TOWARDS AN UNDERSTANDING OF LUTEINIZING HORMONE RECEPTOR BINDING

by

GREGORY B. FRALISH

(UNDER THE DIRECTION OF DR. DAVID PUETT)

Human chorionic gonadotropin (hCG) is critical to the maintenance of pregnancy and male sexual differentiation. It is a member of the family of glycoprotein hormones, which includes luteinizing hormone (LH), follicle stimulating hormone and thyroid stimulating hormone. This family of hormones is characterized by a heterodimeric structure, consisting of a common α-subunit and a hormone specific β-subunit. hCG (and LH) bind to their common receptor, LHR, with high affinity to elicit its biological response(s). LHR is a G protein-coupled receptor with a large, leucine-rich repeat (LRR)-containing extracellular domain (ECD) that is responsible for the high affinity contacts with hormone. Since there is little structural data available on the receptor, our understanding of the binding of hCG and the LHR ECD is limited to indirect observations and models. The goal of this dissertation is to provide direct structural data for this complex system using fusion proteins of the heterodimeric hormone and hormone-receptor complexes. Therefore, an expression and purification system was developed for the two single-chain hormones, YhCG1 (β-α), which has similar bioactivity to native hCG, and YhCG3 (α-β), which has altered binding properties. The purified hormones were characterized structurally using circular dichroism (CD) and limited proteolysis. These studies have revealed that there is a remarkable variability in the structures of the different hormones. Nevertheless, the proteins can functionally interact and activate the receptor suggesting an unexpected conformational tolerance for a productive hormone-receptor interaction. Furthermore, an expression and purification system was developed for three hormone-receptor ECD fusion proteins. By expressing the ligand tethered to the full-length receptor ectodomain (YECD), expression levels of the complex were improved 20-fold over those previously reported, thus making possible detailed structural studies. This fusion complex was shown to be functional, as it can bind hCG with high affinity. Using CD spectroscopy, YECD revealed that the ECD contains between 18-20% α helix and 30% β-strand. These data provide structural evidence in support of the proposed homology model, and suggest that 8 or 9 LRRs are present in the ECD. Further, they provide proof of principle for using ligand receptor fusions to express functional complexes. The structural characterization of two C-terminal truncations of YECD demonstrated a repeating motif in the ECD. Furthermore, amino acid residues in the C-terminal region of the ECD are important in the overall stabilization of the ECD, an intriguing observation considering this region's emerging role as a signal modulator.

INDEX WORDS: human chorionic gonadotropin, LH receptor, baculovirus expression, extracellular domain, single-chain gonadotropin, yoked gonadotropin

TOWARDS AN UNDERSTANDING OF LUTEINIZING HORMONE RECEPTOR BINDING

by

GREGORY B. FRALISH

B.S., The University of Georgia, 1995

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2002

TOWARDS AN UNDERSTANDING OF LUTEINIZING HORMONE RECEPTOR BINDING

by

GREGORY B. FRALISH

Major Professor: David Puett

Committee: Prema Narayan
James Travis
Claiborne Glover
B.C. Wang
Kojo Mensa-Wilmot

Electronic Version Approved:
Maureen Grasso

Dean of the Graduate School

The University of Georgia

December 2002

ACKNOWLEDGMENTS

I would like to acknowledge my major professor David Puett for his unwavering support of my work. His persistent confidence in me as a scientist has afforded me a unique freedom in developing my work. As a result, any future successes that I may have scientifically will be a direct result of his brilliant mentorship. I would also like to acknowledge Prema Narayan. She has been critical to my development as a scientist as well. As an undergraduate, she taught me how to design and execute experiments, and through countless discussions during my graduate education, Prema has also helped me develop critical scientific thinking. Her excellent skill as a writer has been pivotal to my publishing of scientific papers. Perhaps most importantly, I consider both David and Prema friends. I would also like to acknowledge my wife Bolyn. She listened and delivered advice perfectly, exactly as I needed it. I'd like to acknowledge Judy Gray. Her technical and personal support in the lab has been so important to making many of my experiments possible. She made coming to work fun. I'd like to acknowledge Debra Mohnen. I worked in her laboratory first as an undergraduate. Her mentorship and support have been extremely helpful and important to my scientific development. Finally, I'd like to thank my committee members for their advice and critique during all phases of my graduate education.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION

The luteinizing hormone receptor (LHR) is critical for human reproductive capacity. This G protein-coupled receptor (GPCR) is activated following binding of one of its two ligands, luteinizing hormone (LH) or chorionic gonadotropin (CG). Indeed, binding of ligand and activation of receptor are dissociable events. Activation of the receptor results in the stimulation of intracellular second messenger systems, which confer cellular response and biological activity (1,2). LHR is a member of the family of glycoprotein hormone receptors including the receptors for follicle stimulating hormone (FSHR) and thyroid stimulating hormone receptor (TSHR). This group of receptors is in the superfamily of rhodopsin-like GPCRs, and they are characterized by a large (300-400 amino acid residues) extracellular domain (ECD) that is responsible for the high affinity binding. The ECDs share a similar leucine-rich repeat (LRR) motif in their sequence, and the recent identification of several new orphan receptors in fly (3,4), nematode (5) and human (3,6) has necessitated a new designation for this growing class of GPCR; namely, LRR-containing G protein-coupled receptor (LGR).

The receptor's topology, like that for the all GPCRs, is composed of an N-terminal extracellular segment, followed by seven transmembrane (TM) helices connected by three extracellular (ECL) and three intracellular loops (ICL) and an intracellular C-terminal tail that is palmitoylated and thus forms the fourth ICL (Fig. 1.1). LHR binds LH and CG with K_ds in the subnanomolar range. This high-affinity interaction results in the rearrangement of the TM domain and allows receptor coupling

to heterotrimeric G-protein(s), where LHR promotes the exchange of GDP, in the G_α subunit, with GTP. Dissociation of the $G\alpha_s$ subunit from the heterotrimer results in the stimulation of adenylate cyclase activity and a concommitant rise in intracellular levels of cAMP. Temporal regulation is achieved primarily through the α-subunit's inherent GTPase activity, which can be enhanced by accessory proteins (GAPs), as well as by cAMP phosphodiesterases. High intracellular cAMP stimulates an increase in protein kinase A activity (PKA), which phosphorylates several intracellular proteins and alters expression patterns for a variety of genes (7). Most of the reproductive biological action of LH and CG is conferred through the activation of the $G\alpha_s$-adenylate cyclase-cAMP pathway (1,2,8). However, it is apparent that LHR can couple to alternate pathways, namely the inositol phosphate (IP) pathway (9,10). Indeed, most GPCRs are thought to have the ability to couple to multiple G-proteins and thus stimulate multiple pathways (11).

Biological roles of LHR and CG- In females, LHR is expressed in the theca cells of the ovary and is important in the later stages (preovulatory stage) of follicular development as the granulosa cells begin to express LHR as well. The mid-cycle surge in serum levels of LH stimulates an inflammatory-like cascade with the upregulation of α_2-macroglobulin, prostaglandin biosynthesis and several matrix metalloproteinases resulting in release of the ovum (12). The accessory cells of the follicle luteinize to form the corpus luteum, which is the primary site of steroid hormone biosynthesis during pregnancy. Luteinization is marked primarily by the rapid expression of LHR in the luteal cells, the progesterone receptor, COX2, and several enzymes involved in steroidogenic biosynthesis (P450) (7). Following fertilization and implantation, CG is

produced and secreted by the trophoblasts of the placenta. CG binds LHR at the corpus luteum and stimulates the production of progesterone. Progesterone is important in the development of the placenta as well as suppressing menstrual contractions, thus maintaining pregnancy (13). Inactivating mutations of LHR in females prevent ovulation and result in amenorrhea, whereas activating mutations present no apparent phenotype (14).

In males, LHR is expressed primarily in the Leydig cells of the testes. Sexual development of the 46XY fetus is thought to be governed in part by LHR. Leydig cell differentiation and activation of gonadal steroidogenesis, namely testosterone, are stimulated primarily by activation of LHR by placental CG. A functional LHR is critical as children with severe inactivating mutations in the LHR exhibit pseudo-hermaphroditism and Leydig cell hypoplasia (14). Less severe forms of mutation can occur with partial LH receptor function. Activating mutations of LHR cause the early onset of puberty. This phenotype, as stated above, is only manifested in males. The more severe the mutation, that is the higher the basal stimulation of the receptor, the more severe the phenotype. In males with the D578Y mutation, which shows the highest level of basal cAMP levels in vitro (15), puberty onset was at the age of one year of age (16,17). Furthermore, a D578H mutation, which displays high increases in both basal cAMP and IP levels in vitro caused Leydig cell tumors (18).

Binding of CG to LHR- FSH, LH and CG, along with thyroid stimulating hormone (TSH), form the family of glycoprotein hormones that are similar in sharing a heterodimeric structure consisting of a common α-subunit and a hormone specific β-subunit (19). The β-subunits for CG and LH show a high degree of sequence similarity

(85% in the human hormones), with the major difference being in the 30 amino acid residue extension on the C-terminus of CG (C-terminal peptide; CTP) (19), which contains four O-glycosylation sites and is responsible for hCG's increased circulatory half-life compared to LH (20). Thus, it is not surprising that they share the same receptor. hFSH and hTSH possess lower similarity to CG, 36% and 46%, respectively.

The structure of CG was the first in the family of glycoprotein hormones to be elucidated (Fig. 1.2) (21,22). The structure revealed that the two subunits fold with remarkable similarity despite containing little sequence homology. Namely, both subunits fold into three alternating loops, with little well-defined secondary structure and a minimal hydrophobic core. The α-subunit's 10 cysteines form 5 intrachain disulfide bonds while the β-subunit's 12 cysteines form 6 intrachain disulfide bonds. Each subunit has a growth factor-like, cystine knot, formed from three disulfide pairs, where two of the disulfides form a ring that the third passes through. The cystine knots are central to CG's overall structure. They are known to be critical in the folding and dimerization of the hormone (23); because of this, elucidating their role in binding and signal transduction has been enigmatic as mutations to the core cysteines often prevented secretion, hampering detailed analyses. Nevertheless, early studies did suggest that cysteines in both subunits were important for binding to LHR (24).

The folding pathway for the β-subunit of CG has been studied extensively. Interestingly, several non-native disulfide pairs form during folding before the known structure is achieved (25). The last disulfide pairing forms a loop that wraps around the α-subunit and has been appropriately deemed the 'seat-belt' loop. The dimer is stabilized further by a rather extensive network of contacts along the interface of the two

subunits. The other glycoprotein hormones were postulated to fold similarly given the strict conservation of the spacing of the cysteine residues in the various β-subunits. This was confirmed by the recently published crystal structure of FSH (26), which displays an identical overall topology. The solution structure of the free α-subunit was studied using NMR. The free subunit displayed a similar overall three looped structure as was previously described in the CG crystal structure; however, the long loop, which forms extensive contacts with β, was disordered, lacking even the short, five residue α-helix that is present in the intact heterodimer (27).

The structure of hCG has provided a framework for interpretation of the numerous studies involving mutagenesis, chimeras, peptides and antibodies. Briefly, in the β-subunit, deletion of seven N-terminal residues and forty-three C-terminal residues did not affect holoprotein formation or in vitro bioactivity (28). Residues 93-100 of the seat-belt loop, also called the determinant loop, are critical for binding specificity (29,30). In the α-subunit, regions important for binding include the C-terminal amino acids 90-92 (31,32), regions in the hairpin turns 1 and 3, and the α-long loop, which forms a groove with the β determinant loop and contains the α-helix where the glycosylated Asn 52 was found to be critical for signal-transfer (33,34). Thus, binding is achieved through interactions with receptor that map primarily to the central portions of the dimer interface. It has been suggested that the role of high affinity binding is simply to tether the hormone to the receptor, thus allowing for important low affinity interactions with other portions of the hormone and the receptor's extracellular loops and/or TM helices. This model is supported by experiments showing that LHR, with no N-terminal binding domain, can still respond to high concentrations of hCG (35). The tip of the α-subunit

containing the β-turns 1 and 3 is implicated in interactions with the receptor TM domain and extracellular loops (36,37). Indeed, hormone binding and receptor activation are separable events (38).

LHR is composed of two halves of roughly equal size, the extracellular domain (341 residues) and the transmembrane domain (299 residues). The gene for LHR consists of 11 exons. The first ten and a portion of the eleventh code for the extracellular domain, while most of the eleventh codes for the TM domain. LHR binds CG via its large (341 amino acid residues in rat LHR) N-terminal ECD. All of the contacts necessary for high affinity interactions are located in this domain (39,40). The ECD is commonly delineated into three regions: the N-terminal Cys-rich region, the central LRR region and the C-terminal Cys-rich region, also called the hinge region. The N-terminal Cys-rich region consists of the first 29 residues with four cysteine residues interspersed. This sequence has no known homology outside of the glycoprotein hormone receptor family but is important for hormone binding. Mutation of the conserved cysteine residues results in the loss of binding activity (41) while the homologous region in TSHR is not critical for binding of TSH (42). Furthermore, deletion of this region results in the loss of binding activity (43).

Several homology models have been proposed for the LRR region (residues 29-245) (Fig. 1.3) (37,44,45) based upon the crystal structure of ribonuclease inhibitor (RI), which is composed entirely of 15 LRR's (46). Each of the β-α hairpin repeats of RI lie roughly in a plane with neighboring repeat units (46), and, in tandem, these repeat units form a horseshoe-shaped molecule with parallel β-sheets lining the interior cusp. The flexibility and large, solvent-exposed parallel β-sheet surface of the LRRs in RI are

fundamental to its specific and high affinity interactions with ribonuclease A (47). Several structures of LRR proteins are now available allowing for the classification of several subfamilies wherein LHR has been designated a 'typical' LRR protein (48). Mutations to residues on the interior of the cusp, in regions known to be important for binding, illustrate the importance of charge in the binding of hormone to receptor (44,49). Consistent with these data, the surface charge potential of the ECD's interior cusp is negative (37,44), while the hormone's surface charge is net positive on one face of the hormone. Interestingly, it is this side of the hormone that a majority of the important binding regions occur (as discussed above) (21,22,37). The first eight exons of LHR are thought to encompass most of the binding domain based on peptide competition studies (50), exon deletions (43), and C-terminal truncations (51), and it is within these coding regions that the LRR motif is most closely followed. More recently, LRR's 2 and 4 were shown to be especially important for binding of hCG (52,53).

The C-terminal Cys-rich region (residues 245-341), also denoted the hinge region, contributes little to high affinity binding (43,51). Mutation of the individual cysteine residues cause intracellular retention of the receptor but does not affect hormone binding (41). Homologous protein sequences outside of the LGR family have not been found, however based on the spacing of its cysteine residues, this region was suggested to contain a chemokine-like fold (44). The hinge region, however has an important role in signal transduction (54-56).

The ternary model of GPCR activation predicts that a liganded receptor is in equilibrium between two conformations, the inactive R_o and active R* (57). This model suggests that, in the absence of ligand, the receptor makes short-lived transitions to the

R* conformation, but in the presence of agonist the R* conformation is stabilized by the breaking and reformation of interactions in the TM domain. The R* conformation is competent for coupling to G-protein. This model accounts for the observation that an increase in the number of LH receptors expressed heterologously in cell culture results in an increase in the basal level of cAMP. The active state can be a multitude of conformations, which allows for differential coupling to alternate G-proteins, and the varying levels of constitutive activity observed in mutants (14,57). In LHR, TM helices 6 and 7 are hot spots in naturally occurring mutations that display constitutive activity (14). This region has been studied extensively, and important interactions between TM 6 and TM 7 have been suggested (58,59). With the elucidation of the three dimensional structure of the GPCR rhodopsin (60) , homology modeling of LHR and LHR mutants has been attempted (61,62). Modeling has been utilized with new and existing mutagenesis data on the TM domain to reveal an interesting correlation between the solvent accessible surface area induced at the intracellular side of TM domain by mutation and the activation of the receptor (63). The generation of solvent accessibility may be important for allowing interactions between G-protein and receptor to occur. These interactions are thought to occur primarily at the third ICL and TM 6 (64).

The mechanism by which hormone binding induces receptor activation is still poorly understood. Proposed models of the molecular nature of the activation are most commonly based upon mutation data, which may describe an alternate activated state (R*) that is not necessarily the state achieved by hormone binding. As mentioned previously, there are different models for receptor activation upon ligand binding. The most common model suggests that the ECD acts essentially as a tether, for low affinity

interactions between the hormone and receptor TM domain, with data supporting an interaction between the α-tip of the hormone and the ECLs of the TM domain. However, recent data have suggested a more active role for the ECD in receptor activation. In the homologous TSHR, removal of the receptor's ECD resulted in the constitutive activation of the receptor, suggesting a role for the ECD in constraining the TM domain in an inactive state (65). Data consistent with this idea were observed for FSHR and LHR, where regions of LRR 4 were shown to interact with both hormone and regions of the TM domain (53,66). Furthermore, in LHR, substitution of a conserved Ser residue in the hinge region (Ser 277 of human LHR, Ser 255 in rLHR) caused activation of the receptor with hydrophobic residues displaying the strongest induction of the active state (67). Furthermore, peptides derived from residues (rLHR amino acid residues 250-268) in the hinge region could interact directly with hormone (the α-subunit) and this interaction was inhibited by ECL 2 (56). Independently, using chimeric receptors, Nishi et al. observed a direct interaction between exoloop 2 and the conserved region in exon 9 of the human LHR ECD (residues 269-285) (55).

An emerging model of receptor activation has been proposed based upon the convergence of the aforementioned data: hCG interacts with LRRs 2 and 4 in the ectodomain, and the α-subunit displaces the hinge region from interactions with ECL 2 which causes relaxation of the receptor allowing coupling to G-protein. Specificity is achieved through interactions with the β-subunit with specific LRR's, while the α-subunit has a common role in glycoprotein receptor activation in displacing constraining interactions between the hinge region of the ECD and ECL 2 (55). It should be mentioned that the role of other residues in the hinge region have been characterized as

well. Specifically, the charged residues Glu^{332} and Asp^{333} in exon 11 are critical for signal transfer (54), whereas exon 10 can be completely deleted without loss of function (68,69).

There are no direct structural data available for the above models. Indeed, direct structural measurements on the receptor are extremely limited as even structural work on just the ECD domain has been hampered in large part by the inefficient production of the LHR ECD (as well as the ECD's for the other glycoprotein hormone receptors) with commonly utilized expression systems. Recently, we and others have reported significant improvements in receptor ECD expression levels as well as purification strategies for the proteins (70-72). We observed considerable increases in levels of secreted protein, when the single chain ligand (Yoked hCG) was fused to the N-terminus of the ECD (residues 1-341). The protein was secreted from insect cells at levels nearly 20-fold higher than previously reported, and it retained the ability to bind hormone with high affinity. It was suggested that expression problems with the LHR ECD could be overcome by coexpression with ligand that results in the reduction in the flexibility or an increase in stability of the ECD through the interactions with hormone. Indeed, a similar conclusion was made when high levels of expression were observed in Chinese hamster ovary cells (CHO) when the ligand hCG was coexpressed with a shortened version of the the porcine LHR (residues 1-277). No functional data were measured for the CHO-expressed ECD, however a CD spectrum was recorded for the complex with limited analysis of the spectrum, representing the first structural data of any kind on a glycoprotein hormone receptor (70). A CD spectrum for FSHR was reported subsequently (72), and both spectra were in general agreement with the presence of

significant α and β secondary structure in the ECDs of FSHR and LHR. Thus, these data provided the first structural evidence, albeit limited, for the presence of the LRR fold in the ECD's of the glycoprotein hormone receptors.

The feasibility of our fusion protein approach for producing high quantities of protein for structural studies was bolstered by numerous reports demonstrating the functionality of a single chain hCG, namely the C-terminus of the β-subunit was fused to the N-terminus of the α-subunit (N-β-α-C) (73,74). This configuration was chosen by both groups independently primarily to maintain the free C-terminus on the α-subunit, since this region is important for binding, and to enable the placement of the CTP of the β-subunit as an endogenous linker. Indeed, the CTP is not resolved in the crystal structure of hCG and is thought to contain significant conformational flexibility (21,22). The single chain hCG, deemed yoked hCG (YhCG1) or tethered hCG, displayed in vitro (73,74) and in vivo (74) activity similar to that of native heterodimeric hCG. Yoked CG was also extremely useful for the detailed analysis of hormone variants, which contained mutations that hindered dimerization and secretion. In being translated as a single polypeptide chain, the obligatory dimerization step was bypassed, allowing for more efficient secretion of these mutant hormones.

Single chain CG was used to analyze the effects of mutations to the cysteines involved in forming the cystine knot motif in both subunits (75,76). It was shown, remarkably, that disruption of this motif, thought to be critical in the maintenance of structural integrity in the hormone, did not affect in vitro bioactivity. This suggested a diminished role for this motif in binding and transducing a signal in LHR. Furthermore, it was shown that the intrachain disulfide pair, Cys93-Cys100, formed last in the folding

of the β-subunit (25) to define the determinant loop, which is intimately involved in receptor recognition (29), is not absolutely required for activity either. Furthermore, mutations that prevent native quaternary structure were introduced in the single chain CG. Analysis of the mutants revealed the expendable nature of a complete set of native quaternary interactions for bioactivity (77).

Reversal of the configuration to N-α-CTP-β-C (YhCG3) was also attempted. The expression level of this hormone was reduced compared to the β-α, and it was shown to bind to the receptor with a reduced affinity (20-50 fold reduction in IC_{50}) (78-80). Despite the reduced binding affinity, YhCG3 displayed the ability to stimulate receptor as well as YhCG1 and hCG, thus uncoupling to some degree the binding activity from the signal transduction capability of a ligand (78). Moreover, a single chain hormone consisting of a tandem array of β-subunits from FSH and hCG fused to a single α-subunit was expressed and shown to display bioactivity with both FSHR and LHR (81). This protein not only revealed interesting structure-function relationships, but it potentially represents an important molecule clinically as well.

Therefore, these studies with single chain CGs have indicated a remarkable tolerance of the receptor to bind to and recognize hormones with non-native tertiary and quaternary structure. However, without any direct structural measurements of the single chain hormones available, the nature and extent of the conformational perturbations are not well understood. Indeed, the mutations could conceivably cause only minor local changes in structure and not affect the global identity of the hormone. Nevertheless, it is assumed that YhCG1, which displays good secretion kinetics and in vitro and in vivo activity similar to native hCG, shares a similar if not identical conformation with hCG.

As such, numerous single chain versions of all of the glycoprotein hormones have now been generated (82,83). Indeed, YhFSH was reported, but not shown, to have a CD spectrum identical to that of heterodimeric FSH (72).

YhCG1 was tethered to the N-terminus of LHR using the CTP as a linker, and a Factor Xa protease recognition site was engineered between the hormone and receptor ECD (84). This yoked hormone receptor complex (YHR) was expressed in HEK 293 and COS 7 cells, and it was shown to be constitutively active for the induction of cAMP production. Moreover, the receptor was unable to bind ^{125}I-hCG. However, when soluble membrane preparations from cells expressing YHR were treated with Factor Xa to cleave the hormone from the receptor, binding was recovered. So, it was hypothesized that the yoked hormone, despite being fused to the receptor, was still able to make the specific and appropriate contacts with the receptor that blocked exogenous binding and altered receptor conformation in a way that promoted signal transduction. Indeed, steric effects were found not to be involved. When prolactin, a similarly sized protein as hCG, was fused to LHR, there was no basal cAMP response, and ^{125}I-hCG was able to bind without digestion with Factor Xa (85).

The above data justified the expression and characterization of a yoked CG-LHR ECD (YECD) as a potentially valuable protein for the production of significant quantities of LHR ECD. The following chapters contain the published report (71) documenting the expression and characterization of YECD (later called Y341). In demonstrating the functionality of YECD, the subsequent study documents the first CD spectrum for the full-length LHR ECD, as well as the expression, purification and structural characterization of several other yoked hormone receptor complexes. In the final study

presented, yoked hormones YhCG1 (β-α) and YhCG3 (α-β) are expressed, purified and structurally characterized. The structural characteristics were compared to those measured for the native heterodimeric hCG.

The specific goals of this dissertation are:

I. Establish and optimize an expression and purification system for obtaining high levels of purified hormone and hormone-receptor complexes.

II. Obtain structural information on YECD to test the current homology models.

III. Obtain direct structural measurements of YhCG1 and YhCG3 to determine the effects of: (i) the translation of hCG as a single polypeptide chain, and (ii) the subunit configuration on the overall fold of the proteins.

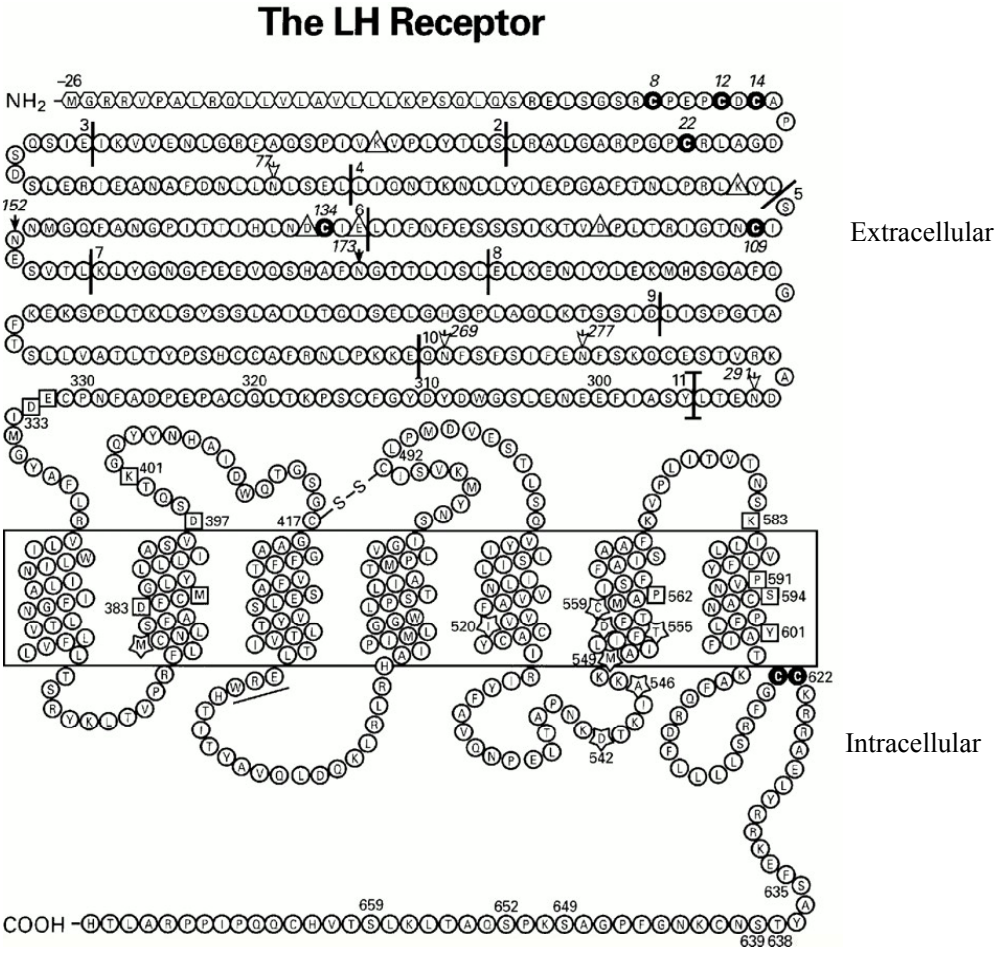

Figure 1.1. LH Receptor schematic. One dimensional representation of the GPCR LHR. The exons of the receptor are identified with bars. In the N-terminal extracellular domain, Cysteine residues that are essential for binding are highlighted in black. Putative N-glycosylation sites are designated with arrows. In the endodomain, residues that inactivate the receptor when mutated are represented with squares and ones that activate the receptor when mutated are represented with stars. Adapted from Dufau, 1998(1).

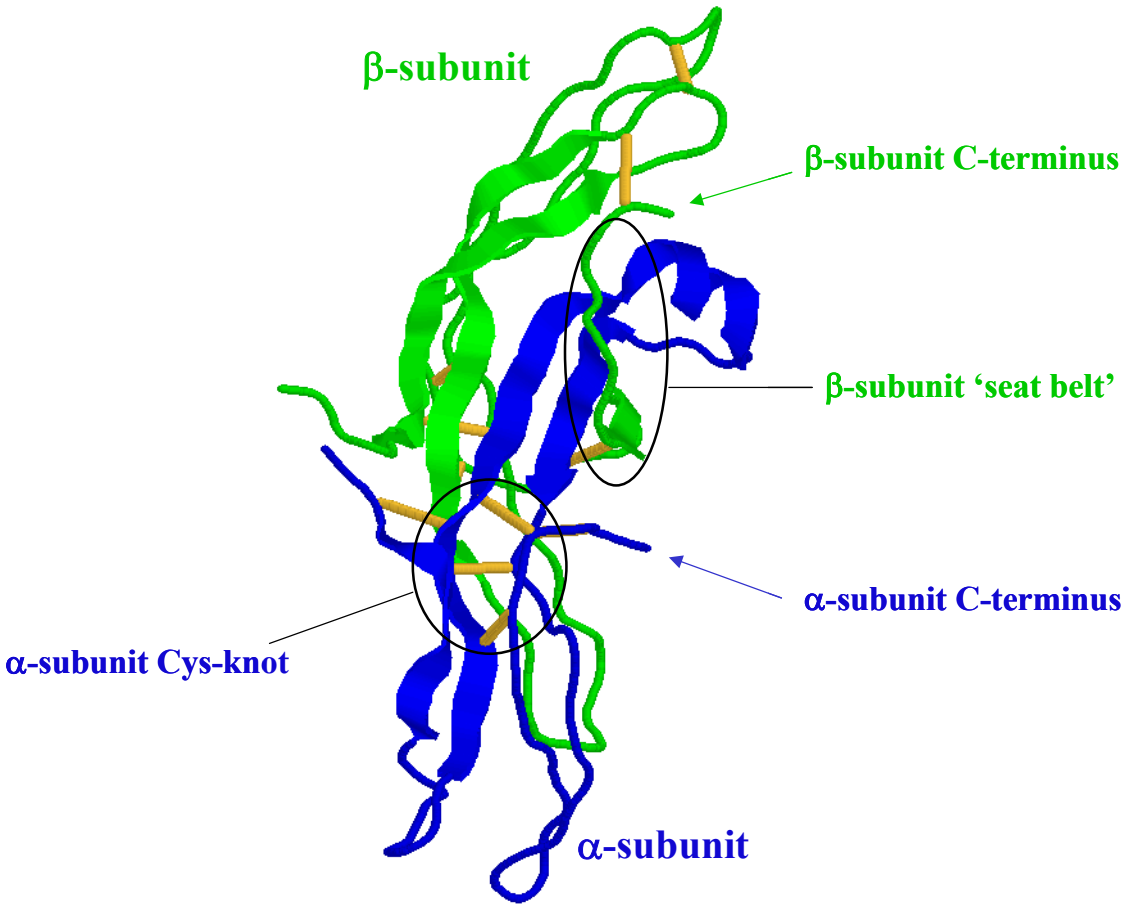

Figure 1.2. hCG crystal structure. The α- and β-subunits are delineated in blue and green, respectively. Disulfide bonds are colored in yellow. The α-subunit Cys-knot is circled, while the β-subunit's is obscured in this view. This figure was generated using the coordinates for the crystal structure of hCG (pdb accession code, 1HCN) (22).

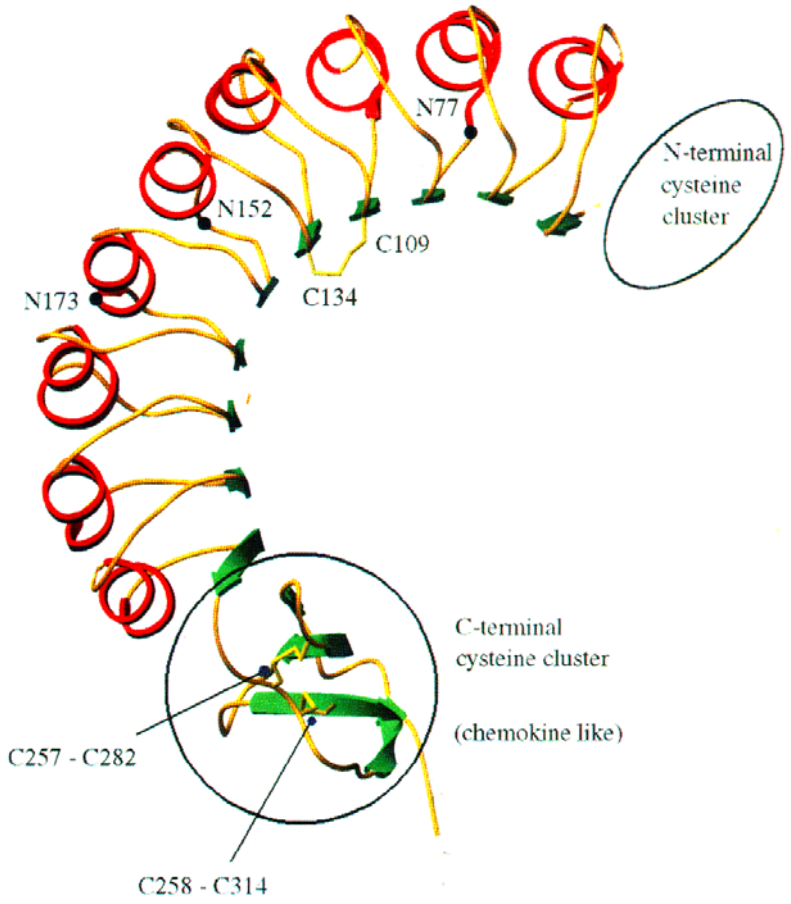

Figure 1.3. Homology model of the LHR Extracellular domain. The N-terminal Cys-rich cluster was not modeled. Putative glycosylation sites are indicated and the C-terminal Cys-rich cluster was modeled using a chemokine-like fold based solely upon the spacing of the cysteine residues. Adapted from Bhowmick *et al.*, 1996 (44).

References

1. Dufau, M. L. (1998) *Annu Rev Physiol* **60,** 461-496

2. Ascoli, M., Fanelli, F., and Segaloff, D. L. (2002) *Endocr Rev* **23,** 141-174

3. Nishi, S., Hsu, S. Y., Zell, K., and Hsueh, A. J. (2000) *Endocrinology* **141,** 4081-4090

4. Hauser, F., Nothacker, H. P., and Grimmelikhuijzen, C. J. (1997) *J Biol Chem* **272,** 1002-10010

5. Kudo, M., Chen, T., Nakabayashi, K., Hsu, S. Y., and Hsueh, A. J. (2000) *Mol Endocrinol* **14,** 272-284

6. Hsu, S. Y., Liang, S. G., and Hsueh, A. J. (1998) *Mol Endocrinol* **12,** 1830-1845

7. Richards, J. S. (1994) *Endocr Rev* **15,** 725-751

8. Richards, J. S., Russell, D. L., Robker, R. L., Dajee, M., and Alliston, T. N. (1998) *Mol Cell Endocrinol* **145,** 47-54

9. Kuhn, B., and Gudermann, T. (1999) *Biochemistry* **38,** 12490-12498

10. Gudermann, T., Birnbaumer, M., and Birnbaumer, L. (1992) *J Biol Chem* **267,** 4479-4488

11. Gudermann, T., Kalkbrenner, F., and Schultz, G. (1996) *Annu Rev Pharmacol Toxicol* **36,** 429-459

12. Richards, J. S. (2001) *Endocrinology* **142,** 2184-2193

13. Gougeon, A. (1996) *Endocr Rev* **17,** 121-155

14. Themmen, A. P. N., and Huhtaniemi, I. T. (2000) *Endocr Rev* **21,** 551-583

15. Laue, L., Wu, S. M., Kudo, M., Hsueh, A. J., Cutler, G. B., Jr., Jelly, D. H., Diamond, F. B., and Chan, W. Y. (1996) *Biochem Mol Med* **58,** 192-198

16. Muller, J., Gondos, B., Kosugi, S., Mori, T., and Shenker, A. (1998) *J Med Genet* **35,** 340-341

17. Kremer, H., Martens, J. W., van Reen, M., Verhoef-Post, M., Wit, J. M., Otten, B. J., Drop, S. L., Delemarre-van de Waal, H. A., Pombo-Arias, M., De Luca, F., Potau, N., Buckler, J. M., Jansen, M., Parks, J. S., Latif, H. A., Moll, G. W., Epping, W., Saggese, G., Mariman, E. C., Themmen, A. P., and Brunner, H. G. (1999) *J Clin Endocrinol Metab* **84,** 1136-1140

18. Liu, G., Duranteau, L., Carel, J. C., Monroe, J., Doyle, D. A., and Shenker, A. (1999) *N Engl J Med* **341,** 1731-1736

19. Hearn, M. T., and Gomme, P. T. (2000) *J Mol Recognit* **13,** 223-278

20. Matzuk, M. M., Hsueh, A. J., Lapolt, P., Tsafriri, A., Keene, J. L., and Boime, I. (1990) *Endocrinology* **126,** 376-383

21. Lapthorn, A. J., Harris, D. C., Littlejohn, A., Lustbader, J. W., Canfield, R. E., Machin, K. J., Morgan, F. J., and Isaacs, N. W. (1994) *Nature* **369,** 455-461

22. Wu, H., Lustbader, J. W., Liu, Y., Canfield, R. E., and Hendrickson, W. A. (1994) *Structure* **2,** 545-558

23. Suganuma, N., Matzuk, M. M., and Boime, I. (1989) *J Biol Chem* **264**(32)**,** 19302-19307

24. Furuhashi, M., Ando, H., Bielinska, M., Pixley, M. R., Shikone, T., Hsueh, A. J., and Boime, I. (1994) *J Biol Chem* **269,** 25543-25548

25. Ruddon, R. W., Sherman, S. A., and Bedows, E. (1996) *Protein Sci* **5,** 1443-1452

26. Fox, K. M., Dias, J. A., and Van Roey, P. (2001) *Mol Endocrinol* **15,** 378-389

27. Erbel, P. J., Karimi-Nejad, Y., De Beer, T., Boelens, R., Kamerling, J. P., and Vliegenthart, J. F. (1999) *Eur J Biochem* **260,** 490-498

28. Huang, J., Chen, F., and Puett, D. (1993) *J Biol Chem* **268,** 9311-9315

29. Campbell, R. K., Dean-Emig, D. M., and Moyle, W. R. (1991) *Proc Natl Acad Sci U S A* **88,** 760-764

30. Grossmann, M., Szkudlinski, M. W., Wong, R., Dias, J. A., Ji, T. H., and Weintraub, B. D. (1997) *J Biol Chem* **272,** 15532-15540

31. Chen, F., Wang, Y., and Puett, D. (1992) *Mol Endocrinol* **6,** 914-919.

32. Birken, S., Gawinowicz Kolks, M. A., Amr, S., Nisula, B., and Puett, D. (1986) *J Biol Chem* **261,** 10719-10727

33. Matzuk, M. M., and Boime, I. (1989) *Biol Reprod* **40,** 48-53

34. Sairam, M. R., and Bhargavi, G. N. (1985) *Science* **229,** 65-67

35. Ji, I. H., and Ji, T. H. (1991) *J Biol Chem* **266,** 13076-13079

36. Ji, I., Zeng, H., and Ji, T. H. (1993) *J Biol Chem* **268**(31)**,** 22971-22974

37. Jiang, X., Dreano, M., Buckler, D. R., Cheng, S., Ythier, A., Wu, H., Hendrickson, W. A., and el Tayar, N. (1995) *Structure* **3,** 1341-1353

38. Ji, I., and Ji, T. H. (1993) *J Biol Chem* **268,** 20851-20854

39. Xie, Y. B., Wang, H., and Segaloff, D. L. (1990) *J Biol Chem* **265,** 21411-21414

40. Tsai-Morris, C. H., Buczko, E., Wang, W., and Dufau, M. L. (1990) *J Biol Chem* **265,** 19385-19388

41. Zhang, R., Buczko, E., and Dufau, M. L. (1996) *J Biol Chem* **271,** 5755-5760

42. Kosugi, S., Ban, T., Akamizu, T., and Kohn, L. D. (1992) *Biochem Biophys Res Commun* **189,** 1754-1762

43. Thomas, D., Rozell, T. G., Liu, X., and Segaloff, D. L. (1996) *Mol Endocrinol* **10,** 760-768

44. Bhowmick, N., Huang, J., Puett, D., Isaacs, N. W., and Lapthorn, A. J. (1996) *Mol Endocrinol* **10,** 1147-1159

45. Moyle, W. R., Campbell, R. K., Rao, S. N., Ayad, N. G., Bernard, M. P., Han, Y., and Wang, Y. (1995) *J Biol Chem* **270,** 20020-20031

46. Kobe, B., and Deisenhofer, J. (1993) *Nature* **366,** 751-756

47. Kobe, B., and Deisenhofer, J. (1995) *Nature* **374,** 183-186

48. Kajava, A. V. (1998) *J Mol Biol* **277,** 519-527

49. Bhowmick, N., Narayan, P., and Puett, D. (1999) *Endocrinology* **140,** 4558-4563

50. Roche, P. C., Ryan, R. J., and McCormick, D. J. (1992) *Endocrinology* **131,** 268-274

51. Hong, S., Phang, T., Ji, I., and Ji, T. H. (1998) *J Biol Chem* **273,** 13835-13840

52. Song, Y. S., Ji, I., Beauchamp, J., Isaacs, N. W., and Ji, T. H. (2001) *J Biol Chem* **276,** 3426-3435

53. Song, Y. S., Ji, I., Beauchamp, J., Isaacs, N. W., and Ji, T. H. (2001) *J. Biol. Chem.* **276,** 3436-3442

54. Alvarez, C. A., Narayan, P., Huang, J., and Puett, D. (1999) *Endocrinology* **140,** 1775-1782

55. Nishi, S., Nakabayashi, K., Kobilka, B., and Hsueh, A. J. (2002) *J Biol Chem* **277,** 3958-3964

56. Zeng, H., Phang, T., Song, Y. S., Ji, I., and Ji, T. H. (2001) *J Biol Chem* **276,** 3451-3458

57. Gether, U. (2000) *Endocr Rev* **21,** 90-113

58. Kosugi, S., Mori, T., and Shenker, A. (1996) *J Biol Chem* **271,** 31813-31817

59. Angelova, K., Narayan, P., Simon, J. P., and Puett, D. (2000) *Mol Endocrinol* **14,** 459-471

60. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., and Miyano, M. (2000) *Science* **289,** 739-745

61. Fanelli, F. (2000) *J Mol Biol* **296,** 1333-1351

62. Fanelli, F., Themmen, A. P., and Puett, D. (2001) *IUBMB Life* **51,** 149-155

63. Angelova, K., Fanelli, F., and Puett, D. (2002) *J Biol Chem* **17,** 17

64. Abell, A. N., and Segaloff, D. L. (1997) *J Biol Chem* **272,** 14586-14591

65. Zhang, M., Tong, K. P., Fremont, V., Chen, J., Narayan, P., Puett, D., Weintraub, B. D., and Szkudlinski, M. W. (2000) *Endocrinology* **141,** 3514-3517

66. Jeoung, M., Phang, T., Song, Y. S., Ji, I., and Ji, T. H. (2001) *J. Biol. Chem.* **276,** 3443-3450

67. Nakabayashi, K., Kudo, M., Kobilka, B., and Hsueh, A. J. (2000) *J Biol Chem* **275,** 30264-30271

68. Zhang, F. P., Rannikko, A. S., Manna, P. R., Fraser, H. M., and Huhtaniemi, I. T. (1997) *Endocrinology* **138,** 2481-2490

69. Zhang, F. P., Kero, J., and Huhtaniemi, I. (1998) *Mol Cell Endocrinol* **142,** 165-174

70. Remy, J. J., Nespoulous, C., Grosclaude, J., Grebert, D., Couture, L., Pajot, E., and Salesse, R. (2001) *J Biol Chem* **276,** 1681-1687

71. Fralish, G. B., Narayan, P., and Puett, D. (2001) *Endocrinology* **142**(4)**,** 1517-1524

72. Schmidt, A., MacColl, R., Lindau-Shepard, B., Buckler, D. R., and Dias, J. A. (2001) *J Biol Chem* **276**(26)**,** 23373-23381

73. Narayan, P., Wu, C., and Puett, D. (1995) *Mol Endocrinol* **9,** 1720-1726

74. Sugahara, T., Pixley, M. R., Minami, S., Perlas, E., Ben-Menahem, D., Hsueh, A. J., and Boime, I. (1995) *Proc Natl Acad Sci U S A* **92,** 2041-2045

75. Sato, A., Perlas, E., Ben-Menahem, D., Kudo, M., Pixley, M. R., Furuhashi, M., Hsueh, A. J., and Boime, I. (1997) *J Biol Chem* **272,** 18098-18103

76. Ben-Menahem, D., Kudo, M., Pixley, M. R., Sato, A., Suganuma, N., Perlas, E., Hsueh, A. J., and Boime, I. (1997) *J Biol Chem* **272,** 6827-6830

77. Jackson, A. M., Berger, P., Pixley, M., Klein, C., Hsueh, A. J., and Boime, I. (1999) *Mol Endocrinol* **13,** 2175-2188

78. Narayan, P., Gray, J., and Puett, D. (2000) *Endocrinology* **141,** 67-71

79. Sen Gupta, C., and Dighe, R. R. (2000) *J Mol Endocrinol* **24,** 157-164

80. Ben-Menahem, D., Jablonka-Shariff, A., Hyde, R. K., Pixley, M. R., Srivastava, S., Berger, P., and Boime, I. (2001) *J Biol Chem* **276,** 29871-29879

81. Kanda, M., Jablonka-Shariff, A., Sato, A., Pixley, M. R., Bos, E., Hiro'oka, T., Ben-Menahem, D., and Boime, I. (1999) *Mol Endocrinol* **13,** 1873-1881

82. Puett, D., and Narayan, P. (2002) *Indian Journal of Experimental Biology* **40,** 415-423

83. Garcia-Campayo, V., and Boime, I. (2001) *Trends Endocrinol Metab* **12,** 72-77

84. Wu, C., Narayan, P., and Puett, D. (1996) *J Biol Chem* **271,** 31638-31642

85. Narayan, P., Gray, J., and Puett, D. (2002) *Mol Endocrinol,* in press.

CHAPTER 2

HIGH-LEVEL EXPRESSION OF A FUNCTIONAL SINGLE CHAIN HCG-LH RECEPTOR ECTODOMAIN COMPLEX IN INSECT CELLS[1]

[1] Fralish, G. B., Prema Narayan, and David Puett (2001) *Endocrinology* **142**(4)**,** 1517-24. Reprinted here with permission of publisher, ©The Endocrine Society.

Abstract

Reproductive capacity in primates is dependent on the high affinity binding of the glycoprotein hormones LH and hCG to the large ectodomain (ECD) of their common receptor (LHR). Our understanding of the precise molecular determinants of hormone binding is limited, since there are no structural data for any of the glycoprotein hormone receptors. Over-expression of the ECD of the receptor has been attempted in various expression systems. Prokaryotic expression does not yield properly folded ECD. Eukaryotic expression, on the other hand, results in mostly heterogeneous, intracellularly trapped protein, but the secreted ECD is completely folded. Accordingly, we have tethered the single chain hormone, yoked hCG (YhCG), to the N-terminus of LHR-ECD (YECD). YhCG is secreted at high levels, binds LHR with high affinity, and, when tethered to the N-terminus of full length LHR, it binds and constitutively activates the receptor. Using recombinant baculovirus, YECD is secreted from insect cells at levels greater than 1 μg/ml, nearly 20-fold higher than that previously reported in eukaryotic expression systems. The protein was purified and binds exogenous ^{125}I-hCG with high affinity, but, significantly, only after protease treatment to remove the tethered hormone. Thus, the fusion protein appears to form a functional hormone-receptor complex that is expressed at levels sufficient for its biophysical characterization.

Introduction

Normal reproductive function in primates and equids is dependent upon the actions of three gonadotropins, pituitary-derived LH and FSH, and the placental-derived hCG. Along with TSH, these molecules form the family of glycoprotein hormones that share a similar heterodimeric architecture of a common α-subunit and a hormone-specific β-subunit. Their biological signals are transduced at the target tissue via their respective G protein-coupled receptors (1). In the case of LH and hCG, the same receptor (LHR) is utilized. This is not surprising given the high sequence similarity (85%) between these two hormones, with the major difference being ascribed to the 30 amino acid residue C-terminal extension of hCG (C-terminal peptide, CTP)(1), whose, four sites of O-glycosylation are important for the increased circulatory half-life of hCG (2). The crystal structure of hCG has been determined (3, 4), and it has provided a framework on which to interpret the numerous studies involving mutagenesis, peptides and chimeras and has aided in our understanding of hormone binding and signal transfer. However, there are no structural data for any of the glycoprotein hormone receptors.

LHR binds hCG via its large, soluble extracellular domain (ECD) (341 amino acids in rat LHR). This glycosylated domain is apparently responsible for most if not all of the contacts necessary to form a specific and high affinity interaction with hCG (K_d = 0.2 nM) (5). The ECD of LHR and the other glycoprotein hormone receptors (FSHR and TSHR) have an imperfect but discernable leucine-rich repeat (LRR) motif (6). The structure of the inhibitor of ribonuclease A, which consists entirely of LRR's, revealed the extended and nonglobular fold of the motif, whose tandem α/β hairpin repeats form a cusp (7). This motif is utilized frequently in nature for protein-protein interactions and

provides an extended binding surface for numerous contact sites with ligand (8). Accordingly, several labs have proposed homology models of the LHR-ECD, based on the ribonuclease A inhibitor structure (6, 9, 10). While significant, the models cannot substitute for direct biophysical data of the receptor or receptor-hormone complex. Indeed, a crystal structure of the ECD would profoundly enhance our understanding of hormone binding and signal transfer; moreover, a structure of the ECD-ligand complex would solidify or reform current paradigms.

LHR presents major hurdles for its resolution at a molecular level. In being an integral membrane protein, it is difficult to express and purify in significant amounts. Therefore, several groups have attempted expression of the soluble LHR-ECD, as well as the ECD's for TSHR and FSHR (Table 2.1). Limited data of significant production of the ECD's for FSHR and TSHR have been reported. The highest expression level of ECD was reported for insect cell-expressed porcine (p) LHR-ECD at approximately 100 μg/ml, while the highest level of secreted ectodomain for the same receptor was 70 ng/ml in mammalian cells. In eukaryotic expression systems, secreted ectodomain displays nearly homogeneous binding activity and has a more complex set of glycans, suggesting a rigorous folding requirement from the cellular protein folding components in the ER and golgi (11). Moreover, there are no published reports of the generation of significant quantities of ligand-receptor complexes for any of the glycoprotein hormone receptors for biophysical characterization.

To improve the production of this challenging glycoprotein, our laboratory has cloned a fusion protein consisting of a single chain yoked hCG (YhCG), N-β-α-C, linked to the N-terminus of the rat (r) LHR-ECD(12). Bioactive YhCG is secreted at high

levels, and, when fused to the full-length receptor, forms a complex with the receptor (yoked hormone receptor; YHR) that promotes signal transduction and inhibits binding of exogenous ^{125}I-hCG (12, 13). Significant binding of ^{125}I-hCG can be recovered only after treatment with Factor Xa protease that cleaves at the engineered site between the hormone and receptor and allows for the dissociation of YhCG. Therefore, if a yoked hormone-extracellular domain (YECD) folded similarly as when connected to the full-length receptor, then it would represent a functional hormone-receptor complex. Accordingly, with future scale-up production of the protein in mind, baculovirus-promoted expression in insect cells was utilized for the characterization of YECD as opposed to mammalian cells, which were used exclusively in previous studies on YHR(13).

Herein we describe the cloning, expression, characterization and purification of YECD. When YhCG is fused to the ECD, the protein is secreted at levels well over 1 μg/ml, i.e. 20-fold greater than secreted levels reported previously (Table 2.1), indicating a practical expression level for scale-up production and biophysical characterization. The protein can be purified to a high degree in one step using an affinity tag. Significantly, the purified protein can bind ^{125}I-hCG with high affinity, but only after dissociation of YhCG is facilitated by digestion with Factor Xa protease, suggesting a functional hormone-receptor association. This approach of fusing ligand to its cognate receptor to increase expression and form a biologically significant ligand-receptor complex is, in principle, applicable to the other glycoprotein hormone receptors and to other receptor systems that involve protein-protein or peptide-protein interactions.

Materials and Methods

Cloning and generation of recombinant virus for epitope tagged YECD

The cDNA for a His tagged (HT) YECD was constructed using the yoked hormone receptor cDNA (YHR), i.e. N-β-α-LHR-C, as a template. PCR was used to amplify a product containing the entire 341 amino acid residues of the ECD of rLHR fused to YhCG via residues 116-145 of the CTP of the β-subunit, with a six-histidine tag on the C-terminus. The product was then cloned into the transfer vector pVL1393 (Pharmingen). In order to generate a second tagged product, the HT in the initial construct was replaced with a flag tag (N-DYKDDDDK-C) using PCR with the following primers: 5'-ATGGAGATATTCCAGGGGCTGCTG-3', 5'-TTACTTATCGTCATCGTCCTTGTAGTCCCTAAGGAAGGCATAGCC-3'. The product was directly cloned into the pCR2.1-TOPO vector (Invitrogen) and subcloned into pVL1393. Recombinant baculoviruses, coding for YECD-HT and for YECD-flag, were generated as previously described for YhCG (12). The virus was plaque purified and amplified, with viral titers of greater than 1 x 10^8 plaque forming units/ml being routinely obtained.

YECD-HT and YECD-flag expression

Plaque purified and amplified recombinant virus was used to infect Sf9 cells, both in monolayer and in suspension cultures. For initial expression screening, monolayer infections were performed. Approximately 2.5 x 10^6 cells were seeded on 60 mm dishes, and virus was added with a multiplicity of infection (MOI) of 1.0 and incubated at 27°C for 1 h in a reduced volume (1 ml) followed by addition of medium to 4 ml. Medium was

then collected after 72 h and, following centrifugation, the amount of YECD was quantified using the β-hCG RIA (ICN), which recognizes free hCG-β and the heterodimer. For expression optimization and purification, cells were adapted in suspension cultures per manufacturer's instructions (Gibco BRL) and infected at a density of 8 x 10^5 cells/ml with an MOI of 0.1, 1.0, or 10.0. Medium was collected, centrifuged and assayed using the β-hCG RIA.

Polyacrylamide gel analysis

Samples were resolved on 10% polyacrylamide gels under reducing conditions. Gels used for Western analysis were transferred to PVDF membrane (Immobilon P). The membrane was probed for the CTP with the rabbit polyclonal antiserum raised against a peptide corresponding to amino acid residues 109-145 of the β-subunit of hCG. The flag tag (Sigma) and the HT (Invitrogen) were detected using commercially available monoclonal antibodies per manufacturer's instructions, and all blots were visualized by the ECL method (Amersham Pharmacia). For total protein detection, gels were silver stained (Bio-Rad).

Purification of YECD-HT and YECD-flag

For purification of YECD-HT and YECD-flag, one-liter Sf9 cell suspension cultures were clarified by centrifugation at 2000 x g for 20 min, and the supernatant was filtered (0.45 μM). For YECD-HT, the expression medium was concentrated 10-fold using an Amicon ultrafiltration cell fitted with a 30 kDa MW cutoff membrane (Millipore). The concentrated medium was then dialyzed at 4°C overnight against "start buffer" (20 mM

sodium phosphate, pH 7.9, 0.5 M NaCl, 10 mM imidazole). The dialyzed sample was incubated with pre-equilibrated (start buffer) superflow Ni^{2+}-NTA resin (Qiagen) overnight at 4°C and then packed into a column for FPLC. After an initial wash with 10 column volumes of start buffer, a linear gradient of 10-200 mM imidazole was applied to the column. Elution of YECD-HT was followed using the β-hCG assay. For purification of YECD-flag, clarified and filtered expression medium was adjusted to a pH of 7.5 and then loaded on a column (Amersham Pharmacia C10/10 column) packed with M2-flag Ab resin (Sigma) that was pre-equilibrated with TBSa (50 mM Tris-HCl, pH 7.4, 0.15 M NaCl) at 0.5 ml/min (4°C). The column was then washed extensively with TBSa (500 ml per inch of gel) at 1 ml/min. The protein was eluted with 0.1 M glycine, pH 3.5, and the eluted sample was immediately collected in tubes containing 1M Tris-HCl, pH 8.0.

Factor Xa digestion of YECD-HT and YECD-flag

Purified samples of YECD were concentrated and dialyzed against Factor Xa buffer (20 mM Tris-HCl, pH 8.0, 0.15 M NaCl, 2 mM $CaCl_2$) overnight at 4°C. The purified, concentrated and dialyzed sample was quantified using the β-hCG RIA. YECD was then digested with Factor Xa enzyme (New England BioLabs) in a 1:25 enzyme to substrate mass ratio for 4-6 h at room temperature. Undigested samples were incubated at room temperature without the addition of enzyme.

Binding of ^{125}I-hCG to YECD-HT and YECD-flag

Purified and protease-treated samples (50-100 ng) were blotted onto PVDF membrane (Millipore) using the Biodot Dot blotting manifold (Bio-Rad). The membrane was then

blocked in TBSb (20 mM Tris-HCl, pH 7.4, 0.5 M NaCl) with 0.2% Tween 20, 3% BSA for 4 h at room temperature. Individual squares (6x6 mm) were cut from the loaded and blocked membrane and added to tubes containing blocking solution with increasing amounts of label (^{125}I-hCG) in the presence or absence of excess unlabeled hCG. The membranes were shaken overnight at room temperature, then washed 4x in TBSb containing 1% NP-40, followed by two washes in TBSb. The squares were then air dried, counted for γ-radiation, and exposed to film overnight at -70°C with Cronex intensifying screens (Kodak). Data were analyzed by non-linear regression using the Prism software (GraphPad Prism, San Diego, CA).

Results

Using YHR as a template, PCR was used to generate the YECD construct. Targeting to the ER is directed by the β-subunit's signal sequence, and the CTP of hCG-β was used as a linker between the C-terminus of the α-subunit and the ECD. The Factor Xa recognition site was retained in YECD, and either a 6-histidine or flag tag was added to the C-terminus to generate the construct shown in Figure 2.1A. High titer recombinant baculovirus was used to infect Sf9 cells, and expression was detected by RIA and Western blot analysis. The RIA recognizes the monomeric or heterodimeric form of the β-subunit of hCG, and it was a useful assay for the rapid analysis of the medium for YECD expression. However, the assay assumes an equal affinity of the antiserum for the tracer antigen, the standards and the unknown (YECD). Since YECD may represent a functional hormone-receptor complex, the affinity of the antiserum for this protein may be different given suggestions of conformational differences in the hormone when bound

to receptor (14, 15). Furthermore, epitopes important for binding of the antiserum to β-hCG may be masked in YECD, and the glycosyl chains present on an insect-expressed YECD are likely to be quite different than those of the human-derived standards. Despite these limitations, positive recognition by the assay corresponded reproducibly with the presence of YECD on Westerns, and dilutions within a given assay gave consistent values for the complex. Western blot analysis of the expression medium was performed using the polyclonal rabbit CTP antibody and the mouse-derived monoclonals specific to the particular C-terminal tags, 6-histidine and flag (Fig. 2.1B). Medium taken from cells infected with YECD-(HT or flag), showed a major band with an apparent molecular weight between 83-89 kDa. In the case of YECD-HT, minor bands at 63 kDa and 28 kDa were also present. The exact nature of these proteins has not been investigated, but they probably represent degradation products or minor cross-reacting species in the expression medium. In medium from cells infected with YECD-flag virus, the lower 28 kDa protein was not present.

Expression of YECD was then optimized in suspension cultures of Sf9 cells. Using the RIA, media from suspension cultures of Sf9 cells infected with various MOIs were assayed over a time period of four days (Fig. 2.2, which shows results for YECD-flag only). At 72 h, the amount of protein produced by an infection of 0.1 MOI was equivalent to that induced by the 1.0 MOI infection, and the highest level was secreted after 96 h. At yields exceeding 1 μg/ml, this is nearly 20-fold greater than the highest reported levels of secreted LHR-ECD (Table 2.1).

Before functional analyses of YECD were performed, the protein from the serum-free medium was purified using the C-terminal tags on the different constructs for affinity

chromatography. Overnight batch binding of YECD-HT to the Ni^{2+}-charged NTA resin was employed because only low amounts of protein bound with a single pass through a packed column. Elution of YECD-HT from the Ni^{2+}-NTA column revealed a significant amount of nonspecific binding to the column despite relatively stringent starting conditions (Fig. 2.3A). More highly purified fractions, eluting towards the end of the gradient, were pooled and used in subsequent analyses (Fig. 2.3A, lanes 7, 8 and right panel). Despite good purification, only 50% of the YECD-HT bound to the column, and, in order to achieve higher purity, significant losses were encountered in collecting the most homogenous fractions from the gradient. This resulted in yields of less than 5% of the starting material. For YECD-flag, binding to the antibody-conjugated resin was dramatically more specific than the metal affinity chromatography method utilized with YECD-HT (Fig. 2.3B). Routinely, 60%-70% of the starting material bound to the column, and elution from the column resulted in nearly homogeneous protein. In this case, final yields were much improved over that obtained with YECD-HT. For both proteins, fractions were pooled and concentrated (Fig. 2.3A, B, pooled and concentrated lanes).

The purified and concentrated samples were treated with Factor Xa protease, which cleaves after the sequence, IEGR, present between the hormone and receptor (Fig. 2.4A). Silver stained gels of the digested protein revealed an efficient and specific cleavage (YECD-flag only shown, Fig. 2.4B). The hormone and the receptor components migrate with the same apparent molecular weight of 44 kDa (Fig. 2.4B) and were distinguished using Western blot analyses (Fig. 2.4C). The CTP antibody is specific to the linker regions of the hormone, thus identifying this component of the

protein. Upon digestion, the hormone migrates with an apparent mobility of 42 kDa (Fig. 2.4C, lane 2). While the molecular weight is higher than predicted by the amino acid sequence, the observed value is consistent with the presence of eight putative sites for O-glycosylation and four sites for N-glycosylation. Indeed, the diffuse band is characteristic of extensive glycosylation. The flag antibody recognizes the tag at the C-terminus of the receptor, and, after treatment with protease, the ECD migrated with an apparent molecular weight of 47 kDa (Fig. 2.4C, lane 1). Again the observed mass of the protein is greater than the predicted size, but the receptor contains four putative sites for N-glycosylation that could account for this difference. Analysis with both antibodies revealed that the cleavage specificity of Factor Xa under the conditions used was excellent, and that most of the YECD-flag was cleaved by the enzyme. The difference in the amount of undigested YECD-flag seen with two antibodies can be explained by the presence of two CTP's in the YhCG portion of the complex, thus enhancing the sensitivity of detection when probing with this antibody.

To assess the functionality of YECD, binding of ^{125}I-hCG to the purified protein was tested. Initially the binding assays were performed by receptor binding to polyethylenimine charged (GF/B) filters (16) or by precipitation with polyethylene glycol (17). Despite obtaining positive results for binding using these methods, the background or nonspecific counts (counts obtained in the presence of an excess of unlabeled hCG) were inconsistent and high. Therefore, a method commonly used for ligand blotting was modified (18). The purified protein was dot-blotted on PVDF membranes and then incubated in binding solution containing the appropriate amounts of labeled and unlabeled hCG. These blots were incubated in the presence of relatively high

concentrations of ^{125}I-hCG (0.5-1 nM) overnight at room temperature, conditions chosen to favor the dissociation of the digested, yoked hormone from the receptor and allow for association of ^{125}I-hCG.

Indeed, using capture with PVDF and these binding conditions, background binding was dramatically reduced, and the specific binding of ^{125}I-hCG to YECD could be readily assessed (Fig. 2.5A). In the presence of 1 nM ^{125}I-hCG, YECD bound specifically to the labeled hormone. The specific radioactivity bound to the membrane was increased considerably when the YECD was digested with factor Xa (Fig. 2.5A, B). In the presence of lower concentrations of label (<0.5 nM), there was no detectable significant specific binding of ^{125}I-hCG to membranes blotted with undigested YECD (Fig. 2.5C). However, YECD that had been digested with factor Xa bound ^{125}I-hCG efficiently. The protein complex was analyzed further by saturation binding experiments. The digested YECD bound with an affinity (K_d = 1-2 nM) somewhat lower or comparable to full-length LHR and previously measured affinities of the ECD (Fig. 2.6) (5, 17, 19-21). Although, the calculated B_{max} was surprisingly low (1-2 pg ^{125}I-hCG/ng of YECD), the data support the idea that the cleaved yoked hormone can still interact productively with the ECD and reduce the total binding of labeled hormone.

Discussion

The binding of LH and hCG to the ECD of their common receptor is essential for male sexual differentiation and normal reproduction. Despite its importance, our understanding of this event is based on indirect observations and models. With no biophysical data available for any of the glycoprotein hormone receptors, many

laboratories are attempting to produce the ECDs of the receptors at levels sufficient for structural studies (Table 2.1). Herein we show that a fusion of single chain ligand (YhCG) and receptor (ECD) via the innocuous CTP as a linker results in the secretion of immunoreactive protein (YECD) at levels nearly 20-fold greater than previously reported (22). By utilizing a flag tag versus the more standard 6-histidine tag, affinity purification of YECD with the monoclonal Ab-conjugated resin resulted in significantly higher purity and final yields compared to Ni^{2+} affinity chromatography.

The purified YECD (both –HT and –flag) binds ^{125}I- hCG with high affinity but only after the treatment with Factor Xa. This important observation suggests that hormone and receptor ECD are able to fold successfully as independent domains while in a tethered context, and moreover they are able to associate in a productive manner to form a functional hormone-receptor complex. This interpretation is supported by previous studies in which expression of YhCG fused to the full-length receptor (YHR) resulted in a constitutively active complex (13). Binding of exogenously added hormone is blocked by the tethered hormone but can be recovered by treatment of YHR with Factor Xa. Further support was provided when the porcine LHR-ECD and transmembrane domains and subunits of hCG were coexpressed as separate polypeptide chains in mammalian cells (23). The proteins were able to coalesce into an active receptor, indicating the ability of the proteins to fold independently while in an overexpressed environment and presumably in close proximity. Intracellular association of hCG with the ECD was not determined.

The lowest MOI used (0.1) produced the most secreted protein. This is in accordance with another report involving an intracellular protease where very low MOIs produced considerably more protein at late harvest times (24). At low MOIs, there is an increased cell survival time, which may result in more robust secretion at later time points.

The ECD of LHR, as well as those for FSHR and TSHR, is responsible for most of the interactions necessary for high affinity ligand binding (5). It has been expressed heterologously in *E. coli*, mammalian cell lines and insect cells. The bacterially expressed protein was trapped in a misfolded state in inclusion bodies (25). Refolded ECD was an active tetramer, but the reported expression levels did not indicate a recovery after refolding, a process which can often be laborious and time consuming. Baculovirus-promoted expression in insect cells has been successful in producing significant quantities of ECD, but most of this protein was trapped intracellularly and was incompletely processed causing an inherent heterogeneity (21). It has been observed that secreted LHR-ECD is more homogeneous in binding activity, and this is probably a result of the cell's protein folding machinery in the secretion pathway, which aids in the more complete folding and processing of the protein (11). By tethering the YhCG to the ECD, we have observed dramatic increases in secretion of the ECD to levels near 1.5 μg/ml, indicating a practical expression level for biophysical approaches to structure determination. To a lesser extent, this phenomenon was observed with the FSH receptor in which coexpression of the ECD with individual subunits or with the hormone increased secretion of the FSHR-ECD, although expression values were not reported (26).

The increased secretion of YECD compared to the ECD may be attributed to several factors. First, i*n vivo* during the first trimester of pregnancy and *in vitro* in various expression systems, hCG is expressed at high levels, and this property may contibute to the improved secretion of challenging proteins that are trapped intracellularly, like the LHR-ECD. Second, LRR proteins are hydrophobic and flexible. While these properties enable such structures to form diverse interactions with other proteins (27), they may hinder their efficient folding and secretion in overexpression systems. With the ligand tethered, the flexibility of the ECD may be reduced and its inherent hydrophobicity may be further buried at the interface of the hormone and receptor. This intramolecular interaction may facilitate folding and processing of the ECD by stabilization of the protein intracellularly and reduce the requirement for host chaperone machinery, which can be limiting to protein production in baculovirus-promoted expression in insect cells (28). Third, the yoked hormone may stabilize the receptor. YECD can retain binding activity for up to three months after preparation when stored at 4°C. Secretion in insect cells of the very low density lipoprotein receptor ECD was not observed until coexpressed with the receptor associated protein, which is known to associate with high affinity (K_d = 0.7 nM) to the lipoprotein receptor (29). This reinforces the concept of associating proteins increasing the stability and expression of their partners and enhancing progression through the secretory pathway.

This novel approach of fusing ligand to receptor will be useful in generating yoked hormone-ECD complexes for the other glycoprotein hormones and may improve expression levels of these receptors as well. The availability of such fusion proteins will facilitate the determination of the structures of these hormone-receptor complexes and

provide an appreciation for the precise determinants for hormone selectivity in this family of related hormones and receptors. In expressing the ECDs in tandem with their cognate ligands, the need for co-crystallization of two large glycoproteins is bypassed, which may facilitate the process of deciphering their structures. Also, in the case of hCG (3, 4) or FSH (30), having the ligand in association with the receptor may permit the use of molecular replacement for phasing of the diffraction data, thus avoiding the need for heavy atom isomorphous replacement. Furthermore, this would eliminate the need for Se-Met incorporation, which is inefficient in insect cells (31).

In summary, for the first time a ligand has been fused to the ECD of a G protein-coupled receptor to promote the expression and secretion of the protein and to form a functional complex. Secretion of the protein in insect cells was improved to levels near 1.5 μg/ml, and better purification was achieved using antibody affinity, rather than Ni^{2+} affinity chromatography. This approach could be extended to the other glycoprotein hormone receptors and other receptor systems that involve protein-protein or protein-peptide interactions to improve their production and/or as a source of hormone-receptor complexes for structural studies.

Acknowledgements

We thank Dr. Vernon Stevens for providing the CTP antiserum and Dr. Jan Potempa for his helpful suggestions regarding the binding assay on PVDF. We would also like to thank Dr. Chengbin Wu, who assisted with the initial set-up of the project and suggested the flag tag for purification.

TABLE 2.1. Expression summary for the ectodomains of the glycoprotein hormone receptors LHR, TSHR and FSHR.

Receptor	Expression System	Expression Level (μg/ml)	Cellular location of expressed protein	Comments	Reference
rLHR	*E.coli*	1	Inclusion bodies	Refolded protein, tetrameric form active, low apparent yields.	(25)
pLHR	Recombinant Baculovirus, Sf9 Cells	100	Intracellular	Heterogeneous nature, only 0.04% could bind ligand.	(21)
pLHR	HC11 (mouse mammary epithelial cell line)	0.07	Secreted	Highest amount of secreted ectodomain reported to date.	(22)
hTSHR	*E.coli*	0.01-0.05	Intracellular	Expression level reported as final purified product.	(32)
hTSHR	Recombinant Baculovirus, HiFive cells	NR	Intracellular	Reported milligram quantities of purified ectodomain.	(33)
hTSHR	Vaccinia Virus/HeLa cells	0.1-0.2	Intracellular	Purified to 70% homogeneity.	(34)
hFSHR	Recombinant Baculovirus, HiFive cells	NR	Intracellular and Secreted	Expression level not reported, protein was not purified	(26)

NR, not reported; r, rat; p, porcine; h, human.

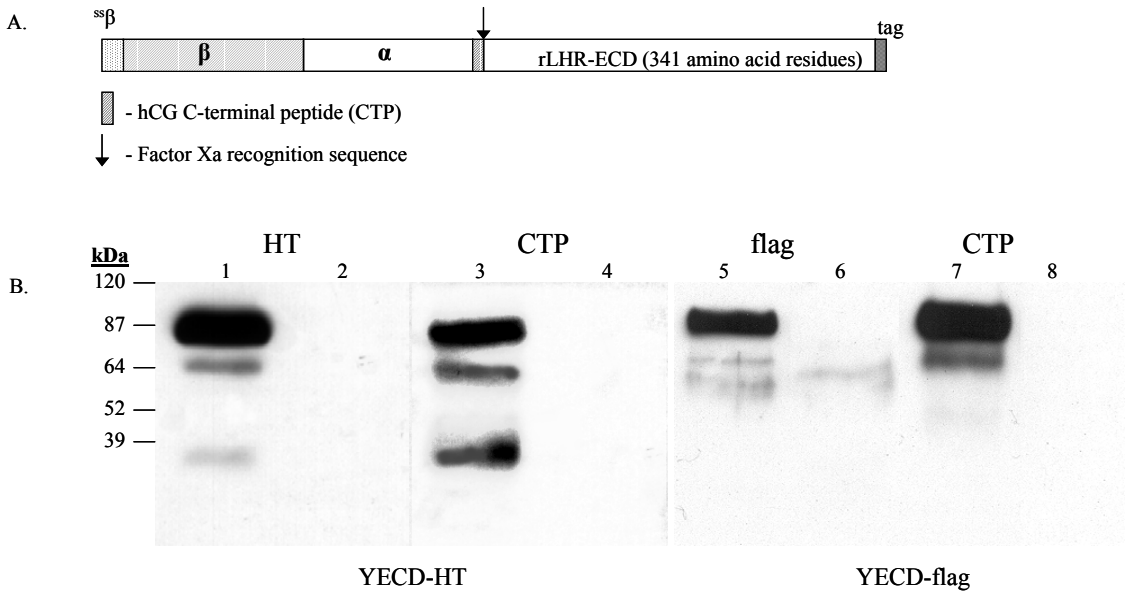

FIG. 2.1. Secretion of YECD by insect cells. A. Schematic representation of YECD. YhCG is tethered to the N-terminus of the ECD of the receptor via amino acid residues 116-145 comprising the CTP of β-hCG. The Factor Xa recognition sequence is present between the hormone and receptor, and the C-terminus of the construct contains either a six-histidine or flag tag. The β-subunit's signal sequence (^{ss}β) is utilized for targeting the protein for secretion. B. Expression of YECD-HT and YECD–flag was detected by Western blot analysis using antibodies against the CTP and the two C-terminal tags. All lanes were loaded with 50 μg total protein. Odd number lanes contain samples collected from the medium of cells infected with the indicated YECD virus. Even number lanes contain medium from cells infected with a control virus expressing the *XylE* protein. The sizes of the molecular weight standards are indicated.

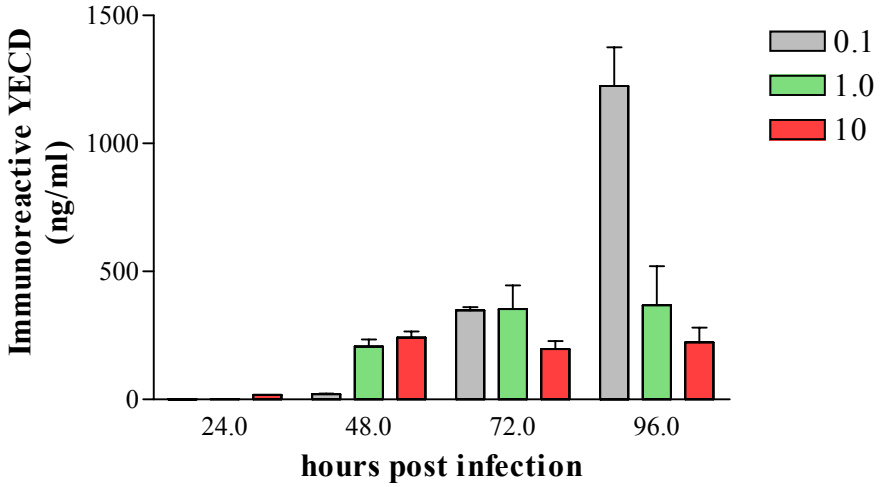

FIG. 2.2. Expression of YECD-flag in insect cells. Sf9 cells were adapted to suspension cultures and infected at various MOIs. Medium was harvested from the cultures every 24 h for four days. The results are the mean ± SEM of two expression surveys assayed by the β-hCG RIA.

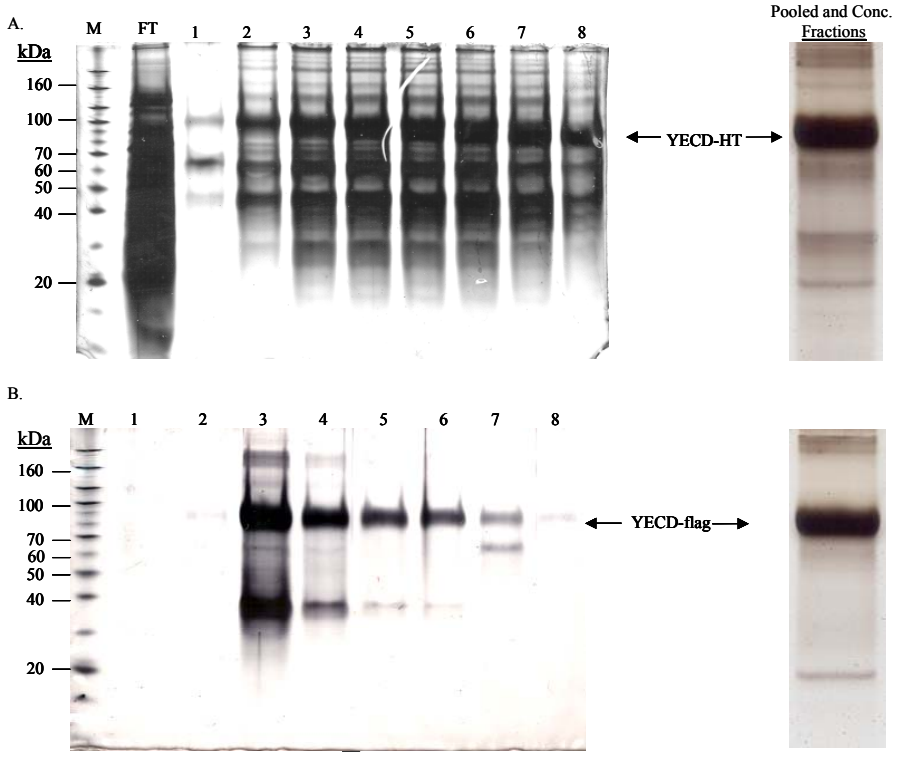

FIG.2.3. Purification of YECD. A. SDS-PAGE of fractions eluted from Ni^{2+}-charged NTA resin loaded with medium from cells infected with YECD-HT virus. FT, 50 μg of total protein was loaded from the flow through fraction; lanes 1-8, 20 μl aliquot from 1 ml fractions along the imidazole gradient. The right panel contains pooled and concentrated (30 kDa MW cutoff) fractions from lanes 7 and 8. B. M2-Ab-conjugated resin purification of YECD-flag from medium from cells infected with YECD-flag virus. Lanes 1-8 of the SDS-PAGE gel represent fractions collected during elution of YECD-flag from the resin with 0.1 M glycine, pH 3.5; equivolume samples (20 μl) were loaded. The right panel is a pool of all fractions from the elution that was concentrated (30 MW kDa cutoff). The gels were silver stained.

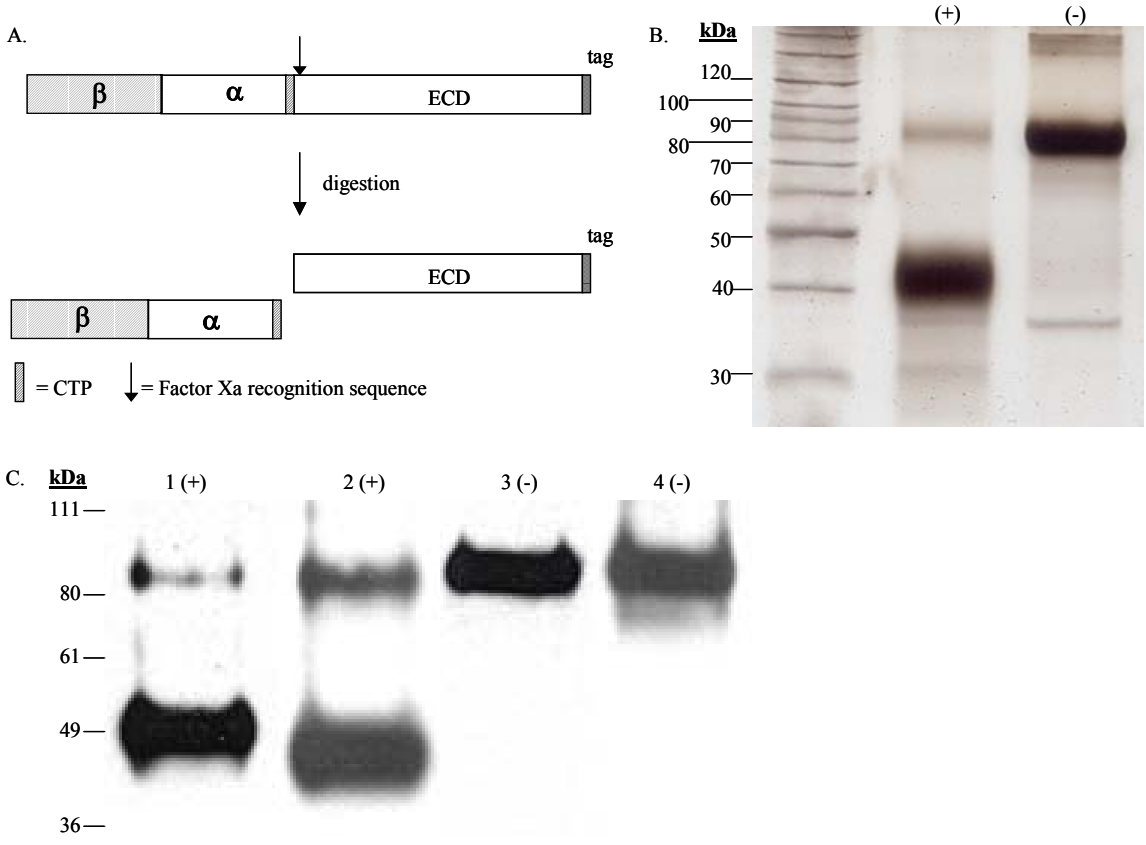

FIG. 2.4. Factor Xa digestion of YECD-flag. A. Schematic representation of digestion of YECD with Factor Xa. Digestion by the protease results in cleavage of the tether between the yoked hormone and ECD. The yoked hormone retains the two CTP's of the construct. B. Silver stained gel comparing the mobility changes of YECD-flag digested with Factor Xa (+) and undigested YECD-flag (-). Approximately 10 ng of YECD-flag was loaded. C. Western blot analysis of the digestion of YECD-flag using the anti-CTP and anti-flag antibodies. All lanes contain 10 ng of digested or undigested sample. Lanes 1 and 3 were probed using the anti-flag antibody, and lanes 2 and 4 were probed using the anti-CTP antibody. (+) = factor Xa treated; (-) = untreated.

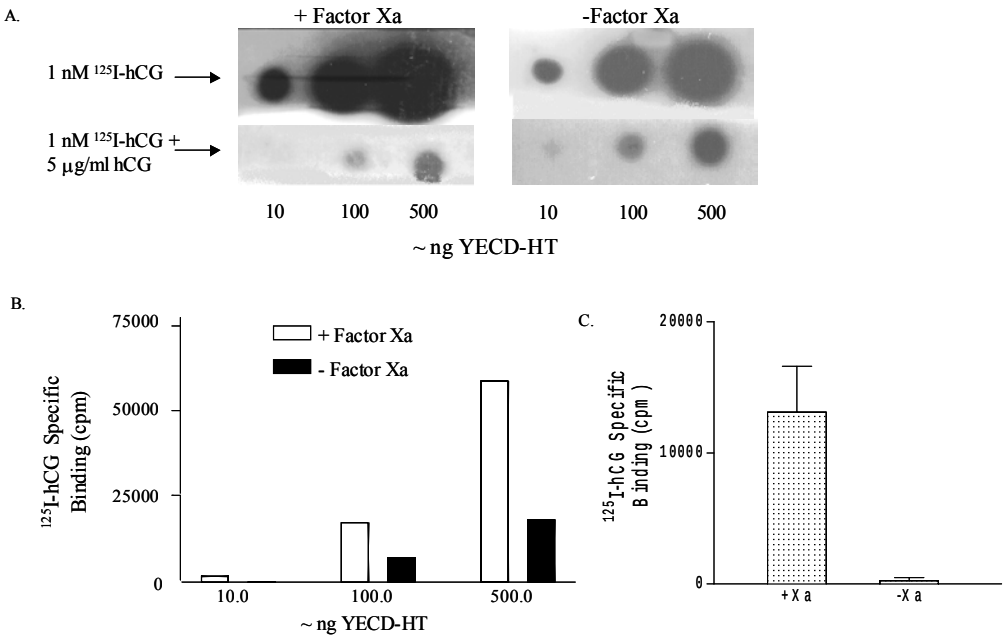

FIG. 2.5. Binding of ^{125}I-hCG to YECD-HT. A. PVDF membranes were blotted with different amounts of Factor Xa-digested and undigested YECD and incubated with solutions containing either 1 nM ^{125}I-hCG alone or 1 nM ^{125}I-hCG plus 5 µg/ml hCG. The membranes were washed and exposed to film overnight at -70°C. B. Squares from these membranes were cut out and counted for γ-radiation. The nonspecific counts (from membranes incubated in the presence of 5 µg/ml hCG) were subtracted from the total. C. Squares incubated with 0.5 nM ^{125}I-hCG alone or 0.5 nM ^{125}I-hCG plus 5 µg/ml hCG were washed, dried and counted. The nonspecific counts were subtracted from the total. The data shown are from a representative experiment. These experiments have been repeated at least 5 times, with independent YECD preparations.

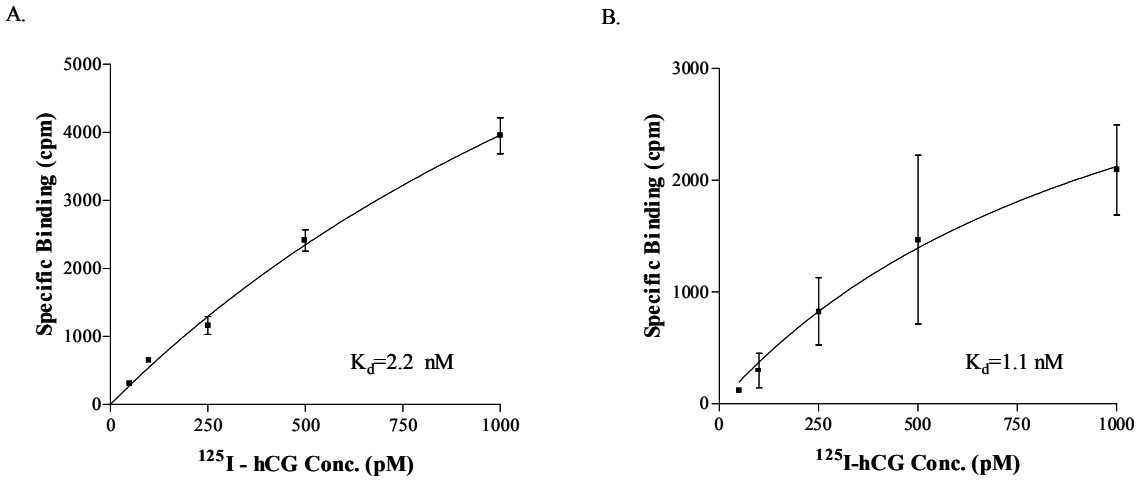

FIG. 2.6. A representative saturation binding experiment with Factor Xa-digested YECD-HT (panel A) or YECD-flag (panel B). Digested YECD protein (50 ng) was blotted to PVDF membrane cut into squares and incubated with increasing amounts of ^{125}I-hCG in the presence or absence of excess unlabeled hCG (5 μg/ml). The squares were washed, dried and counted for γ-radiation. The data shown represent the average ± range of duplicate determinations, and the experiments were repeated at least twice with different protein preparations.

References

1. **Pierce JG, Parsons TF** 1981 Glycoprotein hormones: structure and function. Annu Rev Biochem 50:465-495
2. **Matzuk MM, Hsueh AJ, Lapolt P, Tsafriri A, Keene JL, Boime I** 1990 The biological role of the carboxyl-terminal extension of human chorionic gonadotropin [corrected] beta-subunit [published erratum appears in Endocrinology 1990 Apr;126(4):2204]. Endocrinology 126:376-383
3. **Lapthorn AJ, Harris DC, Littlejohn A, Lustbader JW, Canfield RE, Machin KJ, Morgan FJ, Isaacs NW** 1994 Crystal structure of human chorionic gonadotropin [see comments]. Nature 369:455-461
4. **Wu H, Lustbader JW, Liu Y, Canfield RE, Hendrickson WA** 1994 Structure of human chorionic gonadotropin at 2.6 A resolution from MAD analysis of the selenomethionyl protein. Structure 2:545-558
5. **Xie YB, Wang H, Segaloff DL** 1990 Extracellular domain of lutropin/choriogonadotropin receptor expressed in transfected cells binds choriogonadotropin with high affinity. J Biol Chem 265:21411-21414
6. **Jiang X, Dreano M, Buckler DR, Cheng S, Ythier A, Wu H, Hendrickson WA, el Tayar N** 1995 Structural predictions for the ligand-binding region of glycoprotein hormone receptors and the nature of hormone-receptor interactions. Structure 3:1341-1353
7. **Kobe B, Deisenhofer J** 1993 Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. Nature 366:751-756

8. **Kobe B, Deisenhofer J** 1994 The leucine-rich repeat: a versatile binding motif. Trends Biochem Sci 19:415-421

9. **Moyle WR, Campbell RK, Rao SN, Ayad NG, Bernard MP, Han Y, Wang Y** 1995 Model of human chorionic gonadotropin and lutropin receptor interaction that explains signal transduction of the glycoprotein hormones. J Biol Chem 270:20020-20031

10. **Bhowmick N, Huang J, Puett D, Isaacs NW, Lapthorn AJ** 1996 Determination of residues important in hormone binding to the extracellular domain of the luteinizing hormone/chorionic gonadotropin receptor by site-directed mutagenesis and modeling. Mol Endocrinol 10:1147-1159

11. **Pajot-Augy E, Bozon V, Remy JJ, Couture L, Salesse R** 1999 Critical relationship between glycosylation of recombinant lutropin receptor ectodomain and its secretion from baculovirus-infected insect cells. Eur J Biochem 260:635-648

12. **Narayan P, Wu C, Puett D** 1995 Functional expression of yoked human chorionic gonadotropin in baculovirus-infected insect cells. Mol Endocrinol 9:1720-1726

13. **Wu C, Narayan P, Puett D** 1996 Protein engineering of a novel constitutively active hormone-receptor complex. J Biol Chem 271:31638-31642

14. **Hong SH, Ji IH, Ji TH** 1999 The beta-subunit of human choriogonadotropin interacts with the exodomain of the luteinizing hormone/choriogonadotropin receptor and changes its interaction with the alpha-subunit. Mol Endocrinol 13:1285-1294

15. **Cosowsky L, Rao SN, Macdonald GJ, Papkoff H, Campbell RK, Moyle WR** 1995 The groove between the alpha- and beta-subunits of hormones with lutropin (LH) activity appears to contact the LH receptor, and its conformation is changed during hormone binding. J Biol Chem 270:20011-20019

16. **Roche PC, Bergert ER, Ryan RJ** 1985 A simple and rapid method using polyethylenimine-treated filters for assay of solubilized LH/hCG receptors. Endocrinology 117:790-792

17. **Hong S, Phang T, Ji I, Ji TH** 1998 The amino-terminal region of the luteinizing hormone/choriogonadotropin receptor contacts both subunits of human choriogonadotropin. I. Mutational analysis. J Biol Chem 273:13835-13840

18. **Thomas DM, Segaloff DL** 1994 Hormone-binding properties and glycosylation pattern of a recombinant form of the extracellular domain of the luteinizing hormone/chorionic gonadotropin receptor expressed in mammalian cells. Endocrinology 135:1902-1912

19. **Angelova K, Narayan P, Simon JP, Puett D** 2000 Functional role of transmembrane helix 7 in the activation of the heptahelical lutropin receptor. Mol Endocrinol 14:459-471

20. **Lobel LI, Rausch P, Trakht I, Pollak S, Lustbader JW** 1997 Filamentous phage displaying the extracellular domain of the hLH/CG receptor bind hCG specifically. Endocrinology 138:1232-1239

21. **Pajot-Augy E, Couture L, Bozon V, Remy JJ, Biache G, Severini M, Huet JC, Pernollet JC, Salesse R** 1995 High-level expression of recombinant porcine

LH receptor in baculovirus- infected insect cells or caterpillars. J Mol Endocrinol 14:51-66

22. **Pajot-Augy E, Attal J, Theron MC, Devinoy E, Fontaine ML, Houdebine LM, Salesse R** 1997 Recombinant expression and secretion of a natural splicing variant containing the ectodomain of porcine LH receptor in HC11 mammary epithelial cells. Protein Expr Purif 10:107-114

23. **Remy JJ, Bozon V, Couture L, Goxe B, Salesse R, Garnier J** 1993 Reconstitution of a high-affinity functional lutropin receptor by coexpression of its extracellular and membrane domains. Biochem Biophys Res Commun 193:1023-1030

24. **Liebman JM, LaSala D, Wang W, Steed PM** 1999 When less is more: enhanced baculovirus production of recombinant proteins at very low multiplicities of infection. Biotechniques 26:36-38, 40, 42

25. **Chen W, Bahl OP** 1993 High expression of the hormone binding active extracellular domain (1-294) of rat lutropin receptor in Escherichia coli. Mol Cell Endocrinol 91:35-41

26. **Stevis PE, Deecher DC, Lopez FJ, Frail DE** 1999 Pharmacological characterization of soluble human FSH receptor extracellular domain: facilitated secretion by coexpression with FSH. Endocrine 10:153-160

27. **Kobe B, Deisenhofer J** 1995 A structural basis of the interactions between leucine-rich repeats and protein ligands. Nature 374:183-186

28. **Ailor E, Betenbaugh MJ** 1999 Modifying secretion and post-translational processing in insect cells. Curr Opin Biotechnol 10:142-145

29. **Sato A, Shimada Y, Herz J, Yamamoto T, Jingami H** 1999 39-kDa receptor-associated protein (RAP) facilitates secretion and ligand binding of extracellular region of very-low-density-lipoprotein receptor: implications for a distinct pathway from low-density-lipoprotein receptor. Biochem J 341:377-383

30. **Dias JA, Fox KM, Van Roey P** 2000 Crystal Structure of Glycosylated Fully Active Human Follicle Stimulating Hormone. Program of the Endocrine's Society 82nd Annual Meeting, Toronto, Ontario, Canada, 2000 p133 (Abstract)

31. **Bellizzi JJ, Widom J, Kemp CW, Clardy J** 1999 Producing selenomethionine-labeled proteins with a baculovirus expression vector system. Structure Fold Des 7:R263-267

32. **Bobovnikova Y, Graves PN, Vlase H, Davies TF** 1997 Characterization of soluble, disulfide bond-stabilized, prokaryotically expressed human thyrotropin receptor ectodomain. Endocrinology 138:588-593

33. **Huang GC, Page MJ, Nicholson LB, Collison KS, McGregor AM, Banga JP** 1993 The thyrotrophin hormone receptor of Graves' disease: overexpression of the extracellular domain in insect cells using recombinant baculovirus, immunoaffinity purification and analysis of autoantibody binding. J Mol Endocrinol 10:127-142

34. **Lee MH, Park JY, Cho BY, Chae CB** 1999 Expression of the functional extracellular domain of human thyrotropin receptor using a vaccinia virus system: its purification and analysis of autoantibody binding. J Clin Endocrinol Metab 84:1391-1397

CHAPTER 3

CIRCULAR DICHROIC SPECTROSCOPY OF YOKED CHORIONIC GONADOTROPIN-LUTROPIN RECEPTOR ECTODOMAIN COMPLEXES[2]

[2]Fralish, G. B., Brian Dattilo, David Puett. Submitted to Molecular Endocrinology, 9/20/2002.

The abbreviations used in this chapter are: LHR, luteinizing hormone (lutropin) receptor; CD, circular dichroism; CG, chorionic gonadotropin (choriogonadotropin); CTP, C-terminal peptide (of hCG-β); ECD, ectodomain; FSH, follicle stimulating hormone (follitropin); GPCR, G protein-coupled receptor; hCG, human CG; hFSH, human FSH; LRR, Leu-rich repeat; TBS, 50 mM Tris-HCl, pH 7.5, 150 mM NaCl; YECD, yoked (single chain) hCG-LHR ECD complex (N-hCGβ-α-CTP-Factor Xa site-LHR ECD-Flag-C); YhCG, yoked (single chain) hCG (N-hCGβ-α-Flag-C); Y341, same as YECD (complete ECD, residues 1-341); Y294, YECD with C-terminal deletion (residues 1-294); Y180, YECD with a C-terminal deletion (residues 1-180).

ABSTRACT

High affinity binding of the glycoprotein hormones, luteinizing hormone (LH) and chorionic gonadotropin (hCG), occurs to the relatively large, N-terminal ectodomain (ECD) of the heptahelical LH receptor (LHR). Given the paucity of structural data on this member of the G protein-coupled receptor superfamily, our understanding of the molecular characteristics of the ECD is limited to models, all of which are based on a leucine-rich repeat motif. We previously reported the expression, purification and characterization of a functional yoked hCG-LHR ECD (YECD) complex, N-hCGβ-α-ECD-C. In this study, we have obtained and analyzed circular dichroic (CD) spectra of several single chain hCG-ECD complexes: YECD (Y341, containing all 341 amino acid residues of the ECD, encoded by exons 1-10 and a portion of 11) and two C-terminally truncated fragments of the ECD, Y294 (residues 1-294 encoded by exons 1- 10) and Y180 (residues 1-180 encoded by exons 1-7). By expressing the full-length and truncated ECDs as fusion proteins with ligand, high levels of protein secretion were achieved. The single chain proteins were expressed, purified and characterized by Western blot analysis and circular dichroic (CD) spectroscopy. Subtraction of the CD contributions of the single chain ligand from the spectrum of YECD revealed a possible β/α fold of the ECD consistent with homology models. The CD spectra of Y294 and Y180, compared to that of full-length Y341, revealed a marked decrease in mean residue ellipticity (and helicity), but, surprisingly, the two spectra were comparable to each other. These results suggest that the C-terminal 47 amino acid segment of the ECD, i.e. residues 295 – 341, stabilizes the structure (helicity) of the hCG-ECD or, less likely, is itself

helical in nature. Furthermore, the results support a repeating fold, consistent with that of a typical leucine-rich repeat protein.

INTRODUCTION

The G protein-coupled luteinizing hormone receptor (LHR[1]) plays a critical role in human reproductive physiology and male sexual differentiation, with important processes such as gonadal steroidogenesis and ovulation regulated by high affinity binding of luteinizing hormone (LH) or chorionic gonadotropin (CG) to LHR (1,2). The two ligands, LH and CG, are members of the family of glycoprotein hormones, which also includes follicle stimulating hormone (FSH) and thyroid stimulating hormone (TSH) (3). These proteins all share a heterodimeric structure, consisting of a common α-subunit and a hormone specific β-subunit. LH and CG exhibit a high degree of homology (> 90% similarity), with most of their dissimilarity arising from a 30 amino acid residue C-terminal peptide (CTP) of hCG-β. Indeed, the two hormones display similar (albeit distinct) binding affinities for LHR (3). The crystal structures of hCG (4,5) and hFSH (6) reveal a common fold for the growth factor-like cystine-knot motif that is probably shared by all of the family members.

The three glycoprotein hormone receptors, i.e. for LH/CG, FSH and TSH, form a subfamily in the group of rhodopsin-like G protein-coupled receptors (GPCR) (1,2). They are characterized by a large N-terminal ectodomain, comprising about half the mass of the receptor and composed of an N-terminal Cys-rich region, a LRR region, and a C-terminal hinge region also rich in cysteine residues, that binds ligand with high affinity in the absence of the transmembrane domain (Fig. 3.1A) (1,2,7). Homology modeling of a large portion of the LHR ECD has been done by our laboratory, in collaboration with

Drs. Isaacs and Lapthorn, and others (Fig. 3.1B) (8-10). The number and delineation of the LRR's in the different models vary, but all are in agreement that the repeats found in the LHR ECD sequence could accommodate the hallmark β/α repeat fold. However, the intrastrand regions of different LRR proteins display diverse structural characters that are dependent upon the length of this region (11). Thus, the precise nature of the fold is unclear. None of the models included the conserved N-terminal Cys-rich region, and only one of the reports suggested a possible fold for a portion of the C-terminal Cys-rich cluster (10). From the modeling studies and subsequent mutagenesis mapping of the ECD, it has become apparent that charge-charge interactions play a critical role in the high affinity binding of hCG to LHR ECD (9,10,12,13), as do hydrophobic interactions (13). Further, peptide competition studies (14,15) suggest a noncontiguous hormone footprint indicative of an extended hormone-receptor interface, while independent LRR and exon truncation (16,17) and deletion (18) studies associate the N-terminal region of the ECD with high affinity binding.

Direct structural data of the ECD for LHR and the other glycoprotein hormone receptors are extremely limited. This is due in large part to the incompatibility of the ECDs with most commonly employed overexpression systems. Recently, two groups have reported high levels of expression of the LHR ECD in bacteria (19) and CHO cells (20). We have successfully expressed and characterized a novel yoked hCG-LHR ECD (1-341) fusion protein (YECD) in insect cells (21). The feasibility of this approach was supported by studies on the previously described yoked hormone, YhCG (22), and the yoked hormone-receptor complex (YHR) (23,24). From studies on YHR it was apparent that hCG fused to the N-terminus of LHR was capable of fully stimulating the receptor

and blocking binding of exogenous hCG. By extension, YECD was cloned, expressed and characterized. Functional protein complex was secreted from insect cells at high levels (1.5 mg/L) permitting unique structural studies on the single chain hCG-LHR ECD (21).

We report herein the structural characterization (using CD spectroscopy) of yoked hCG-LHR ECD (1-341) (Y341, i.e., the number of ECD amino acid residues in the rat construct) and expression, purification and structural characterization of two C-terminally truncated hormone-receptor fusion proteins, yoked hCG-LHR ECD (1-294) (Y294) and yoked hCG-LHR ECD (1-180) (Y180). The CD spectrum for the fused hormone, YhCG, was also obtained and when subtracted from the YECD spectrum yielded a difference spectrum consistent of that of a β/α protein. Estimates of percent helix from the CD spectra of Y294 and Y180 were found to be similar to each other but less than Y341. These results suggest that: (i) the C-terminal 47 amino acid residues of the ECD stabilize the helicity of the entire ECD or possibly, but less likely, are themselves helical in nature, and that (ii) the ECD has a repetitive folding pattern that may extend further into the C-terminal Cys-rich cluster than expected (10). This approach of fusing ligand and receptor has enabled high expression levels of complexes to be achieved and has provided a new approach for obtaining direct structural data for determining the molecular characteristics of GPCR ECDs.

MATERIALS AND METHODS

Cloning and expression of constructs - YhCG was cloned as described previously (22). A Flag tag (N–DYKDDDDK–C) was added to the 3' end of the YhCG sequence

using PCR with primers YhCG1-5' (5'–ATGGAGATATTCCAGGGGCTGCTG–3') and YhCG1-3' (5'-TTACTTATCGTCATCGTCCTTGTAGTCAGATTTGTGATAATAAAC–3'). The PCR product was purified and cloned into the TOPO TA cloning vector (Invitrogen, Carlsbad, CA). Colonies were analyzed by restriction analysis and sequencing and then subcloned into the baculovirus transfer vector pVL1393 (BD Biosciences, Pharmingen, San Diego, CA). Low-titer baculovirus stocks were generated through cotransfection of the pVL1393/YhCG and Baculogold, linearized baculoviral DNA (BD Biosciences, Pharmingen). Medium from the transfected cells was used in serial amplifications to generate high-titer recombinant baculovirus (~10^8 plaque forming units/ml). YECD, containing a histidine tag, was cloned as previously described (21) and used as a template for PCR to generate Flag-tagged Y294 and Y180 cDNAs using methods described above with only the 3' primer sets differing: Y294-3' primer (5'-TTACTTATCGTCATCGTCCTTGTAGTCAAGCGTCTCGTTATCTGC - 3') and Y180-3' (5'–TTACTTATCGTCATCGTCCTTGTAGTCCAGCGAGATTAGAGTC 3'). Thus, the truncated constructs only varied in the length of the ECD, retaining the exact Flag tag, subunit, linker arrangement and Factor Xa recognition sequence. The PCR products were confirmed by restriction enzyme analysis and sequencing, and subcloned from the TOPO TA vector into the pVL1393 transfer vector.

The high-titer recombinant baculovirus stocks for these constructs were produced as described previously (21) and were used to induce the heterologous expression of the proteins in SF9 insect cells. The SF9 cells were grown in suspension in orbital shaker cultures at a density of 8 x 10^5 cells/ml. Expression of the proteins at low multiplicities

of infection (MOI = 0.1) was sufficient for their characterization. Expression was monitored using RIAs specific to the β-subunit (ICN Pharmaceuticals, Costa Mesa, CA) and heterodimeric hCG (Diagnostic Products Corporation, Los Angeles, CA).

Purification of recombinant proteins: Y341, Y294, Y180 and YhCG- Medium from cells infected with the recombinant baculovirus was harvested and clarified with two sequential centrifugations at 1,400g and 17,700g. Ammonium sulfate was slowly added to the clarified medium to 60% saturation (80% saturation for YhCG) at 4˚C and stirred vigorously overnight. The precipitated protein was pelleted by two centrifugations at 17,700g. The pellet was resuspended in TBS and dialyzed extensively at 4˚C against TBS. The resuspended and dialyzed protein sample was cycled at low flow rates (0.5 ml/min) over M2-Flag Ab affinity resin (Sigma-Aldrich, St. Louis, MO) overnight. Binding of protein to the column was followed using RIA of the eluent fractions. The loaded column was washed extensively with TBS, and protein was eluted with 0.1 M glycine (pH 3.5). The column fractions were collected in tubes containing 1 M Tris-HCl, pH 8.0 to immediately neutralize the solutions. Fractions containing the protein were pooled and concentrated for subsequent analyses. Purity of the proteins was assessed using SDS-PAGE and silver staining, and molecular weight estimates were based upon the migration of protein standards (Invitrogen).

Factor Xa digestion and Western blot analysis of Y341, Y294 and Y180- Purified proteins were treated with Factor Xa protease (New England Biolabs, Beverly, MA) as previously described (21). Briefly, the proteins were dialyzed overnight against a buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl and 2 mM $CaCl_2$, then treated with protease (1:25 enzyme:substrate ratio) for 4 h at room temperature. The digestion was

analyzed by SDS-PAGE and Western blot analysis using the M2-flag mouse monoclonal antibody (Sigma), which recognizes the C-terminal flag sequence. An anti-mouse IgG HRP-conjugated secondary antibody was used for detection of the protein using standard ECL reagents (Amersham Pharmacia). A polyclonal rabbit CTP antibody (1:1000 dilution from original stock), kindly provided by Dr. Vernon Stevens, (Ohio State University) was employed to probe for hormone portions of the complexes (23).

Soluble binding assays- The ability of the yoked hCG-ECD proteins to bind exogenously added ^{125}I-hCG was assessed using a method previously detailed (21). Purified, Factor Xa-treated (and untreated) protein (10 ng) was added to PVDF membrane using the Bio Rad Dot Blot apparatus. The membrane was cut into squares and then incubated with 1 nM ^{125}I-hCG, with and without 5 μg/ml unlabeled hCG (kindly provided by Dr. A.F. Parlow at the NIDDK) in TBS containing 0.2% (v/v) Tween 20 and 3% (w/v) bovine serum albumin. The tubes containing the membranes were shaken overnight at room temperature, washed five times with TBS containing 1% NP-40, dried and counted for gamma radiation. Specific counts bound (cpm) were calculated by subtracting counts obtained in the presence of excess unlabeled hCG. The data presented represent the mean ±SEM of three independent experiments.

Circular dichroism- The CD spectra of the purified proteins were recorded using a Jasco 710 CD spectrometer. Before collection of the spectra, the samples were dialyzed against 5 mM phosphate buffer, pH 6.8 overnight. After dialysis, the sample was centrifuged through a 0.1 μm filter (Millipore, Bedford, MA). The protein concentrations were determined by UV absorption spectroscopy from 220-350 nm with the extinction coefficients for the various proteins being estimated at 280 nm from the primary

sequences assuming completely oxidized cystines (25). CD spectra of the proteins were recorded at room temperature using a 1 mm pathlength cell and protein concentrations of 5-10 μM. The settings for collection of the spectra were as follows: band width, 1 nM; sensitivity, 50 mdeg; response, 2 sec; scan speed, 20 nm/min; step resolution, 0.2 nm; starting wavelength, 250 nm; lowest wavelength, 190 nm; and 5 scans/sample. The data presented represent the average of at least three separate spectral recordings with two or more different protein preparations. The Jasco standard analysis program was used to convert units and smooth data from replicate scans. Secondary structural analysis of YhCG and Y341 were done using the SELCON3 (26-28) and the CONTINLL (29) programs based upon reference protein sets of 37 and 42 proteins, respectively. With the Y294 and Y180 spectra, estimates were based upon CONTINLL and the method described by Hennessey and Johnson (30).

RESULTS

The structural nature of the ECD has been elusive primarily because of difficulities encountered when trying to overexpress the protein in heterologous systems. Recently, we reported the purification and characterization of a functional single chain hCG-ECD protein expressed in insect cells. It was observed that when the ligand was fused to the ECD expression levels near 1.5 mg/L were achieved (21). The boost in expression was proposed to be ligand-associated stabilization of the ECD, which may be flexible and unstable. This suggestion was supported by a subsequent report where the porcine LHR ECD was produced in CHO cells at high levels through coexpression with hCG (20).

The C-terminal region of the ECD (amino acid residues 205-341 of rat LHR ECD) has been shown to contribute little to the high affinity binding of hCG (16-18). Thus, we have generated truncated versions of YECD consisting of the first 180 and 294 amino acid residues of the ECD (denoted Y180 and Y294, respectively) with the identical yoked hormone-receptor arrangement of YECD, while retaining the Factor Xa recognition sequence (Fig. 3.1C). Y294 is of interest given its existence as an alternative splicing product (residues 1-297 in porcine LHR and 1-294 rat LHR), whose mRNA is detectable *in vivo* (31,32). Y180 was chosen because it represents the smallest portion of the ECD that can still bind hormone with significant affinity, albeit some 10-fold reduced compared to intact LHR (17). The cDNAs were cloned, recombinant baculoviruses were generated and expression of the proteins in insect cells was optimized. Interestingly, expression levels for the C-terminally truncated proteins were greater that that of the full-length product, e.g. Y294 and Y180 expressed at respective levels of 1.5-2-fold and 4-5 fold higher than that of Y341 (data not shown).

Purification of the various proteins was achieved using ammonium sulfate precipitation and antibody affinity chromatography, and typical protein yields were in the range of 40-60%. Protein eluted from the M2-Flag column was homogeneous as determined by SDS-PAGE and silver staining (Fig. 3.2). Factor Xa-treated and untreated samples were analyzed by Western blot analysis (Fig. 3.3). Using antibodies to the M2-Flag and CTP, the receptor and hormone components of the protein complex can be discerned. Western blotting with these antibodies indicates intact receptor species resolving at molecular weights greater than would be predicted from just the amino acid sequence consistent with the addition of carbohydrate. For Y341, the receptor is more

diffuse in the gel, suggesting a greater amount of glycosylation heterogeneity compared to Y180 and Y294. Y341 contains three more putative N-glycosylation sites than Y180, but the same number as Y294, suggesting there may be a requirement for the C-terminal region of the ECD for complete maturation. Also, when the Factor Xa-treated samples were probed with the CTP antibody the hormone regions of the protein were resolved to equivalent positions on the gel. The efficiency and specificity of the cleavage is excellent with little undigested protein present.

Proteins were tested for binding activity using a method previously described (21). The data indicate that Y294 can bind exogenous ^{125}I-hCG at a level similar to that of the Y341 (Fig. 3.4). Binding to Y341 and Y294 is only detectable if the proteins are first treated with Factor Xa (Fig. 3.4), suggesting that the fused hormone occupies the binding site for hCG. This is not surprising in these fusion proteins since the ligand is constrained by the CTP-Factor Xa cleavage site linker from diffusing further away from the ECD binding site than the length of the linker, ensuring that the single chain gonadotropin is always either bound to the ECD or in close proximity.

Numerous studies using detergent-solubilized cells have shown that the intact ECD of rat and human LHR binds hCG with the same affinity as the full-length receptor, if not somewhat slightly higher affinity (2). Several C-terminal deletion fragments of the ECD have also been characterized with respect to ligand binding (17), as have a number of deletions within the ECD in full-length LHR in intact and in detergent-solubilized cells (18). Pertinent to the ECD N-terminal fragments investigated herein are several reports on the N-terminal ECD fragments, 1-293, 1-294 and 1-295, of rat and human LHR (17,18,32-36). Expressing the binding results as $(K_d)_{wt}/(K_d)_{mut}$, i.e. the ratio of the K_d of

the full-length receptor to that of the deletion mutant, the mean from these six reports is 1.1 with a range of 0.6-1.6. Thus, the 1-294 fragment we studied binds ligand with essentially the same affinity as does the full-length ectodomain of LHR. This is confirmed with the reduced binding without pretreatment with Factor Xa.

Y180 did not show binding activity under these conditions (Fig. 3.4). There is only one report on hCG binding to the 1-180 fragment, in which a binding ratio, $(K_d)_{wt}/(K_d)_{mut}$, of 0.1 was found (17), i.e. the N-terminal fragment of the LHR ECD binds hCG with an affinity some 10-fold less than that of full-length receptor. Thus, we believe that the reduction in the affinity of this length of ECD for hormone prevents the detection of specific binding in our system under the conditions used (1 nM ^{125}I-hCG), where there are stoichiometric amounts of single chain hCG and ECD 1-180 present after enzymatic digestion. However, since Y294 and Y341 retain high-affinity binding, low levels of specific binding are detectable, even with equivalent amounts of single chain hormone and ECD present.

The CD spectra of Y341 and YhCG were measured and that for Y341 shows a more negative mean residue ellipticity ([θ]) compared to YhCG in the 200-240 nm range, as well as a marked increase in [θ] at wavelengths below 200 nm (Fig. 3.5). These differences indicate the presence of more helicity in the Y341 protein compared to YhCG. Secondary structure estimates, calculated using two commonly employed methods (26-29), revealed values that did not vary significantly based on the two algorithms used (Table 3.1). Since helicity estimates from CD spectra are by far the most reliable of the secondary structural elements (37), the helical content of the free ECD was estimated by analyzing the difference spectrum using the same methods. As shown in

Table 3.1, the helicity of the ECD is estimated to be 18-20%. The estimate of helicity in YhCG from the CD spectrum is slightly higher, but in reasonable agreement with the expected value from the crystal structure (4,5), however the β strand estimate from the recorded spectrum was significantly lower than expected. Similar observations were made with single chain hFSH (38).

The CD spectra of Y294 and Y180 were determined and found to be very similar, when shown as mean residue ellipticity (Fig. 3.6), however, both displayed reduced (less negative) ellipticity compared to Y341 in the 200-240 nm range. Since the spectra were highly reproducible between protein preparations and measurements, we performed an analysis of helical content in each construct. The secondary structural elements of the spectra were extracted as described above and are compared in Table 3.1. Since the percent helix for Y294 and Y180 are about the same, this implies that the larger protein, Y294, contains more helical residues than does Y180, and both yield lower helical values than Y341. The self-consistent method of extracting secondary structural information from CD spectra was able to reliably deconvolute just the Y341 spectrum. Thus, only the results obtained from the CONTINLL program are tabulated for Y294 and Y180. However, for comparison, helical estimates were calculated for Y294 and Y180 using another method (30) that yielded values consistent with those from the CONTINLL method. In addition, estimates of secondary structure were obtained by subtracting the CD spectrum of YhCG from that of Y294 and of Y180, e.g. as described above for Y341, followed by standard analysis. The results are given in Table 3.1 and indicate a significant reduction in the amount of secondary structure present in the C-terminal truncated ECDs.

There are no *ab initio* or homology models of the N-terminal and C-terminal Cys-rich regions, however a suggestion was made that a portion of the C-terminal hinge region may have a cytokine-like fold (10). The β/α motif of a "typical" LRR consists of approximately six amino acid residues forming a β-strand, some 11 or 12 in a turn, and about seven in an α-helix (2,10). Thus, excluding contributions from the two Cys-rich regions of the LHR ECD, the LRR region is expected to contribute about 56-63 and 48-54 amino acid residues in α-helix and β-strand, respectively, depending upon whether one considers 8 or 9 LRRs (9,10). Realizing the imperfect nature of the LHR ECD LRRs and the variable lengths of the individual LRRs, the number of amino acid residues expected in secondary structure can be relaxed somewhat to consider an average of 7 ± 1 in a typical α-helix and 6 ± 1 in a typical β-strand. It then follows that the 8 or 9 LRRs could contribute roughly 48-72 (or 60 ± 12) and 40-63 (or 52 ± 12) residues in α-helix and β-strand, respectively.

The fraction of α-helicity of Y341 estimated from the CD spectrum converts to about 68-87 (or 77 ± 10) amino acid residues including the Flag tag, YhCG, the CTP plus Factor Xa cleavage site, and the complete 341 amino acid residue ECD; subtraction of the YhCG contribution (estimated from the CD spectrum) results in a remainder of about 51-70 (or 60 ± 10) helical residues in the ECD (Table 3.2). This range is in excellent agreement with that expected after subtracting the CD spectrum of YhCG from that of Y341 and analyzing the difference spectrum, 61-68 (or 65 ± 4) residues. Subtraction of the number of helical residues in crystallized hCG (4,5) from the analyzed CD spectrum of Y341 gives a somewhat higher range, 72±12, but nonetheless one that is similar to that

predicted by the model. These results, along with those obtained based upon identical analyses of the Y294 and Y180 spectra, are summarized in Table 3.3.

The assumptions made in these calculations from the model must be emphasized, namely that the Flag tag, Factor Xa cleavage site, and the two Cys-rich regions do not contain α-helical residues and that there are no major changes in α-helicity of YhCG or the ECD concomitant with association. Confirmation of these assumptions must await additional studies. It should be pointed out, however, that earlier CD studies on hCG indicated negligible helicity (39) consistent with the low percentage noted in the crystal structure (4,5). CD spectra of the hCG subunits (40,41), including disulfide-intact subunits and reduced, S-carboxymethylated subunits, tryptic fragments of the reduced, S-carboxymethylated subunits, and the CTP (42) revealed that only limited α-helicity could be induced in the helix-promoting solvent, trifluoroethanol, but only with reduced, S-carboxymethylated subunits and certain tryptic fragments. Thus, it seems unlikely that YhCG would undergo a large increase in helicity upon binding to the ECD. Furthermore, numerous secondary structure prediction programs suggest that the amino acid residuces in the N- and C-terminal Cys-rich regions have no significant propensity to form secondary structure (unpublished data, DeMars G, Fralish GB, Puett D).

The same approach to estimate the number of residues in β-strands was employed (Table 3.2, values in parentheses). We emphasize, as discussed by others (37), that estimates of β-structure from CD spectra tend to be low in proteins with a high content of β-strands, e.g. hCG (4,5). Furthermore, the accuracy of β-strand predictions is significantly lower than that obtained for helix (37). Thus, we have placed less

confidence in these values and less emphasis on these results here and in the subsequent discussion.

DISCUSSION

This study reports, for the first time, the CD spectrum of the full-length LHR ECD (amino acid residues 1-341)-yoked hCG complex (Y341), as well as the first reported CD spectrum of a single chain hCG. Of note, the spectrum of YhCG was distinct from that of native hCG (Fralish, GB, Narayan, P, Puett, D, unpublished observation), in contrast to a report where the CD spectrum for single chain hFSH was stated to be identical to that of native FSH (38). The spectra were analyzed extensively with multiple, commonly utilized algorithms. Consistent values for α-helix and β-strand were obtained regardless of the method used, with the helical estimates being far more reliable (nearly 90% accuracy can be expected) (37). Using the spectrum of the free YhCG, a variety of analyses were performed where the hormone's spectral contribution to that of the complex was subtracted to glean an estimate of the type and amount of secondary structure present in the ECD alone. The results of this analysis were compared to values that would be expected based upon the proposed homology model of the ECD (10) and considering the presence of either 8 or 9 LRR's. The findings suggest that the ECD contains 18-20% helix, some 6-7 % lower than was suggested based upon a CD spectrum of a shorter porcine LHR ECD (residues 1-277)-hCG complex (20).

The LHR ECD, like that of the other glycoprotein hormone receptors, is delineated by three distinct regions: an N-terminal Cys-rich region (Exon 1, residues 1-28 in rat LHR), which is important for binding (15,18), a central LRR-containing region (exons 2-8, most of exon 9, residues 29-245 in rat LHR) also important in binding(1,2),

and a C-terminal Cys-rich region (exons 9, 10 and part of exon 11, residues 245-341 in rat LHR), which is not considered to be important for binding (17,43), but is critical in signal transduction (36,44,45). The results obtained for Y341 (summarized in Table 3.2) suggest that most of the secondary structure (α-helicity) is expected to lie in the purported 8 (13) or 9 (10) 'typical' LRR's, which are predicted to fold with a repeating motif of: β-sheet (6 residues), turn (11-12 residues) and α helix (6-8 amino acids). Thus, the C-terminal Cys-rich region, which has been suggested to fold in a chemokine-like manner (10), is not predicted to contain significant amounts of helix or strand. This inference is based upon: (i) Previous secondary structure predictions (PhD), where sequences within the LRR region were estimated (with some 74% accuracy, in a comparative analysis with controls) to contain considerable fractions of α-helix and β-sheet (9), (ii) The analysis (albeit limited) of the CD spectrum of porcine LHR ECD (1-277)-hCG complex, which lacked most of the C-terminal Cys-rich domain, contained 25% helix (20), (iii) The C-terminal Cys-rich region (and N-terminal Cys-rich region) contains no significant propensity to form secondary structure based upon numerous predictive methods (unpublished data, DeMars G, Fralish GB, Puett D). Thus, the spectral analyses of Y341 and YhCG provide direct structural evidence in favor of the proposed homology models, accommodating 8 or 9 LRRs.

Upon truncating the receptor's C-terminus to 294 amino acids, we observed, with high reproducibility, a marked reduction in the mean residue ellipticity, [θ]. The drop in [θ] represents a 50% reduction in the number of amino acids involved in helical secondary structure. This observation suggests that residues encoded by exon 11 of the ECD either contain a significant portion of helicity or are important for the stabilization

of the structure of the entire ECD. For the former to be the case, a strong propensity for helix in this region would be required, as most of the truncated 47 residues would be helical. As discussed above, this seems unlikely.

Alternatively, these data suggest stabilization of the secondary structure of the ECD via amino acid residues present in the C-terminal hinge region. This intriguing observation suggests a mechanism by which the N-terminal regions of the ECD, which are important for binding and hormone recognition, communicate structurally to the C-terminal regions of the ECD, which are critical in the transfer of signal to the TM domain. An emerging model of receptor activation suggests that hormone binding results in the disruption of an interaction between regions of the hinge-domain in exon 9 (residues 250-268) and exoloop 2 of the TM domain (44). When this interaction is disrupted by mutation (46), by peptide competition (36) or through the generation of chimeras (44), the receptor is activated. As such, interactions between the ECD and TM domain are suggested to restrict the receptor in a conformation not productive for coupling to G-protein, and when hormone binding occurs the receptor is 'relaxed' by disrupting this interaction. This model may extend to other members of this class of GPCRs, as the TSH receptor is constitutively active when its N-terminal ECD is removed (47). The critical nature of the amino acid residues in exon 11 for secondary structural integrity in the entire ECD suggests that interactions exist between the receptor's C-terminal hinge region and the central LRR region. Indeed, sequences truncated in Y294 have been shown to be critical in the transduction of signal, where mutation of the conserved charged residues Glu-332 and Asp-333 in rLHR resulted in a dramatic reduction of the ability of the hormone to activate receptor (45). Therefore, these data

represent direct, structural support for the hinge region's potential role in signal transduction, although it is unclear how this structural transition may directly affect this process.

It should be noted that the secondary structure present in the full-length Y341 is not required for binding as Y294 displayed similar binding activity as Y341. This observation is consistent with numerous studies comparing soluble LHR ECDs of similar length to those studied here (1,2,7,16-18,32-35). The receptor's tolerance to structural variability in the hormone has been demonstrated (24,48,49), however, this study illuminates the possibility that the receptor itself can be structurally altered while still retaining the ability to recognize hormone. Thus, the interactions responsible for high affinity recognition by either the hormone or receptor are not necessarily determined by a strict three-dimensional (lock and key) configuration. Indeed, further studies are needed to understand the extent of perturbation of the putative binding site.

The CD spectrum (mean residue ellipticity, [θ]) of the Y180 protein was nearly identical to that of Y294. These data suggest a repeating structural unit. For instance, if a region of a protein has a repeating β/α fold, as is predicted for the ECD, then the percentage of the secondary structure should remain constant regardless of the size of the truncation. Indeed, the predicted % helix is the same for both proteins (6-7%). Thus, these data provide direct structural evidence for a repeating motif in the LHR ECD, greatly bolstering the argument for a typical LRR fold.

In summary, we have utilized the yoked hormone-receptor system to produce significant quantities of functional hormone-receptor ectodomain complexes. In doing so, we were able to use CD spectroscopy for direct structural measurements on the

proteins. Various analyses of the spectra for the full-length Y341 and the free YhCG suggest the receptor contains 18-20% helix supporting the current model with 8 or 9 LRRs. Surprisingly, C-terminal truncation of the ECD by 47 amino acid residues resulted in the destabilization of considerable secondary structure of the ECD with no loss in binding activity. The destabilization observed provides direct structural evidence for the interaction of this region with other parts of the receptor and offers intriguing possibilities for the role of the hinge region as a signal modulator. Single chain hormone-receptor complexes have been useful for the production of significant quantities of proteins that are challenging to most expression systems. It is anticipated that they may be useful for other classes of ligand-receptor pairs.

ACKNOWLEDGMENTS

We thank Dr. Prema Narayan for her interest and critical reading of the manuscript, and Ron Seidel for helpful instruction on the CD spectrometer.

Table 3.1

Secondary structure analysis of YhCG, Y341, Y294 and Y180 from CD spectra and of LHR ECD 1-341, 1-294 and 1-180 from CD difference spectra

Secondary structure analysis based on CD spectra of fusion proteins and CD difference spectra, i.e. fusion protein minus that of YhCG. Aperiodic refers to amino acid residues that not in α-helices, β-sheets (or strands), or β-turns, i.e. often denoted as random coil.

Fusion protein analyzed	**Helix(%)**	**Sheet(%)**	**Turn(%)**	**Aperiodic(%)**
YhCG	7[a], 8[b]	35[a], 36[b]	24[a], 22[b]	30[a], 34[b]
Y341	11[a], 14[b]	35[a], 34[b]	19[a], 22[b]	26[a], 32[b]
Y294	8[c], 6[b]	40[b]	22[b]	32[b]
Y180	7[c], 6[b]	40[b]	21[b]	31[b]
ECD (Fusion protein – YhCG)				
1-341[d]	18[a], 20[b]	30[a], 30[b]	19[a], 21[b]	24, 30[b]
1-294[d]	9[c], 6[b]	38[b]	23[b]	33[b]
1-180[d]	8[c], 6[b]	38[b]	23[b]	33[b]

[a] Estimates based on analysis with SELCON3 (26-28).
[b] Estimates based on analysis with ContinLL (29).
[c] These estimates of α-helicity were based on the method of Hennessey and Johnson (30).
[d] Based on manual subtraction of the CD spectrum of YhCG from that of each fusion protein, i.e. with θ in mdeg at each wavelength with both proteins at identical μM concentrations.

Table 3.2

Estimates of the number of amino acids in α-helices (and β-strands) in full-length and C-terminal deletion fragments of the LHR ECD

Estimates of α-helicity are based on the CD spectra of Y341, Y294 and Y180, corrected for that of YhCG and hCG. Note: we emphasize the helicity values since estimates of β structure (value in parentheses) are much less accurate (37). See Results section for detailed analysis of rationale.

Type of analysis	**1-341**	**1-294**	**1-180**
Model prediction[a]	60±12 (52±12)	60±12 (52±12)	42±3 (36±3)
$YECD_{CD} - YhCG_{CD}$[b]	60±10 (135)	24±6 (144)	16±4 (98)
$(YECD\text{-}YhCG)_{CD}$[c]	65±4 (102)	24±4 (111)	13±2 (60)
$YECD_{CD}\text{-}\ hCG_{XRAY}$[d]	72±12 (123)	36±6 (132)	28±4 (86)

[a]Expected from the LLR model of the the ECD (10). The values denote mean ± range as expected from 8 or 9 LRR's.
[b]The number of amino acid residues in YhCG was estimated from the CD spectrum ($YhCG_{CD}$) and subtracted from the estimates based on the spectra of Y341, Y294 and Y180 ($YECD_{CD}$). The mean ± range is given based on multiple methods for secondary structure determination from CD spectra.
[c]The CD spectrum of YhCG was subtracted from that of Y341, Y294 and Y180 and the resulting difference spectrum, denoted "$(YECD\text{-}YhCG)_{CD}$", was analyzed for secondary structure.
[d]The CD spectra of Y341, Y294 and Y180 were analyzed and then the number of amino acid residues in helix and β-structure of crystallized hCG (4,5) was subtracted.

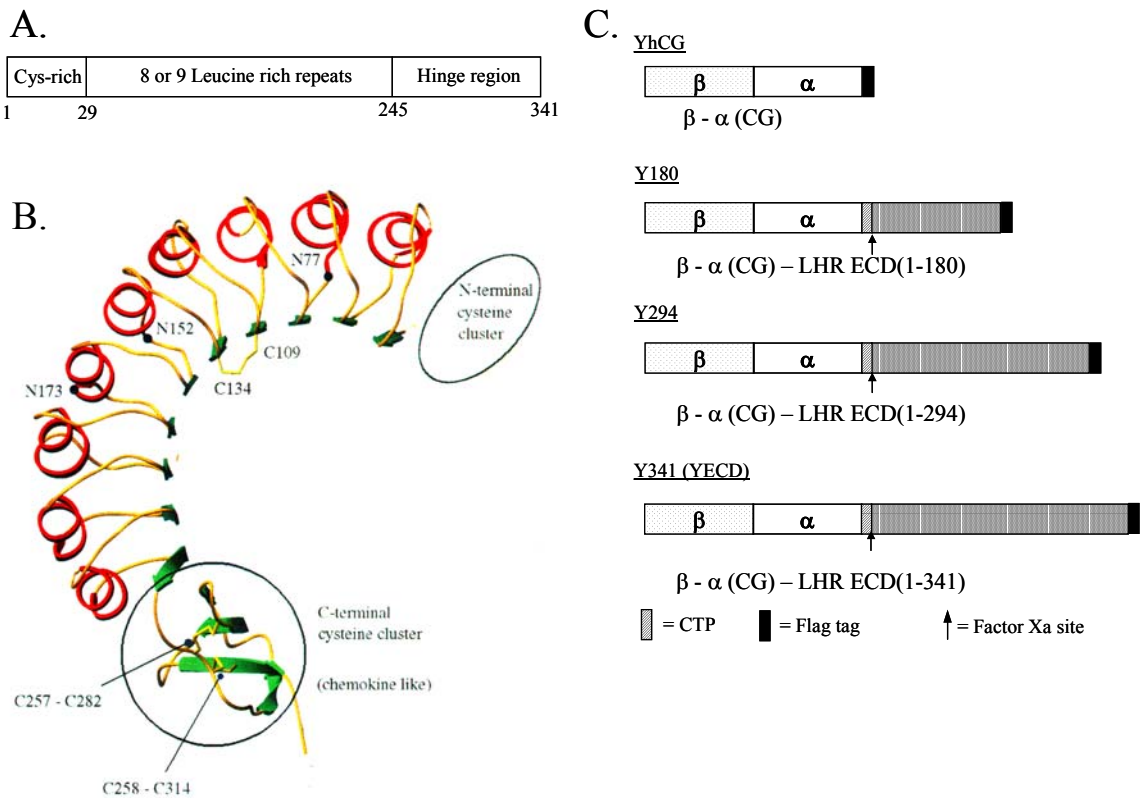

Figure 3.1. ECD model and fusion protein schematics. A. Schematic representation of the tripartite delineation of the ECD. The residues purported to reside in each region are indicated. B. Homology model of the LHR ECD (10). Printed with permission of the Endocrine Society. C. Schematic representation of Y341, Y294, Y180 and YhCG. YhCG has been characterized previously (22). Y341, Y294 and Y180 represent the soluble hormone receptor fusions. Y341 and Y294 are predicted to contain 8 or 9 LRRs (9,10), while Y180 is predicted to contain 6 LRRs. The naturally occurring CTP in hCG-β was used to link the subunits of the hormone and an additional CTP was used to link the hormone to the receptor. All fusion proteins contain identical arrangements of the hCG subunits and Factor Xa protease recognition sequence between hormone and receptor.

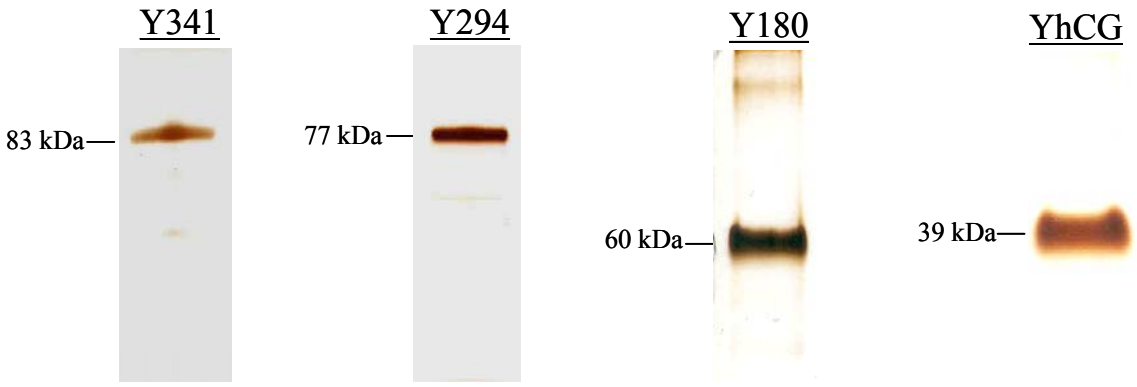

Figure 3.2. SDS-PAGE of Y341, Y294, Y180 and YhCG. The purity of the individual proteins was assessed using 10 % SDS-PAGE and silver staining. All lanes were loaded with 50-100 ng of protein and silver stained. The approximate molecular weights of the proteins estimated from standards are indicated.

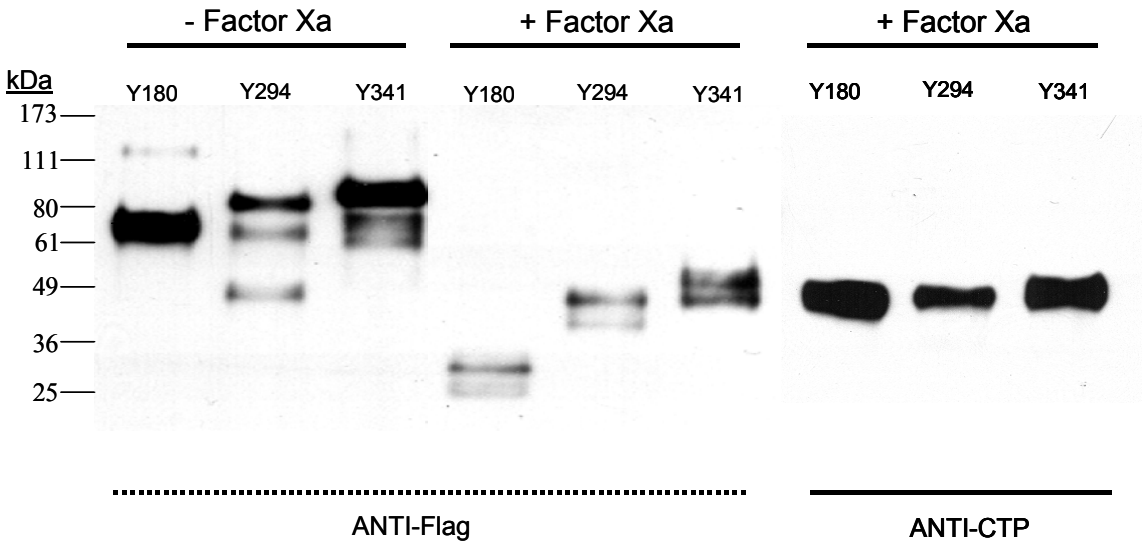

Figure 3.3. Factor Xa digestion and Western blot analysis of YECDs. Y341, Y294 and Y180 were treated with Factor Xa protease for 4 h at room temperature. The samples (5-10 ng protein) were resolved by 10% SDS-PAGE, then transferred to PVDF membrane and probed for the Flag or CTP epitope. Intact controls were probed for the Flag epitope.

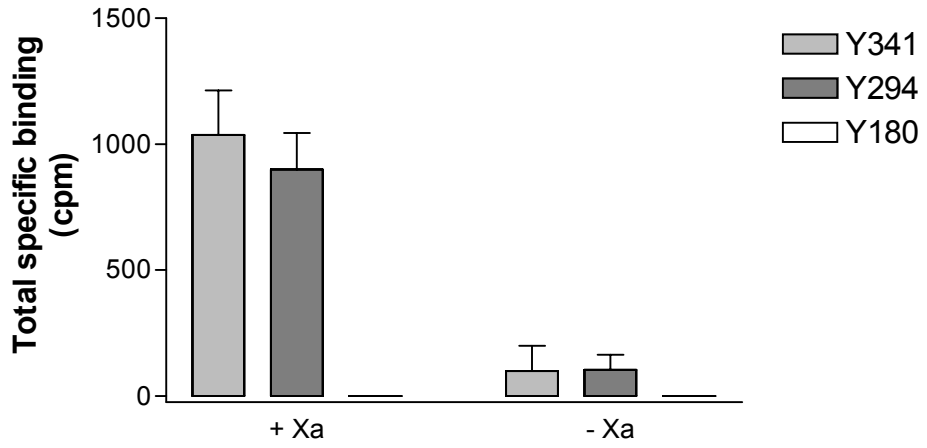

Figure 3.4. Total specific binding of Factor Xa-treated and untreated samples of Y341, Y294 and Y180. Soluble binding assays were performed as previously described (21). The ordinate represents total specific counts (cpm) bound to PVDF membrane squares blotted with 10 ng purified protein that was either treated or not treated with Factor Xa. Following incubation with 1 nM ^{125}I-hCG, specific counts were calculated by subtracting counts obtained in the presence of excess unlabeled hCG from total bound ^{125}I-hCG. Stoichiometric amounts of YhCG and ECD were present in the incubations with ^{125}I-hCG, and no corrections were made for competition of YhCG and hCG binding.

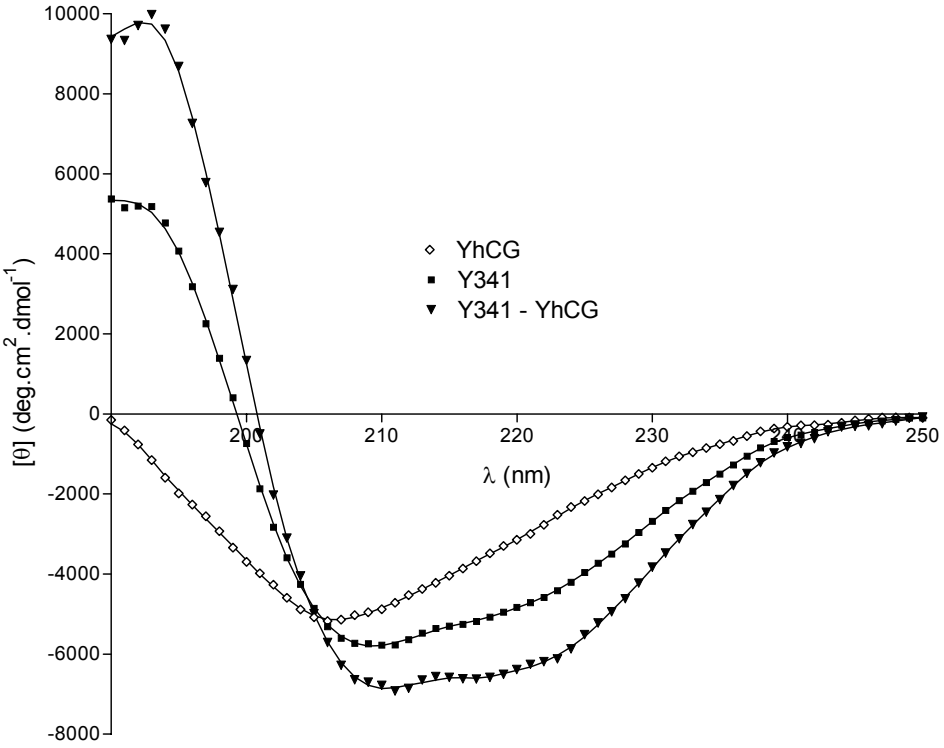

Figure 3.5. Circular dichroic spectra of Y341 and YhCG. Representative CD spectra of the purified Y341 and YhCG proteins were recorded and are presented as [θ], mean residue ellipticity. The spectrum of YhCG was subtracted from that of Y341 (with θ in mdeg) and the calculated difference spectrum (Y341-YhCG) was then converted to the appropriate units for comparison. Estimates of secondary structure are given in Table 3.1.

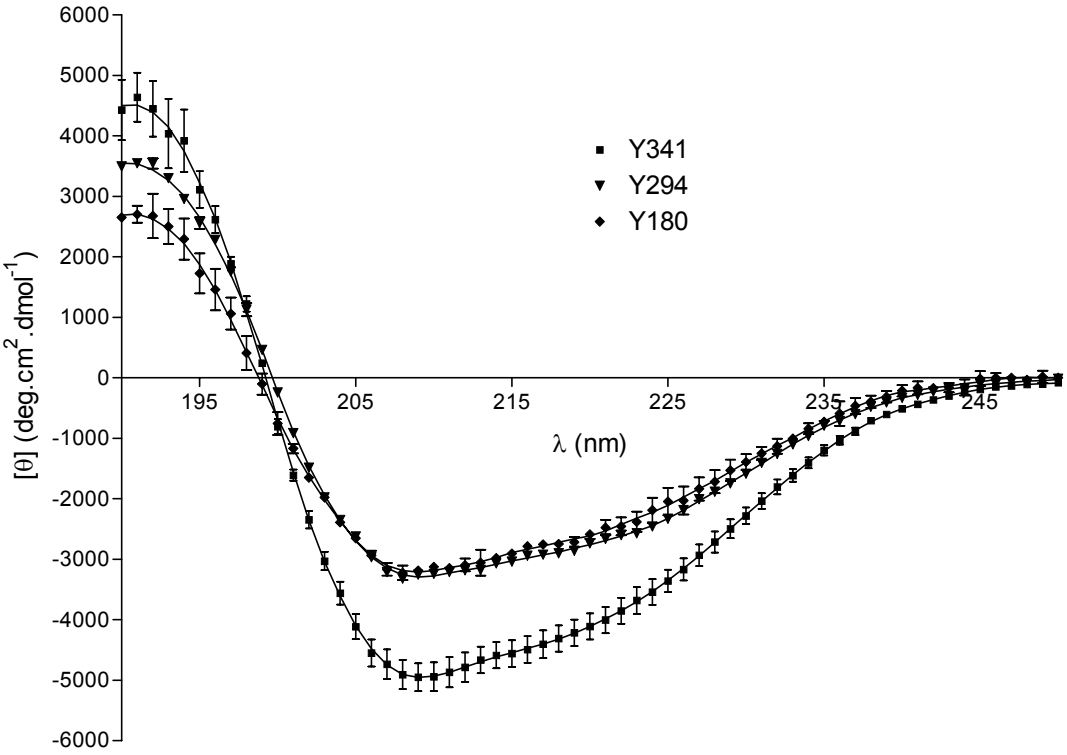

Figure 3.6. Circular dichroic spectra of Y341, Y294 and Y180. The CD spectra of Y341, Y294 and Y180 were measured. The spectra are shown as mean ± SEM to demonstrate the reproducibility of the results between protein preparations. The same subtractive analyses were done with Y294 and Y180 as for Y341 in Fig. 5 to obtain the secondary structure estimates provided in Table 3.1.

REFERENCES

1. Dufau, M. L. (1998) *Annu Rev Physiol* **60,** 461-496
2. Ascoli, M., Fanelli, F., and Segaloff, D. L. (2002) *Endocr Rev* **23,** 141-174
3. Hearn, M. T., and Gomme, P. T. (2000) *J Mol Recognit* **13,** 223-278
4. Wu, H., Lustbader, J. W., Liu, Y., Canfield, R. E., and Hendrickson, W. A. (1994) *Structure* **2,** 545-558
5. Lapthorn, A. J., Harris, D. C., Littlejohn, A., Lustbader, J. W., Canfield, R. E., Machin, K. J., Morgan, F. J., and Isaacs, N. W. (1994) *Nature* **369,** 455-461
6. Fox, K. M., Dias, J. A., and Van Roey, P. (2001) *Mol Endocrinol* **15,** 378-389
7. Xie, Y. B., Wang, H., and Segaloff, D. L. (1990) *J Biol Chem* **265,** 21411-21414
8. Moyle, W. R., Campbell, R. K., Rao, S. N., Ayad, N. G., Bernard, M. P., Han, Y., and Wang, Y. (1995) *J Biol Chem* **270,** 20020-20031
9. Jiang, X., Dreano, M., Buckler, D. R., Cheng, S., Ythier, A., Wu, H., Hendrickson, W. A., and el Tayar, N. (1995) *Structure* **3,** 1341-1353
10. Bhowmick, N., Huang, J., Puett, D., Isaacs, N. W., and Lapthorn, A. J. (1996) *Mol Endocrinol* **10,** 1147-1159
11. Kobe, B., and Kajava, A. V. (2001) *Curr Opin Struct Biol* **11,** 725-732
12. Bhowmick, N., Narayan, P., and Puett, D. (1999) *Endocrinology* **140,** 4558-4563
13. Song, Y. S., Ji, I., Beauchamp, J., Isaacs, N. W., and Ji, T. H. (2001) *J Biol Chem* **276,** 3426-3435
14. Keutmann, H. T. (1992) *Mol Cell Endocrinol* **86,** C1-6
15. Roche, P. C., Ryan, R. J., and McCormick, D. J. (1992) *Endocrinology* **131,** 268-274

16. Braun, T., Schofield, P. R., and Sprengel, R. (1991) *Embo J* **10,** 1885-1890

17. Hong, S., Phang, T., Ji, I., and Ji, T. H. (1998) *J Biol Chem* **273,** 13835-13840

18. Thomas, D., Rozell, T. G., Liu, X., and Segaloff, D. L. (1996) *Mol Endocrinol* **10,** 760-768

19. Lobel, L. I., Pollak, S., Klein, J., and Lustbader, J. W. (2001) *Endocrine* **14,** 205-212

20. Remy, J. J., Nespoulous, C., Grosclaude, J., Grebert, D., Couture, L., Pajot, E., and Salesse, R. (2001) *J Biol Chem* **276,** 1681-1687

21. Fralish, G. B., Narayan, P., and Puett, D. (2001) *Endocrinology* **142,** 1517-1524

22. Narayan, P., Wu, C., and Puett, D. (1995) *Mol Endocrinol* **9,** 1720-1726

23. Wu, C., Narayan, P., and Puett, D. (1996) *J Biol Chem* **271,** 31638-31642

24. Narayan, P., Gray, J., and Puett, D. (2000) *Endocrinology* **141,** 67-71

25. Gill, S. C., and von Hippel, P. H. (1989) *Anal Biochem* **182,** 319-326

26. Sreerama, N., and Woody, R. W. (1993) *Anal Biochem* **209,** 32-44

27. Sreerama, N., and Woody, R. W. (1994) *J Mol Biol* **242,** 497-507

28. Sreerama, N., Venyaminov, S. Y., and Woody, R. W. (1999) *Protein Sci* **8,** 370-380

29. Provencher, S. W., and Glockner, J. (1981) *Biochemistry* **20,** 33-37

30. Hennessey, J. P., Jr., and Johnson, W. C., Jr. (1981) *Biochemistry* **20,** 1085-1094

31. VuHai-LuuThi, M. T., Misrahi, M., Houllier, A., Jolivet, A., and Milgrom, E. (1992) *Biochemistry* **31,** 8377-8383

32. Tsai-Morris, C. H., Buczko, E., Wang, W., and Dufau, M. L. (1990) *J Biol Chem* **265,** 19385-19388

33. Chen, W., and Bahl, O. P. (1993) *Mol Cell Endocrinol* **91,** 35-41

34. Zhang, R., Cai, H., Fatima, N., Buczko, E., and Dufau, M. L. (1995) *J Biol Chem* **270,** 21722-21728

35. Hong, S. H., Ji, I. H., and Ji, T. H. (1999) *Mol Endocrinol* **13,** 1285-1294

36. Zeng, H., Phang, T., Song, Y. S., Ji, I., and Ji, T. H. (2001) *J Biol Chem* **276,** 3451-3458

37. Greenfield, N. J. (1996) *Anal Biochem* **235,** 1-10

38. Schmidt, A., MacColl, R., Lindau-Shepard, B., Buckler, D. R., and Dias, J. A. (2001) *J Biol Chem* **276,** 23373-23381

39. Holladay, L. A., and Puett, D. (1975) *Arch Biochem Biophys* **171,** 708-720

40. Birken, S., Gawinowicz Kolks, M. A., Amr, S., Nisula, B., and Puett, D. (1986) *J Biol Chem* **261,** 10719-10727

41. Puett, D., and Birken, S. (1989) *J Protein Chem* **8,** 779-794

42. Puett, D., Ryan, R. J., and Stevens, V. C. (1982) *Int J Pept Protein Res* **19,** 506-513

43. Thomas, D. M., and Segaloff, D. L. (1994) *Endocrinology* **135,** 1902-1912

44. Nishi, S., Nakabayashi, K., Kobilka, B., and Hsueh, A. J. (2002) *J Biol Chem* **277,** 3958-3964

45. Alvarez, C. A., Narayan, P., Huang, J., and Puett, D. (1999) *Endocrinology* **140,** 1775-1782

46. Nakabayashi, K., Kudo, M., Kobilka, B., and Hsueh, A. J. (2000) *J Biol Chem* **275,** 30264-30271

47. Zhang, M., Tong, K. P., Fremont, V., Chen, J., Narayan, P., Puett, D., Weintraub, B. D., and Szkudlinski, M. W. (2000) *Endocrinology* **141,** 3514-3517

48. Ben-Menahem, D., Jablonka-Shariff, A., Hyde, R. K., Pixley, M. R., Srivastava, S., Berger, P., and Boime, I. (2001) *J Biol Chem* **276,** 29871-29879

49. Jackson, A. M., Berger, P., Pixley, M., Klein, C., Hsueh, A. J., and Boime, I. (1999) *Mol Endocrinol* **13,** 2175-2188

CHAPTER 4

CONSEQUENCES OF TRANSLATION AS A SINGLE POLYPEPTIDE ON THE STRUCTURES OF TWO YOKED CHORIONIC GONADOTROPIN ANALOGS IN α-β AND β-α CONFIGURATIONS[3]

[3] Fralish G.B., Prema Narayan, David Puett. Submitted to Molecular Endocrinology, 9/05/2002.

ABSTRACT

Human chorionic gonadotropin (hCG) is a placental-derived heterodimeric glycoprotein hormone which, through the binding and activation of its heptahelical receptor (the luteinizing hormone receptor), rescues the corpus luteum and maintains pregnancy. The three-dimensional structure of hCG is known, however the relevance of its fold to bioactivity is unclear. While both subunits (α and β) are required for activity, recent data with single chain analogs have suggested a diminished role for the cystine-knot and an intact heterodimeric interface in binding and receptor activation *in vitro.* Herein, we report the purification and structural characterization of two yoked (Y) hCG analogs, YhCG1 (β-α) and YhCG3 (α-β). The fusion proteins yielded higher IC_{50}s and EC_{50}s than those of hCG; the maximal hCG-mediated cAMP production, however, was the same. Purified YhCG1, YhCG3 and heterodimeric hCG were structurally characterized using circular dichroic (CD) spectroscopy. Remarkably, the three proteins exhibit distinct far UV CD spectra, with YhCG1 containing somewhat more secondary structure than YhCG3 and hCG. Limited proteolysis gave products that were dissimilar for the three hormones, confirming and extending the conformational data obtained with CD. Native hCG was more resistant to limited proteolysis with proteinase K than the single chain analogs. The α-subunit of hCG was completely resistant to proteinase K, while the α-subunits in the yoked analogs displayed unique fragmentation. The β-subunit in hCG produced only one detectable cleavage product, but when tethered in the β-α configuration, it was nearly completely digested and, conversely, displayed more tolerance to proteinase K treatment in the α-β context. Taken together, the data suggest that the N-terminally translated subunit in the single chain analogs is folded differently

from that of hCG. The data presented herein provide direct structural evidence for altered three-dimensional conformations in the two single chain hCG analogs. Thus, the cognate G protein-coupled receptor can recognize and functionally respond to multiple ligand conformations.

INTRODUCTION

Reproduction in primates is governed by the actions of the pituitary-derived gonadotropins, LH and FSH, and the placentally-derived CG. LH and FSH promote gonadal steroidogenesis and the growth and maturation of follicles, while a mid-cycle surge in LH levels induces ovulation. The implanted, fertilized egg leads to the production of CG by the placenta, which in turn promotes the production of progesterone, thus maintaining pregnancy.

TSH, FSH, LH and CG form the family of glycoprotein hormones, which are characterized by their heterodimeric quaternary structure consisting of a common α-subunit and a hormone specific β-subunit (1). Biological activity of the glycoprotein hormones is conferred through their respective G protein-coupled receptors. TSH and FSH bind to specific receptors (TSHR and FSHR, respectively), while LH and CG bind to a common receptor (LHR). This is not surprising given the high degree of sequence homology (85%) shared between CG and LH β-subunits, with most of their dissimilarity occurring from the C-terminal 30 amino acid residue extension (CTP, C-terminal peptide) of hCG-β (1). The CTP, which contains four O-linked glycosylation sites, is important for the relatively long circulatory half-life of hCG (2). LHR binds hCG and LH with high affinity (K_d = 0.1-1.0 nM) primarily through its large, N-terminal extracellular domain (3).

The crystal structure of hCG (4, 5) revealed that, surprisingly, the α- and β-subunits have nearly identical folds despite having very little sequence homology. Both subunits form a three-looped structure maintained by a central growth factor-like cystine knot. The heterodimer is not constrained by a hydrophobic core and is devoid of significant helical structure. However, it is stabilized by a significant number of contacts at the dimer interface and by the β-subunit 'seat-belt' loop, which wraps around a portion of α and forms an intramolecular disulfide bond during the later stages of hCG folding in the endoplasmic reticulum (6). Based upon the structure of hCG and the conservation of the key cysteine residues, the other glycoprotein hormones were postulated to fold in a similar manner. This was recently confirmed with the structure of hFSH (7).

The role of the three-dimensional structure in the binding and signal transducing activity of the hormone is unclear. It is evident that heterodimerization is a prerequisite for efficient hCG secretion, as well as biological activity, and the individual subunits cannot activate LHR (8). The 'seat-belt' loop is intimately involved in conferring specificity (9), and the C-terminus of α is important in high affinity binding (10-13). These critical regions are located on one face of the hormone (4, 5) which carries a net positive surface charge that is complementary to the receptor's net negative surface charge reported in several homology models of the extracellular domain (14, 15). These data, coupled with the evidence of only subtle conformational changes during binding (16), strongly suggest the importance of the three dimensional structure of CG for its biological activity.

We and others have generated a single-chain hCG, consisting of the C–terminus of the β-subunit yoked or fused to the N-terminus of the α-subunit, a configuration

chosen primarily to maintain the free α-carboxy terminus, as well as having the endogenous CTP as a flexible linker. The N-β-α-C fused hCG (YhCG1) was biologically active, displaying *in vitro* and *in vivo* bioactivity similar to that of native hCG (17, 18). It was suggested that the N-β-α-C configuration, with the 30 amino acid residue CTP linker, permitted native-like folding of the fused hormone (17). Our laboratory and others have reversed the order of subunit tether to N-α-CTP-β-C (YhCG3) (19-21). This analog bound LHR poorly, but activated the receptor efficiently, thus uncoupling the two activities (19).

Single chain hCG's have enabled novel mutagenic and chimeric approaches to better understand glycoprotein hormone structure-function relationships since the obligatory dimerization step in secretion is bypassed, thus permitting expression of the variants. It was shown with a single-chain hCG that an intact cystine-knot in either subunit is not required for biological activity (22, 23). Furthermore, multiple β-subunits, when fused to a single α-subunit, could form bifunctional hormones (24, 25). Using monoclonal antibodies, single chain hCG (N-β-α-C) was shown to retain biological activity when some of the native quaternary interactions at the interface between the subunits were disrupted by mutation (26). However, complete disruption of subunit interactions, as in the single chain N-α-β-C (without linker), causes nearly complete loss of activity (21). From these elegant studies, it is apparent that native tertiary and quaternary structure are not a prerequisite to biological activity. Indeed, these structural characteristics have more ascribed importance to intracellular behavior (21-23, 26).

While the studies discussed above were critical in detecting changes in conformation, the nature and extent of the conformational differences between

heterodimeric and single chain hCG are still unclear. Is the conformation affected locally at the dimer interface only or are there global changes to structure? Does translation as a single polypeptide affect the folding of the individual subunits and how does the configuration of the subunits in the polypeptide affect their conformation? To address these questions, we have purified two single chain analogs, YhCG1 and YhCG3. The structural character of the two yoked hormones and native heterdimeric hCG were compared using the complementary techniques of circular dichroism (CD) and limited proteolysis. The CD spectra for the three proteins were all distinct, revealing considerable changes in conformational character. Furthermore, when subjected to limited proteolysis, the hormones displayed proteolytic fragmentation patterns that were distinct from one another. The native heterodimeric hCG was far more resistant to proteinase than the fused analogs, with only the β-subunit showing limited sensitivity. YhCG1 was highly sensitive to proteolysis with very little in the way of stable product. Conversely, YhCG3 displayed an increased tolerance to protease over YhCG1, with stable products consisting of portions of both α- and β-subunits. These results suggest that global conformational differences may be occurring in the single chain glycoprotein hormones. Interestingly, conformational differences occur not only between hormones that display altered bioactivities to native hCG (YhCG3), but also between proteins with similar bioactivity to hCG (YhCG1). Thus, it is apparent that LHR can productively recognize and respond to multiple protein conformers.

RESULTS

In this study, we utilized two previously characterized yoked hCGs (YhCG1 and YhCG3) to assess the possible differences in global conformation of hormones, since differences in bioactivity have been noted between the two and between each and hCG (17-21). The two proteins differ in the linear arrangement of the subunits (Fig. 4.1). YhCG1 contains the N-β-α-C configuration, with the β-subunit's CTP fulfilling the role of flexible linker between the two subunits; YhCG3 contains the reverse configuration, N-α-CTP-β-C, with an additional CTP inserted after α to allow for more conformational freedom for hormone folding (19). Both hormones have a Flag tag (DYKDDDDK) on their C-termini for purification purposes. The tag did not affect the binding and signaling activity of the crude YhCG1 and YhCG3 (data not shown). YhCG1 (17) and other similar constructs retain native-like *in vitro* and *in vivo* (18) activity, while YhCG3 (19) and similarly conceived fusion hCGs (20, 21) display significant reduction in binding affinity while retaining the ability to transduce signal effectively.

Both YhCG1 and YhCG3 were purified to a high level using immunoaffinity chromatography against the Flag tags (Fig. 4.2A,B). YhCG3, with its additional CTP and four additional O-glycosylation sites, migrates at a significantly higher apparent molecular weight (43 kDa) than YhCG1 (36 kDa). Silver stained gels display relatively homogeneous samples that were routinely achieved using the described purification protocol.

Heterodimeric hCG was analyzed for its purity using SDS-PAGE and silver staining (Fig. 4.3). hCG that was not boiled or reduced retained its dimeric configuration (Fig. 4.3A). Silver staining, which can detect as little as 1 ng of protein, shows the high

level of purity of the sample. A minor band was apparent at a molecular weight lower than that of the intact dimer that is too large to be a 'nicked' or β-core fragment (27). Boiling the hCG sample results in the dissociation of the subunits into two bands with approximate molecular weights of 20 kDa and 40 kDa, representing the α- and β-subunits, respectively, as confirmed by Western blot analysis (Fig. 4.3B). The α-subunit stains with a lower intensity than the β-subunit under the experimental conditions used. The α-subunit may absorb less silver than β.

The *in vitro* bioactivities of the purified hormones were assessed using standard competitive binding and cAMP assays. The three hormones displayed similar cAMP dose-response profiles (Fig. 4.4A), and the EC_{50}s of the three hormones were in relatively good agreement with those originally reported (19). While the potencies of YhCG1 and YhCG3 were less than that of hCG (2-fold and 6-fold, respectively), the maximum cAMP values at saturating hormone levels were similar (Fig. 4.4A, Table 4.1). In competitive binding assays, the IC_{50} of YhCG1 and YhCG 3 were 4-fold and 7-fold greater than that of native hCG, respectively (Fig. 4.4B, Table 4.1). Differences in absolute values from those previously reported (19-21) can be attributed to the variation when assaying crude and purified proteins with disparate RIAs. Nonetheless, the same trend is apparent, with YhCG1 displaying more 'native-like' binding activity than YhCG3.

CD spectroscopy is highly sensitive, particularly to the secondary structure of proteins given the spectral properties of the peptide bond in the far-UV (170-250 nm) (28). Accordingly, we have utilized the CD of proteins to compare the overall structures of the three hormones. Interestingly, each of the hormones displays a unique CD spectrum (Fig. 4.5) indicating differences in their solution conformations. Heterodimeric

hCG exhibits two prominent (negative) CD bands at 207 nm and 196 nm in agreement with published spectra (29, 30). The CD spectra for YhCG1 and YhCG3 also exhibit a similar negative band at 207 nm, but there are significant differences between the two fusion proteins and each differs from that of hCG. The mean residue ellipticity of YhCG1 is slightly (but reproducibly) more negative above 219 nm and significantly more positive below 219 nm. YhCG3 displays less negative ellipticity than hCG and YhCG1 below 225 nm and exhibits two shoulders at 230 nm and 220 nm. These spectral differences are consistent with immunogenic data published on a similar single chain analog, where the N-α-CTP-β-C orientation displayed a dramatic reduction in immunoreactivity to antibodies recognizing the individual subunits (21). Taken together, these results suggest a loss of the epitopes and reduction in secondary structure when the subunits are tethered in this configuration.

Secondary structure estimates can be extracted using a variety of algorithms, with α-helical values being by far the most accurate (31). Three commonly employed methods were used to obtain the values in Table 2 (32-34). These methods are based upon relatively large reference sets of standard proteins and are capable of detecting low levels of helicity and β structure. An earlier technique, based on curve resolution into Gaussian bands (29), did not detect the low level of α-helicity in hCG. YhCG1, shows a slight increase in helix content over native hCG, while helical estimates for YhCG3 are less than those obtained in hCG and YhCG1. Correcting for the differences in the number of amino acids residues in hCG, YhCG1 and YhCG3, the average % α-helix from Table 2 yields 11-12, 16-17, and 9-10 helical residues, respectively. That for hCG is an overestimate based on the crystal structure (4, 5), which gives 5 helical residues.

This discrepancy may arise from the inability of the existing programs to correct for glycosylation, aromatic amino acid residues, and disulfides, all which can make minor contributions to the far-UV CD spectrum of hCG (29).

To further test the conformational differences suggested in the CD experiments, limited proteolysis experiments were performed with the proteins. Proteases are sensitive probes of conformation, particularly tertiary structure, as cleavage depends not only on the amino acid sequence of the protein, but also the accessibility of the site. The general specificity protease, proteinase K (aliphatic or aromatic at the P2 site and Ala at the P1 site), was chosen based upon the presence of close to 90 potential cleavage sites throughout the entire sequence of the hormones. Thus, subtle variations in conformational states should be discernable with this enzyme. Silver stained gels of the reactions indicated a dramatic increase in sensitivity of the single chain analogs over hCG (Fig. 4.6). Native hCG was highly resistant to cleavage as only one product was apparent and significant portions of uncleaved material remained after the entire time 80 min digestion. On the other hand, little detectable product was apparent with YhCG1 digestions after 10-20 min, while YhCG3 exhibited a small reduction in molecular weight at 5 min and was reduced to two stable products after the full time period. Proteinase-only reactions did not produce any detectable products (data not shown).

Western analyses of the gels were performed to identify the origins on the products as well as to detect trace amounts of product. Of note, the variation in the molecular weights of the subunits of the heterodimer, and the undigested holoproteins observed in the silver stained gels and the Western blots can be attributed to the different standards utilized for the two analyses. In the native heterodimer, the α-subunit displayed

remarkable tolerance to cleavage (Fig. 4.7), while the β-subunit was stabilized as well, with only one cleavage product evident (Fig. 4.8). Notably with YhCG1, a stable band containing α-subunit was discernable only with Western blotting (Fig. 4.7). Some β-subunit cleavage products were also detectable, but of much lower intensity (Fig. 4.8). In YhCG3, multiple products (between 16-26 kDa) were detectable that contained α-subunit immunoreactivity (Fig. 4.7), but are not readily detected in the silver stain. However, only two distinct, stable products of β-origin were detectable that most likely correspond to the 24 and 20 KDa fragments seen in the silver stain (Fig. 4.8). Of note, samples incubated for the entirety of the experiment with no proteinase K displayed no endogenous proteinase activity ('0' lanes). Since all of the proteins are glycosylated at potentially the same positions, it is not expected that minor differences in glycosyl complexity will result in the marked differences in proteinase K susceptibility observed here. To summarize the overall proteinase K sensitivity of the various heterodimers and their subunits, a qualitative scoring was employed (Table 4.3).

DISCUSSION

The data herein provide the first direct structural measurements comparing conformations of hCG and the single chain proteins, YhCG1 and YhCG3. The CD spectra were recorded for the three bioactive hormones, and they reveal significant conformational differences in the proteins studied. Limited proteolysis experiments support the CD results and suggest that the yoked proteins display an increased flexibility compared to native heterodimer. Furthermore, the N-terminal subunit in the single chain hormones (e.g. α-subunit in YhCG3 and β-subunit in YhCG1) exhibits an increased

sensitivity to proteinase K thus suggesting a detrimental effect on folding of the subunit that is translated first in the single chain. The results obtained add significantly to our understanding of the structural consequences of translating the α and β subunits in a single polypeptide chain and their ultimate folding into bioactive hormones.

The glycoprotein hormones, LH, CG, FSH and TSH, comprise a family of proteins whose biological potency is governed by their ability to form heterodimeric structures intracellularly prior to secretion and release into circulation. The biologically active proteins consist of a common α-subunit associated noncovalently with a hormone-specific β-subunit. The crystal structures of hCG (4, 5) and hFSH (7) have provided considerable insight into the structure-function relationships of this family of cystine-knot-containing hormones. Indeed, the growth factor-like cystine-knot is fundamental to the structural integrity of the hormones by acting as a molecular scaffold on which the rest of the molecule is defined and thus was suggested to be integral to the proper presentation of hormone to receptor (1).

With the advent of single chain glycoproteins (17, 18), the importance of the cystine-knot to the biological activity of the hormones could be addressed, since the obligatory dimerization step in hormone secretion was bypassed. Indeed, several studies with single chain hCG (22, 23) and hFSH (35) have now ascribed the primary function of the cystine-knot to intracellular behavior and not to productive *in vitro* bioactivity. "First generation" yoked hormones were constructed as N-β-α-C so as to utilize the CTP of β as a linker and to retain the free C-terminus of α, which is reported to be important in binding to receptor (11, 12). YhCG1 was able to bind to and activate LHR with properties similar to those of the native heterodimer. More recently, the reverse

configuration (N-α-CTP-β-C; YhCG3) was analyzed by our laboratory and others (19-21). A similar trend was observed in that these hormones bound receptor with lower affinity but retained the ability to transduce signal effectively. Moreover, an N-α-β-C construct, lacking the CTP linker, was able to form heterodimeric-like immunoreactivity but was completely devoid of bioactivity (21). Thus, dimeric determinants alone do not satisfy the requirements for hormone activity. This point is supported by immunological evidence, where single chain mutant hormones, which lack native-like heterodimeric structure, can still bind and activate LHR with wild-type-like activity (26). Therefore, it is becoming more apparent that distinct conformers of glycoprotein hormones can bind and activate their receptor. Importantly, however, these provocative studies have not unveiled the extent or the nature of the purported conformational differences. To that end, we generated and purified significant quantities of YhCG1 and YhCG3 for direct structural comparisons with each other and with native hCG. Using the complementary and sensitive techniques, CD and limited proteolysis, several important characteristics of these hormones were gleaned.

First, both CD and limited proteolysis experiments confirmed that hCG, YhCG1 and YhCG3 represent distinct conformational entities. While it is expected that the structure of YhCG3 may differ from that of YhCG1 and hCG, the marked contrasting properties of hCG and YhCG1 are notable. Second, the yoked CGs were more sensitive to proteinase K than dimeric hCG, suggesting a more compact structure in the native heterodimer. Third, the N-terminal subunit in the yoked CGs displayed a decreased resistance to proteinase K, indicating a more open and flexible structure. Thus, the order

of subunit connectivity is important to the folding of the subunits. The relevance of these general observations are discussed below with the individual constructs.

The CD spectrum for YhCG1 suggests that it may contain more helix than native hormone, while the limited proteolysis experiments for YhCG1 implicate a much less stable overall structure than hCG with a shortened proteolytic half-life and the lack of distinguishable (silver stain) products. These apparent contradictory properties are supported by evidence that linear (reduced, S-carboxymethylated) free β-subunit and β-derived peptides have significant helical propensities (36). Indeed, it is not known if the appropriate disulfide bonding pattern is formed in YhCG1. Although, it is conceivable that, given the inherent complexity of the disulfide bonding network in the β-subunit alone, tethering the α-subunit to the C-terminus may promote mispairing and alter the natural folding pathway for β. Indeed, non-native disulfide bonds have been reported to form during the folding of β (6), and these may be kinetically trapped when the α-subunit is translated proximally. Increased aggregation has also been noted with the single chain hCGs (21). Thus, with a configuration of N-β-α-C, the β-subunit may be destabilized by mispairing of disulfides promoting minor helical formation in the localized regions. Consequently, the α-subunit can fold effectively, hence its stability as determined by Western analysis. The stable α-subunit can partially dimerize with the β-subunit and form an effective, native-like agonist. This hypothesis is supported by the loss of some but not all of the heterodimeric epitopes in the N-β-α-C configuration (26).

The YhCG3 CD spectrum displays a notable reduction in ellipticity, which extracts into a slight reduction in estimated secondary structure (helicity), although it is not significantly different from that of hCG. Also, the limited proteolysis noted in

Western analyses shows that the resistance of the α-subunit to proteinase observed in native hCG and YhCG1 is diminished in YhCG3. Together, these data support the idea that the α-subunit, when tethered in N-α-CTP-β-C context, is unable to achieve its native fold. This change in structure would presumably affect association with β. Thus, the α-subunit may exist in a 'quasi-free' state. Certainly, this would explain the reduction in the intensity of negative ellipticity and the loss of immunoreactivity to α-subunit specific antibodies (21), as the long loop of α would lack a defined structure (37), thus increasing this region's susceptibility to proteolysis. Furthermore, residues α33-51, which are important to hCG binding to receptor (38-41), are also located in this region, therefore the arrangement of these residues would likely be altered, which, in turn, is manifested in the observed reduction in binding affinity.

A unique feature of the quaternary structure of hCG and hFSH is the seatbelt of the β-subunit that wraps around the long loop of the α-subunit. In heterodimeric hCG, it has been proposed that the two subunits combine during biosynthesis with the closure of the seatbelt, i.e. formation of the β-Cys-26-Cys-110 disulfide, occurring toward the latter stages of folding and assembly (6). Yet functionally holoprotein formation of hCG can also occur between subunits in which the β-Cys-26-Cys-110 disulfide bond is intact (42). These and other results led Moyle and co-workers to suggest that the seatbelt in free hCGβ exists in a conformation that differs from that in the heterodimer (42). In their model, subunit association occurs with the long loop of α and the seatbelt of β moving relative to each other to give the stable arrangement in the heterodimer. We propose that the major conformational differences between heterodimeric hCG, α-β, and β-α involve the α long loop and the β seatbelt.

Single chain glycoprotein hormones have become invaluable tools for the study of structure-function relationships. The experiments presented herein add significantly to our understanding of the effects of single chain translation on the solution conformations of the heterodimeric glycoprotein hormones. Through direct structural data it is evident that single chain hCGs possess altered conformations that are not related to their *in vitro* bioactivities. The order in which the subunits are tethered affects the CD spectra of the hormones and the overall stabilities of the subunits, i.e. the N-terminal subunit that is translated first displays a decreased resistance to proteolysis, suggesting a less 'native-like' fold. Furthermore, the marked differences observed in using both CD and limited proteolysis suggest that single chain translation produces global conformational variation, in so much as the differences between native hCG and yoked hCGs are not localized to a single region. The ability of LHR to recognize multiple conformations suggests that the major evolutionary driving force for the heterodimeric structure of glycoprotein hormones may be secretory control and not hormone selectivity at the receptor interface.

MATERIALS AND METHODS

Cloning and expression of YhCG1 and YhCG3

YhCG1 and YhCG3 were cloned as previously described (17, 19). A Flag tag (DYKDDDDK) was added to the 3' end of both hormones using PCR and the 3'–primers,YhCG1:5'-TTACTTATCGTCATCGTCCTTGTAGTCAGATTTGTGATAATAAAC–3' and YhCG3:5'-GGCCTCGAGTTACTTATCGTCATCGTCCTTGTAGTCTTCTGGGAGGATCG-3'. Both hormones were cloned into the baculovirus transfer vector pVL1392/3

(Pharmingen), which utilizes the *polH* promoter for expression very late in the baculovirus infection cycle. Of note, both proteins retained the signal sequences from the original 5' proximal subunit. After co-transfection of the transfer vector and linearized baculovirus DNA (Autographa californica Nuclear Polyhedrosis Virus, AcNPV) Baculogold (Pharmingen) into SF9 insect cells, medium from the cells was used for two more serial infections for amplification of the recombinant virus. High-titer (10^7-10^8 plaque forming units/ml) recombinant virus was used for expression in suspension cultured Sf9 cells. A multiplicity of infection of 0.1 was used and the cell medium was assayed 4 days post-infection. A sandwich radioimmunoassay (IRMA) (Diagnostic Products Corporation) was used to quantify expression levels.

Protein Purification

Media were harvested from suspension cultures and clarified with two sequential centrifugation steps of 1400g and 17,700g respectively. Clarified expression medium was stirred vigorously at 4C while ammonium sulfate was gradually added to 80% saturation followed by stirring overnight at 4C. The protein pellet was harvested from the medium with two high speed spins (17,700g), and the pellet was resuspended in TBS (50 mM Tris-HCl, pH 7.5, 150 mM NaCl). The resuspended pellet was then dialyzed extensively against TBS and again clarified with high-speed centrifugation (17,700g). The dialyzed and clarified protein sample was applied to an immunoaffinity resin containing the M2 Anti-Flag monoclonal Antibody conjugated to agarose beads. Efficient binding of the proteins was achieved by cycling the sample over the column overnight at 4C, after which the column was washed extensively with TBS (500–1000 ml, for a 1 ml column). The proteins were eluted with a 10-15 column volume wash with

0.1 M glycine, pH 3.5. The eluent was immediately neutralized via collection of fractions in microcentrifuge tubes preloaded with 1M Tris-HC,l pH 8.0. The fractions were assayed using IRMA, and those containing hormone were pooled and concentrated as needed. Highly purified native hCG was obtained from Dr. A.F. Parlow and the NIDDK.

SDS-PAGE and Western Analysis

Protein samples were resolved with nonreducing 15% polyacrylamide gels and visualized by silver staining (50-100 ng protein sample loaded) and Western analysis (5-10 ng protein sample loaded) as previously described (43). Briefly, gels were fixed and transferred to PVDF (Millipore) using a tank transfer apparatus. The membrane was blocked (3% BSA, 0.2 % Tween 20, TBS) and incubated with either a 1:5000 dilution of the α3 antiserum (rabbit polyclonal, kindly provided by Dr. Irving Boime, Washington University, St. Louis, MO), which is specific for the α-subunit of hCG and a 1:10,000 dilution of rabbit antiserum specific for the β-subunit of hCG (also obtained from Dr. Irving Boime, Washington University, St. Louis, MO). The membrane was then washed and incubated with a donkey anti rabbit IgG conjugated with horseradish peroxidase. After extensive washing of membrane, the protein was treated with enzyme-linked chemiluminescence reagents (Amersham Pharmacia) with visualization achieved by exposure of the membranes to film.

In vitro Bioassays

Characterization by the binding and cAMP induction patterns of the different hormones was achieved using methods previously described (19). All purified hormones were quantified using IRMA (Diagnostic Products Corp.) for the bioassays. Competitive

binding assays were performed in 12-well tissue culture plates with HEK 293 cells stably expressing LHR. Cells were incubated with a fixed amount of ^{125}I-hCG (100 pM) and increasing amounts of unlabeled hCG overnight at room temperature in Waymouths MB medium containing 0.1% BSA. Cells were washed twice with 1N NaOH. The washes were collected and radioactivity measured using a γ-counter. For cAMP induction, stably transfected LHR cells were stimulated in 12-well tissue culture plates for 30 min with increasing amounts of hormone in Waymouth's MB Medium containing 0.1% BSA and 0.8 mM of the phosphodiesterase inhibitor, isobutylmethylxanthine. The medium was removed, and the cells in each well were lysed with 1 ml of ethanol, and followed by incubation overnight at –20˚C. The ethanol was collected, centrifuged, dried and resuspended in cAMP buffer that is compatible with the RIA used subsequently to quantify cAMP amounts (Perkin Elmer Life Sciences, Inc.).

Circular dichroism

Protein concentrations were determined by UV absorption with the extinction coefficients for the proteins being estimated at 280 nm from the primary sequences assuming completely oxidized cystines (44). Purified protein samples were diluted to 10 μM and dialyzed against 5 mM phosphate buffer, pH 6.8. The dialyzed samples were filtered with a 0.1 μm filter and loaded into a 1 mm pathlength cell. The CD spectra were measured in the far UV at wavelengths between 250–190 nm using a Jasco 710 CD spectrometer. The measurements were obtained with the following spectrometer settings: band width, 1 nM; sensitivity, 50 mdeg; response, 2 sec; scan speed, 20 nm/min; step resolution, 0.2 nm; starting wavelength, 250 nm; lowest wavelength, 190 nm; and 5 scans

per sample. The data are presented as the mean ± standard error of the mean (SEM) of three repeated measurements of 5 scans each with different protein preparations.

Limited proteolysis

Proteinase K was added to purified hormone that was diluted to 10 μM in TBS, pH 7.5, 5 mM $CaCl_2$ for a final enzyme:substrate ratio of 1:100. All reactions were stopped simultaneously by the addition of 40% SDS sample buffer, and the samples were boiled for 5 min and immediately resolved by 15% SDS-PAGE as described above. Nonspecific proteolysis in the hormone preparations and the proteinase K samples was analyzed by the incubation of samples containing hormone only and proteinase only for the longest digestion period. All of the samples were analyzed by Western analysis and silver staining as described above.

ACKNOWLEDGMENTS

We thank Dr. Irving Boime for his generous gift of the α- and β-subunit specific antibodies and Ms. Judy Gray for excellent technical assistance.

Table 4.1. Summary of *in vitro* bioactivities of purified heterodimer and yoked hCGs.

	IC_{50}[a] (ng/ml)	IC_{50}[a] (nM[d])	EC_{50}[b] (ng/ml)	EC_{50}[b] (nM[d])	R_{max}[c] (pmol/ml)	IC_{50}/EC_{50}[e]
hCG	21.04 ± 2.1	0.54 ± 0.05	3.2 ±.39	0.08 ± 0.01	1540 ± 239	1.0
YhCG1	72.4 ± 0.4	2.34 ± 0.48	8.2 ±1.1	0.21 ± 0.03	1174 ± 103	1.7
YhCG3	160.2 ± 15	3.55 ± 0.34	23.2 ±3.4	0.51 ± 0.08	1310 ± 174	1.0

[a] IC_{50} describes the concentration of hormone required to displace 50% of ^{125}I-hCG from HEK 293 cells stably expressing rat LHR.
[b] EC_{50} represents the dose of hormone that achieves 50% maximal cAMP production in HEK 293 cells stably expressing rat LHR.
[c] R_{max} values are indicative of maximal cAMP levels obtained from a saturating dose of 250 ng/ml hormone, after the subtraction of basal cAMP.
[d] Molar concentrations were calculated using molecular weight values described previously (19).
[e] Normalized to 1.0 for hCG.
Data are presented as the mean ± SEM and represent three experiments, each done in duplicate.

Table 4.2. Secondary structure analysis of the CD spectra of hCG and fused analogs, YhCG1 and YhCG3.

	% α-helix [a]	**% β–sheet** [a]	**% turn** [a]	**% nonrepeating** [a]
hCG	5.4, 3.9, 5.4	33.5, 36.6	23.4, 25.2	37.7, 33.6
YhCG1	7.9, 5.1, 7.3	36.1, 36.6	22.3, 23.6	33.7, 33.1
YhCG3	4.8, 4.5, 4.5	40.1, 33.9	21.9, 27.7	33.2, 33.6

[a] The secondary structural composition of the hormones were estimated using two algorithms, Continll (32) giving the first value in each colum, and Cdsstr (34), the second value. Calculation of % helix from the $[\theta]_{220nm}$ was also performed as previously described (33), and these values are the third numbers in the column under "% α helix." The term nonrepeating refers to an aperiodic conformation, i.e. one devoid of secondary structure.

Table 4.3. Qualitative evaluation of holoprotein and subunit resistance to Proteinase K.

	Dimer	**α-subunit**	**β-subunit**
hCG	++++	++++	+++
YhCG1	++	+++	+
YhCG3	++	+	++

Protein and subunit tolerance to proteinase K was qualitatively scored based upon the limited proteolysis experiments. "++++" denotes a high tolerance while '+' denotes high sensitivity.

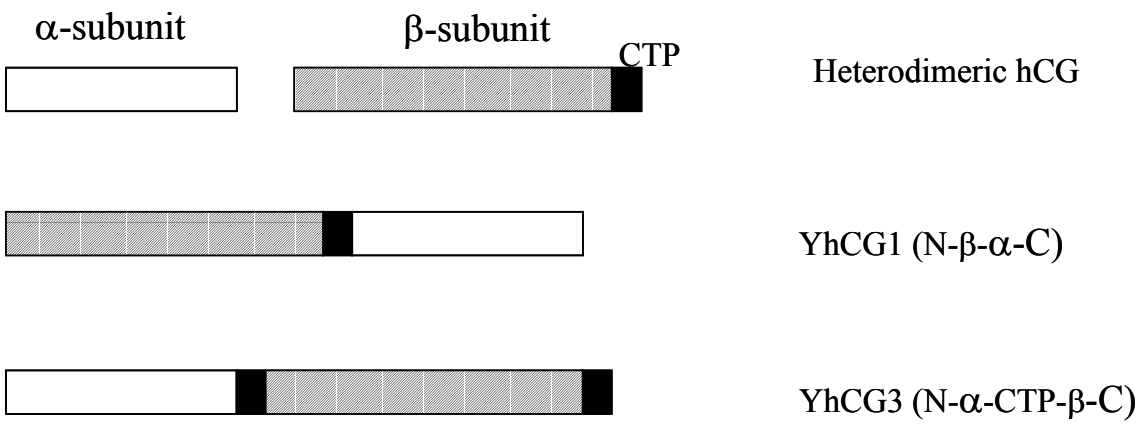

Fig. 4.1. Schematic representations of native hCG, YhCG1 and YhCG3. The C-terminus of the β-subunit is fused to the N-terminus of the α-subunit in YhCG1 as previously described (17). The flexible 30 amino acid residue CTP of the β-subunit was used as a endogenous linker. YhCG3 consists of the C-terminus of the α-subunit fused through a CTP to the N-terminus of the β-subunit. Thus, YhCG3 contains two CTP's (19).

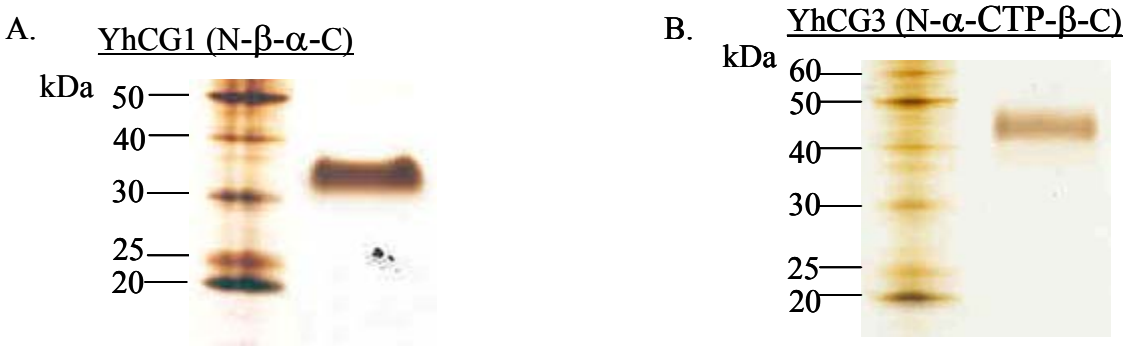

Fig. 4.2. Purification of YhCG1 and YhCG3. Immunoaffinity chromatography was used to purify YhCG1 (A) and YhCG3 (B). Purity was assessed by SDS-PAGE and silver staining of 50-100 ng protein. Molecular weight standards are indicated.

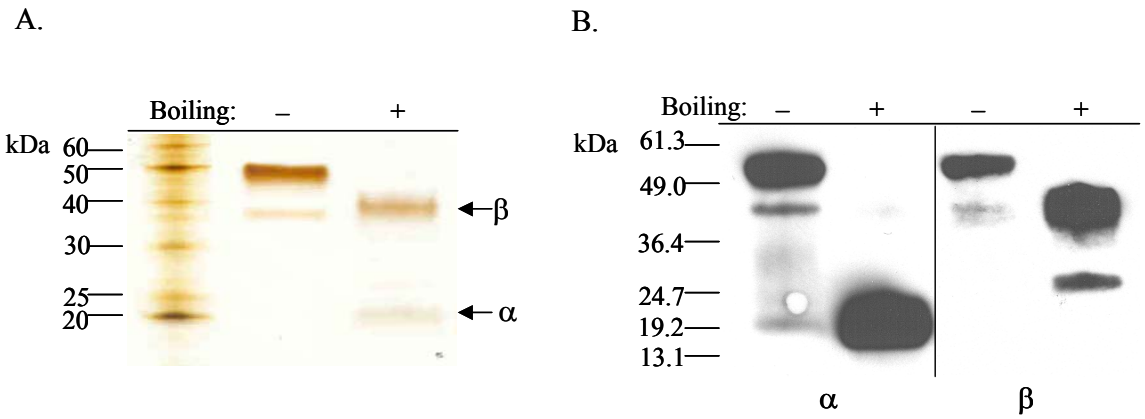

Fig. 4.3. Characterization of heterodimeric hCG. A. 50-100 ng of protein was analyzed by SDS-PAGE under non-reducing conditions (with and without boiling) and silver staining. B. Western blot analysis of 5-10 ng of purified hCG using antibodies specific to either the α-subunit (left panel) or the β-subunit (right panel). The sizes of the molecular weight standards (kDa) are indicated.

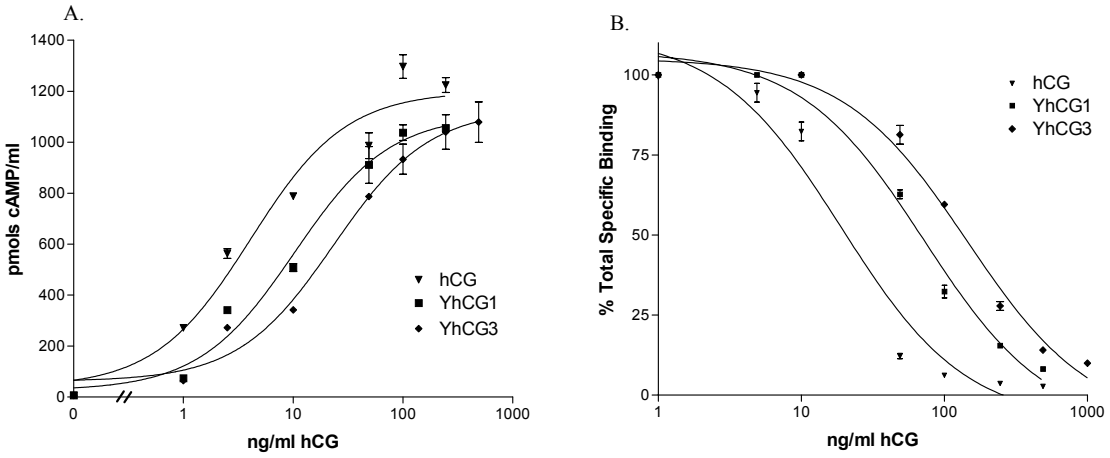

Fig. 4.4. ***In vitro*** **biological activity of heterodimeric hCG and purified single chain analogs YhCG1 and YhCG3.** A. Cells stably expressing LHR were stimulated with increasing amounts of hormone and assayed for the production of cAMP (pmoles/ml). B. Hormones were assayed for their ability to compete with ^{125}I-hCG for binding sites on cells stably expressing LHR. Data shown represent mean ±SEM of three experiments, each performed in duplicate.

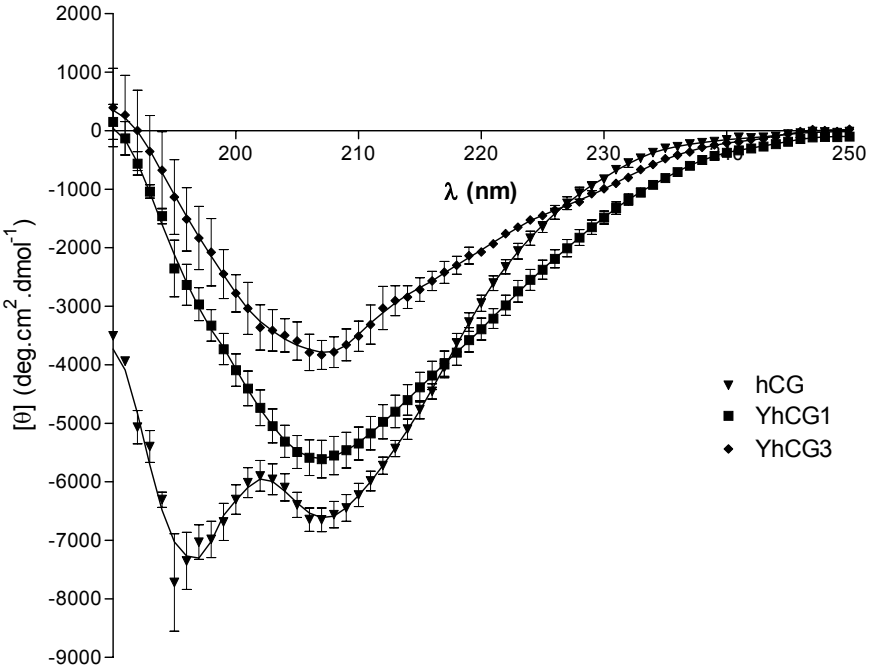

Fig. 4.5. Far UV CD spectroscopy of hCG, YhCG1 and YhCG3. The CD spectra for 5-10 μM purified hormone samples were recorded using a Jasco 710 CD spectrometer. Each spectrum represents the mean ± SEM of at least two separate experiments, of five scans each, with different protein preparations. Data are presented as [θ] mean residue ellipticity vs wavelength.

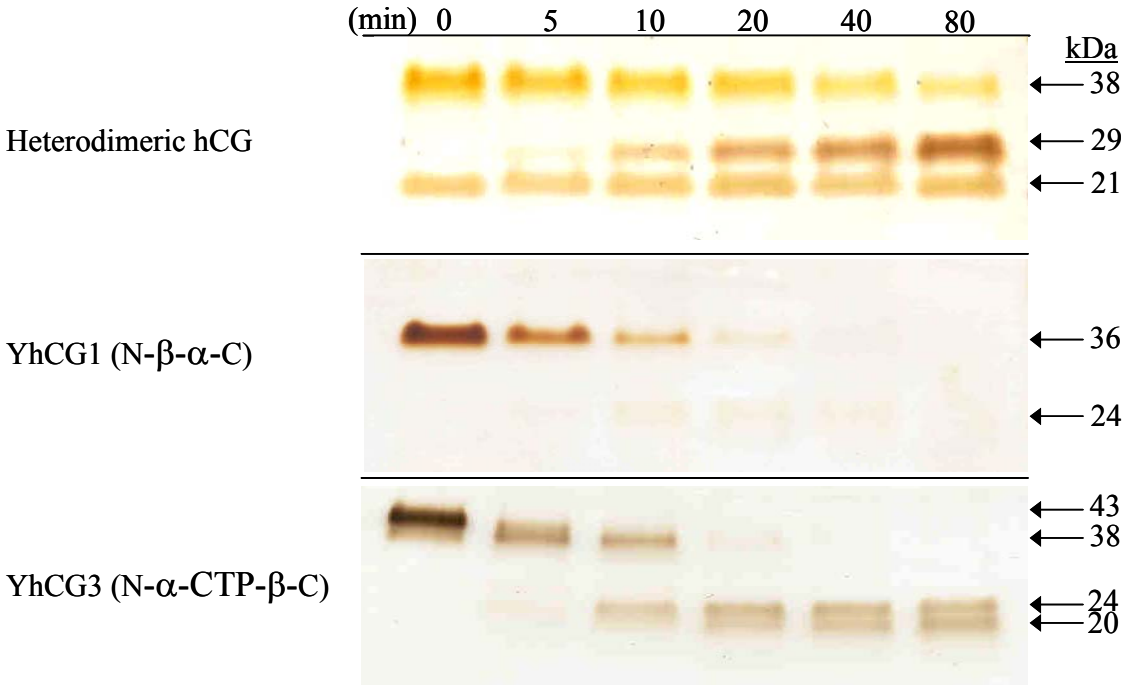

Fig. 4.6. SDS-PAGE and silver staining of limited proteolysis (proteinase K) reactions with hCG, YhCG1 and YhCG3. Hormones were digested with a 1:100 enzyme:substrate (w:w) ratio for various times at room temperature. The reaction was stopped by the addition of sample buffer and boiling. Samples were analyzed by gel electrophoresis (15% acrylamide) and visualized by silver staining. Proteolytic fragments are labeled with molecular weight estimates (kDa) based upon the migration of standards. Reactions containing only hormone (0) or only protease (not shown) were incubated at room temperature for the entire 80 min.

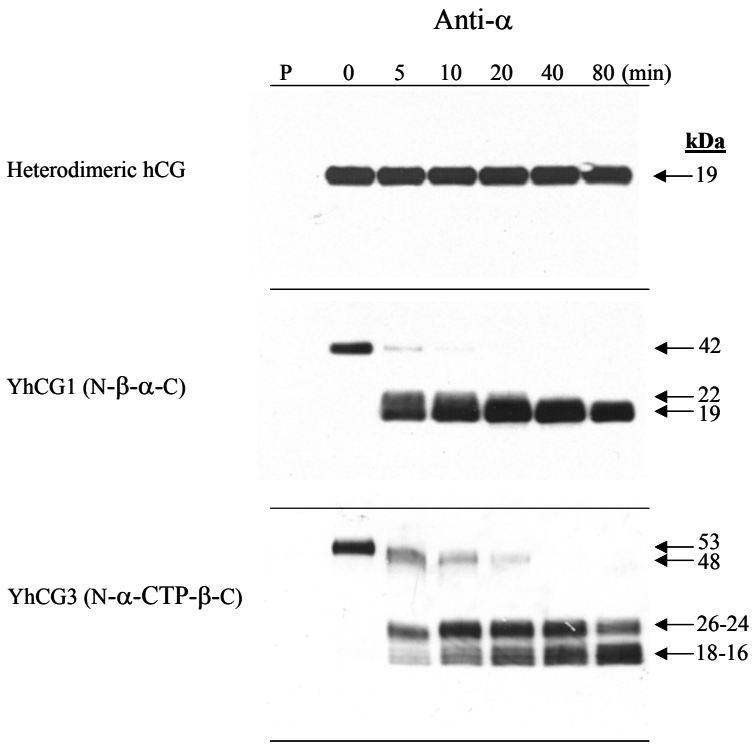

Fig. 4.7. Proteolysis of the α-subunit of hCG, YhCG1 and YhCG3 by proteinase K. Proteinase K was used to digest the hormones for the various times listed. The reactions were subjected to electrophoresis in 15% polyacrylamide gels, and fragments specific to the α-subunit were detected using Western blot analysis. Reactions containing only hormone (0) and only protease (P) were incubated at room temperature for the entire 80 min. Molecular weight estimates of the products were made based upon the migration of standards.

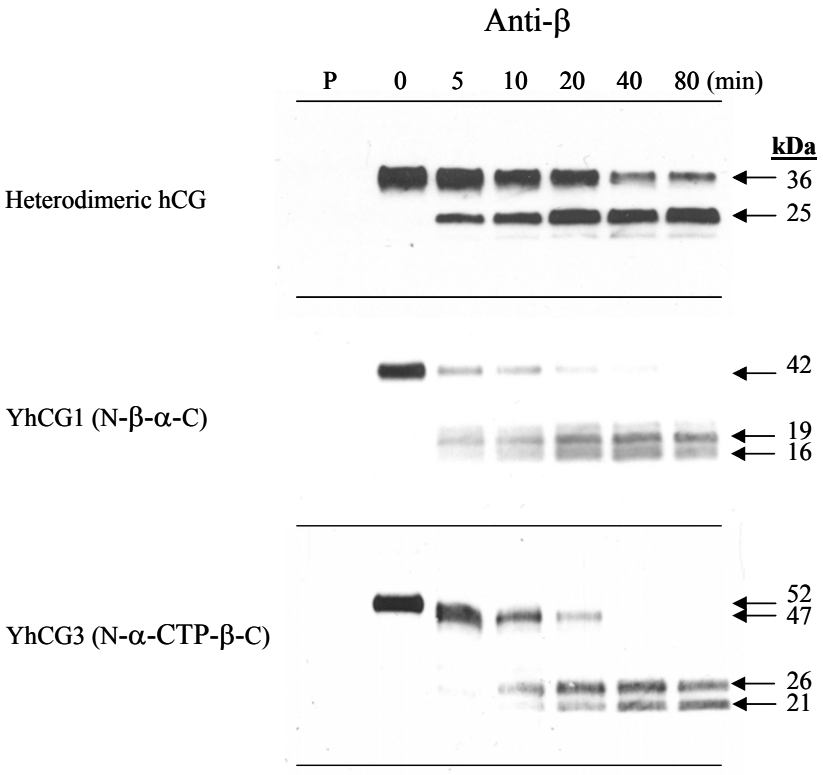

Fig. 4.8. Proteolysis of the β-subunit by proteinase K in the hormones hCG, YhCG1 and YhCG3. Hormones were subjected to proteolysis for the listed times and loaded on to 15% polyacrylamide gel for Western blot analysis with a β-subunit-specific antiserum. Estimations of the molecular weights (kDa) of proteolytic products were made based upon the migration of standards. Reactions containing only hormone (0) or only protease (P) were incubated at room temperature for the entire 80 min.

REFERENCES

1. Hearn MT, Gomme PT 2000 Molecular architecture and biorecognition processes of the cystine knot protein superfamily: part I. The glycoprotein hormones. J Mol Recognit 13:223-278
2. Matzuk MM, Hsueh AJ, Lapolt P, Tsafriri A, Keene JL, Boime I 1990 The biological role of the carboxyl-terminal extension of human chorionic gonadotropin beta-subunit [published erratum appears in Endocrinology 1990 Apr;126:2204]. Endocrinology 126:376-383
3. Xie YB, Wang H, Segaloff DL 1990 Extracellular domain of lutropin/choriogonadotropin receptor expressed in transfected cells binds choriogonadotropin with high affinity. J Biol Chem 265:21411-21414
4. Lapthorn AJ, Harris DC, Littlejohn A, Lustbader JW, Canfield RE, Machin KJ, Morgan FJ, Isaacs NW 1994 Crystal structure of human chorionic gonadotropin [see comments]. Nature 369:455-461
5. Wu H, Lustbader JW, Liu Y, Canfield RE, Hendrickson WA 1994 Structure of human chorionic gonadotropin at 2.6 A resolution from MAD analysis of the selenomethionyl protein. Structure 2:545-558
6. Ruddon RW, Sherman SA, Bedows E 1996 Protein folding in the endoplasmic reticulum: lessons from the human chorionic gonadotropin beta subunit. Protein Sci 5:1443-1452

7. Fox KM, Dias JA, Van Roey P 2001 Three-dimensional structure of human follicle-stimulating hormone. Mol Endocrinol 15:378-389

8. Pierce JG, Parsons TF 1981 Glycoprotein hormones: structure and function. Annu Rev Biochem 50:465-495

9. Grossmann M, Szkudlinski MW, Wong R, Dias JA, Ji TH, Weintraub BD 1997 Substitution of the seat-belt region of the thyroid-stimulating hormone (TSH) beta-subunit with the corresponding regions of choriogonadotropin or follitropin confers luteotropic but not follitropic activity to chimeric TSH. J Biol Chem 272:15532-15540

10. Kundu GC, Ji I, McCormick DJ, Ji TH 1996 Photoaffinity labeling of the lutropin receptor with synthetic peptide for carboxyl terminus of the human choriogonadotropin alpha subunit. J Biol Chem 271:11063-11066

11. Chen F, Wang Y, Puett D 1992 The carboxy-terminal region of the glycoprotein hormone alpha-subunit: contributions to receptor binding and signaling in human chorionic gonadotropin. Mol Endocrinol 6:914-919

12. Yoo J, Zeng H, Ji I, Murdoch WJ, Ji TH 1993 COOH-terminal amino acids of the alpha subunit play common and different roles in human choriogonadotropin and follitropin. J Biol Chem 268:13034-13042

13. Furuhashi M, Ando H, Bielinska M, Pixley MR, Shikone T, Hsueh AJ, Boime I 1994 Mutagenesis of cysteine residues in the human gonadotropin alpha subunit. Roles of individual disulfide bonds in secretion, assembly, and biologic activity. J Biol Chem 269:25543-25548

14. Jiang X, Dreano M, Buckler DR, Cheng S, Ythier A, Wu H, Hendrickson WA, el Tayar N 1995 Structural predictions for the ligand-binding region of glycoprotein hormone receptors and the nature of hormone-receptor interactions. Structure 3:1341-1353

15. Bhowmick N, Huang J, Puett D, Isaacs NW, Lapthorn AJ 1996 Determination of residues important in hormone binding to the extracellular domain of the luteinizing hormone/chorionic gonadotropin receptor by site-directed mutagenesis and modeling. Mol Endocrinol 10:1147-1159

16. Cosowsky L, Rao SN, Macdonald GJ, Papkoff H, Campbell RK, Moyle WR 1995 The groove between the alpha- and beta-subunits of hormones with lutropin (LH) activity appears to contact the LH receptor, and its conformation is changed during hormone binding. J Biol Chem 270:20011-20019

17. Narayan P, Wu C, Puett D 1995 Functional expression of yoked human chorionic gonadotropin in baculovirus-infected insect cells. Mol Endocrinol 9:1720-1726

18. Sugahara T, Pixley MR, Minami S, Perlas E, Ben-Menahem D, Hsueh AJ, Boime I 1995 Biosynthesis of a biologically active single peptide chain containing the human common alpha and chorionic gonadotropin beta subunits in tandem. Proc Natl Acad Sci U S A 92:2041-2045

19. Narayan P, Gray J, Puett D 2000 A biologically active single chain human chorionic gonadotropin analog with altered receptor binding properties. Endocrinology 141:67-71

20. Sen Gupta C, Dighe RR 2000 Biological activity of single chain chorionic gonadotropin, hCGalphabeta, is decreased upon deletion of five carboxyl terminal

amino acids of the alpha subunit without affecting its receptor binding. J Mol Endocrinol 24:157-164

21. Ben-Menahem D, Jablonka-Shariff A, Hyde RK, Pixley MR, Srivastava S, Berger P, Boime I 2001 The position of the alpha and beta subunits in a single chain variant of human chorionic gonadotropin affects the heterodimeric interaction of the subunits and receptor-binding epitopes. J Biol Chem 276:29871-29879

22. Ben-Menahem D, Kudo M, Pixley MR, Sato A, Suganuma N, Perlas E, Hsueh AJ, Boime I 1997 The biologic action of single-chain choriogonadotropin is not dependent on the individual disulfide bonds of the beta subunit. J Biol Chem 272:6827-6830

23. Sato A, Perlas E, Ben-Menahem D, Kudo M, Pixley MR, Furuhashi M, Hsueh AJ, Boime I 1997 Cystine knot of the gonadotropin alpha subunit is critical for intracellular behavior but not for in vitro biological activity. J Biol Chem 272:18098-18103

24. Ben-Menahem D, Hyde R, Pixley M, Berger P, Boime I 1999 Synthesis of multi-subunit domain gonadotropin complexes: a model for alpha/beta heterodimer formation. Biochemistry 38:15070-15077

25. Kanda M, Jablonka-Shariff A, Sato A, Pixley MR, Bos E, Hiro'oka T, Ben-Menahem D, Boime I 1999 Genetic fusion of an alpha-subunit gene to the follicle-stimulating hormone and chorionic gonadotropin-beta subunit genes: production of a bifunctional protein. Mol Endocrinol 13:1873-1881

26. Jackson AM, Berger P, Pixley M, Klein C, Hsueh AJ, Boime I 1999 The biological action of choriogonadotropin is not dependent on the complete native quaternary interactions between the subunits. Mol Endocrinol 13:2175-2188

27. Birken S, Maydelman Y, Gawinowicz MA 2000 Preparation and analysis of the common urinary forms of human chorionic gonadotropin. Methods 21:3-14

28. Woody RW 1995 Circular dichroism. Methods Enzymol 246:34-71

29. Holladay LA, Puett D 1975 Gonadotropin and subunit conformation. Arch Biochem Biophys 171:708-720

30. Erbel PJ, Haseley SR, Kamerling JP, Vliegenthart JF 2002 Studies on the relevance of the glycan at Asn-52 of the alpha-subunit of human chorionic gonadotropin in the alphabeta dimer. Biochem J 364:485-495

31. Greenfield NJ 1996 Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. Anal Biochem 235:1-10

32. Provencher SW, Glockner J 1981 Estimation of globular protein secondary structure from circular dichroism. Biochemistry 20:33-37

33. Chen YH, Yang JT, Chau KH 1974 Determination of the helix and beta form of proteins in aqueous solution by circular dichroism. Biochemistry 13:3350-3359.

34. Johnson WC 1999 Analyzing protein circular dichroism spectra for accurate secondary structures. Proteins 35:307-312

35. Hiro'oka T, Maassen D, Berger P, Boime I 2000 Disulfide bond mutations in follicle-stimulating hormone result in uncoupling of biological activity from intracellular behavior. Endocrinology 141:4751-4756

36. Puett D, Birken S 1989 Helix formation in reduced, S-carboxymethylated human choriogonadotropin beta subunit and tryptic peptides. J Protein Chem 8:779-94

37. Erbel PJ, Karimi-Nejad Y, De Beer T, Boelens R, Kamerling JP, Vliegenthart JF 1999 Solution structure of the alpha-subunit of human chorionic gonadotropin. Eur J Biochem 260:490-498

38. Reed DK, Ryan RJ, McCormick DJ 1991 Residues in the alpha subunit of human choriotropin that are important for interaction with the lutropin receptor. J Biol Chem 266:14251-14255

39. Bielinska M, Boime I 1992 Site-directed mutagenesis defines a domain in the gonadotropin alpha-subunit required for assembly with the chorionic gonadotropin beta-subunit. Mol Endocrinol 6:261-271

40. Liu C, Roth KE, Shepard BA, Shaffer JB, Dias JA 1993 Site-directed alanine mutagenesis of Phe33, Arg35, and Arg42-Ser43-Lys44 in the human gonadotropin alpha-subunit. J Biol Chem 268:21613-21617

41. Xia H, Chen F, Puett D 1994 A region in the human glycoprotein hormone alpha-subunit important in holoprotein formation and receptor binding. Endocrinology 134:1768-1770

42. Xing Y, Williams C, Campbell RK, Cook S, Knoppers M, Addona T, Altarocca V, Moyle WR 2001 Threading of a glycosylated protein loop through a protein hole: implications for combination of human chorionic gonadotropin subunits. Protein Sci 10:226-235

43. Fralish GB, Narayan P, Puett D 2001 High-level expression of a functional single-chain human chorionic gonadotropin-luteinizing hormone receptor ectodomain complex in insect cells. Endocrinology 142:1517-1524

44. Gill SC, von Hippel PH 1989 Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem 182:319-326

CHAPTER 5

CONCLUSIONS

The studies herein have shed light on this highly complex system of protein-protein interactions. The data suggest that the ECD of the LHR contains significant α-helix and β-folds with a repeating motif. C-terminal regions of the ECD are important in the stabilization of the structure of the ECD, providing an intriguing link to new ideas emerging on the role of the ECD in signal transduction. Structural studies on the yoked hormones have implicated global differences in their fold compared to hCG, and the fold does not solely dictate the functionality of the protein. Therefore, the binding of hCG to the ECD of LHR is a flexible interaction that is not restricted by global conformation.

The heterodimeric glycoprotein hormones (30–40 kDa), FSH, TSH, LH and CG, share a common α-subunit and are distinguished by a hormone specific β-subunit, which itself displays significant homology between the hormones (60-90% homology). LH and CG bind to the same receptor, while TSH and FSH display specificity for their own distinct receptors. The glycoprotein hormone receptors form a subgroup in the family of rhodopsin-like G protein-coupled receptors and are distinguished by their large (300-400 amino acid residues) ectodomains. Indeed, the ECDs of the glycoprotein hormone receptors also contain significant homology within their subgroup (50% similarity), yet there is little cross reactivity within this group of receptor-ligand pairs. The relative large size of the interacting partners and the homology between the different hormone-receptor

pairs underscores the complexity and fidelity of this system. The precise determinants of hormone selectivity are still poorly understood.

Following the release of the X-ray structure of the LRR protein, ribonuclease inhibitor, several homology models were proposed for the LHR ECD based upon the presence of an imperfect LRR motif in the primary sequence. The models differ considerably in the delineation, number and structural character of the repeats. However, they have been useful in providing a structural context for the numerous studies of ligand binding.

There is little structural data available for the glycoprotein hormone receptors. This is primarily a result of low expression levels of the ECD in heterologous expression systems. Recently, several groups have reported significant production of the LHR ECD and FSHR ECD. We observed a dramatic increase in levels of secreted ECD when CG was fused to the N terminus (YECD). Similar observations were reported with the porcine LHR-ECD (amino acid residues 1-277). Both studies noted increases in receptor stability, thus the ligand may induce an increased conformational rigidity. YECD was able to bind ^{125}I-hCG with high affinity, but only after the fused hormone was able to dissociate. Thus, in a single chain context, both hormone and receptor were able to fold and effectively associate.

In the studies reported herein, we have utilized yoked hormones and yoked hormone-receptor complexes to provide structural data for this complex system. YECD was shown to be functional, and, using CD spectroscopy, the receptor was shown to contain a significant amount of secondary structure. Extensive analysis revealed consistent values for α and β types of structure. Namely, the ECD is predicted to contain

18-20% α-helix and 30 % β-strand. These values are in agreement with the current prediction of 8 or 9 LRRs. Truncation of the C-terminus of YECD (Y341) by 47 amino acids led to a dramatic reduction in the amount of secondary structure in the ECD, while the protein seemed to bind hCG as well as full-length ECD (Y341). These data suggest that the C-terminal portions of the hinge region may contain a helical domain. Alternatively, the C-terminus may have a role in stabilizing the overall structure of the receptor. This idea is intriguing given the emergence of an increasing body of evidence implicating the hinge region as a signal modulator. Further, truncation of the ECD to just a 180 amino acid residue N-terminal fragment resulted in a stable protein with a similar CD spectrum (mean residue ellipticity [θ]) as Y294. These data provide the first structural evidence for a repeating fold in the LHR ECD.

Further information is needed to the exact interactions necessary for high affinity binding. Higher resolution approaches are obvious next steps in this field. Thus, NMR and X-ray crystallography are sure to become key techniques in the near future. However, more sensitive techniques than CD should be attempted in the event the protein is not amenable to crystallization or is too large for NMR. Hydrogen exchange coupled with mass spectrometry is useful in detecting and identifying regions between two proteins that are tightly associated. Certainly, this technique may afford a logical 'next step' to elucidating the regions of hormone and receptor that are important for binding and signal transfer.

Single chain glycoprotein hormones have become commonly utilized for structure function studies as the expression of mutants that affect dimerization in the native hormones most often will express as a single chain. The first yoked hormones were those

of CG, in the configuration N-β-α-C, published by two labs. This fused hormone (YhCG1) displays similar binding and signal transduction properties as the native hCG. Based on mutagenesis studies with YhCG1, tertiary and quaternary structure in hCG has a diminished role in binding and signal transduction as evidenced with monoclonal antibodies specific for particular epitopes. Furthermore, reversing the configuration of the single chain hormone to N-α-β-C (YhCG3) resulted in a hormone with altered binding activity that was uncoupled from the hormones ability to stimulate the receptor.

Thus, it is becoming evident that there is a level of structural tolerance at the receptor-hormone interface that was not expected before the advent of single chain hormones. Nevertheless, the nature and extent of the structural tolerance is unclear. Since there were no direct structural data available for the yoked CGs, YhCG1 and YhCG3 were expressed and purified in this study. These pure proteins, along with purified native heterodimeric hCG, were structurally characterized using CD spectroscopy and limited proteolysis with proteinase K. These results suggest that the three hormones differ in overall structure, suggesting global differences in structure. Interestingly, these differences were irrespective of their bioactivities. Indeed, YhCG1 displayed unique structural characteristics to hCG, despite having similar binding and signaling properties. Therefore, bioactivity is not a good measure for structural likeness. Further, the order in which the subunits are tethered affects the overall tolerance of the protein to proteinase K. A trend was observed where the N-terminal subunit displayed reduced stability.

These striking differences strongly support the need for further, higher resolution approaches to understand their precise structural nature of the fusion proteins. Certainly,

amide-deuterium exchange is a sensitive technique for detecting differences in protein structure, coupled with either NMR or Mass spectrometry, exact regions of variability would be discernable. Also, X-ray crystallography and NMR are important approaches that should be considered.

In summary, the complex nature of this interaction obviates the need for high-resolution studies. The work presented herein provides a foundation to continue in that direction, towards a 3-dimensional atomic model of the hormone-receptor complex and the yoked hormones. Understanding this interaction in detail will galvanize efforts to develop more specific and effective forms of birth control through the rational design of antagonists, for which there are currently none available. Alternatively, small molecule agonists to the receptor would be useful for the fertility clinic. The reagents and procedures developed here should greatly expedite the progression towards these goals.