan, S. & Venkatesan, K., eds), pp. 13.01–13.23, Indian Academy of Science, Bangalore.
55. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637.
56. Nicholls, A., Sharp, K.A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296.
57. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950.

Received: **20 Apr 1994**; revisions requested: **29 Apr 1994**: revisions received: **6 May 1994**. Accepted: **6 May 1994**.