Journal of Biomolecular NMR, 7 (1996) 295–304
ESCOM

J-Bio NMR 349

Expression of human chorionic gonadotropin uniformly labeled with NMR isotopes in Chinese hamster ovary cells: An advance toward rapid determination of glycoprotein structures

Joyce W. Lustbader[a,*], Steven Birken[a], Susan Pollak[a], Andrew Pound[b], Brian T. Chait[c], Urooj A. Mirza[c], Shakuntala Ramnarain[d], Robert E. Canfield[a] and Jonathan Miles Brown[d]

[a]*Department of Medicine, Columbia University, 630 West 168th Street, New York, NY 10032, U.S.A.*
[b]*Protein Chemistry Core Laboratory, Columbia University, 630 West 168th Street, New York, NY 10032, U.S.A.*
[c]*Rockefeller University, 1230 York Avenue, New York, NY 10021, U.S.A.*
[d]*Martek Biosciences Inc., 6480 Dobben Road, Columbia, MD 21045, U.S.A.*

Received 16 August 1995
Accepted 6 March 1996

Keywords: Glycoprotein hormone; hCG; Non-radioactive labeling; CHO cells

Summary

Most secreted eukaryotic proteins are modified by glycosylation, and it has been difficult to solve their structures by crystallographic or NMR techniques because of problems posed by the presence of the carbohydrate. The structure of a chemically deglycosylated form of the human pregnancy hormone, human chorionic gonadotropin (hCG), has been solved by crystallographic methods. Since chemical deglycosylation may have induced changes in the structure, and since it is known that deglycosylated hCG is biologically inactive, the crystallographic structure requires confirmation by NMR techniques. Also, it has not been possible to determine the structures of the isolated subunits, nor the nature of interactions between the carbohydrate side chains and the protein backbone by crystallographic methods. Structural information via NMR techniques can be obtained from proteins in solution if they can be uniformly labeled with ^{13}C and ^{15}N isotopes. We report the first such uniform labeling of a glycoprotein using a universal ^{13}C- and ^{15}N-labeling medium to express ^{13}C,^{15}N-labeled hCG, suitable for solving the structure in solution of the native, biologically active form of hCG as well as that of its free subunits. The ^{13}C,^{15}N-labeled recombinant hCG and its separated subunits are shown to be nearly identical to urinary hCG reference preparations on the basis of protein chemical studies, immunochemistry, biological activity, and the capability of isolated hormone subunits to recombine to form biologically active hormone. Mass spectrometric analysis and preliminary NMR studies indicate that the isotopic labeling is uniform and greater than 90% after only two growth passages in the labeling media. One unexpected finding during subunit purification was that lyophilization of glycoproteins from trifluoroacetic acid HPLC buffers may result in the loss of a significant portion of sialic acid.

Introduction

Unlike prokaryotic proteins, most eukaryotic proteins that appear on cell surfaces or are secreted by the cell have been posttranslationally modified by glycosylation (Wold, 1981; Wold and Moldave, 1984). Despite the great importance of eukaryotic glycoproteins, few structures have been determined because of difficulties in obtaining diffractable crystals. NMR spectroscopy of such glycoproteins has been compromised by their high sugar content, which obscures the resonances of some of the amino acid side chains and also increases the molecular weight of the compounds.

The solution of the three-dimensional structure of a protein may be facilitated by universally labeling the protein with ^{13}C and ^{15}N. In order to express forms of glycoproteins that are similar to natural products, mammalian cells must be utilized instead of bacterial systems, since the latter do not glycosylate their expressed proteins. However, unlike bacterial cells, eukaryotic cells

*To whom correspondence should be addressed.

TABLE 1
EFFECT OF MEDIUM FORMULATION ON THE hCG LEVEL

Formulation number	Characteristics of formulation	Supplemental amino acids (mg/l)	hCG expression level[a]
1A[b]	Partial hydrolysate, celtone 1 g/l	–	1.3
1B[b]	Partial hydrolysate, celtone 1 g/l	Glutamine (240)	1.1
2A[b,c]	Complete hydrolysate 1 g/l	Cysteine (100)	7.6
2B[b,c]	Complete hydrolysate 1 g/l	Cysteine (100) Glutamine (240)	7.1
3A[b,c]	Complete hydrolysate 1.4 g/l	Cysteine (80) Arginine (200)	12.6
3B[b,c]	Complete hydrolysate 1.4 g/l	Cysteine (80) Arginine (200) Glutamine (480)	16.4
4A[b,c]	Complete hydrolysate 3 g/l	Cysteine (160) Arginine (240)	17.9
4B[b,c]	Complete hydrolysate 3 g/l	Cysteine (160) Arginine (240) Proline (780)	19.3
5A[b,c]	Complete hydrolysate 3 g/l, no Asp	Cysteine (160) Arginine (240)	44.8
5B[b,c]	Complete hydrolysate 3 g/l, no Asp	Cysteine (160) Arginine (240) Proline (780)	41.4
5C[b,c]	Complete hydrolysate 3 g/l, no Asp	Cysteine (160) Arginine (240) Glutamine (330)	46.1

[a] Expression levels were found to be reproducible upon repetition of growth conditions.
[b] No asparagine.
[c] Supplemental glucose and pyruvate (3.9 g/l and 70 mg/l, respectively).

require defined media, containing all the essential amino acids necessary for growth in addition to carbohydrate. For a universally ^{13}C,^{15}N-labeled glycoprotein to be produced, all these components must themselves be universally labeled with ^{13}C and ^{15}N. To date, these materials have not been available. For example, an important effort by Fesik and co-workers (Hansen et al., 1992) fell short of achieving a uniformly ^{13}C,^{15}N-labeled medium because of the unavailability of dual-labeled cysteine. In addition, apparently not all carbon sources present in the medium were labeled. For example, glucose and pyruvate must also be isotopically labeled, or the cells will be provided with other potential sources of unlabeled carbons for biosynthesis. Finally, in their study, only 1 liter of medium was produced, a quantity inadequate for most production schemes.

We report the expression and characteristics of a uniformly ^{13}C,^{15}N-labeled glycoprotein hormone, human chorionic gonadotropin (hCG). This compound is the sole placental member of a family of four glycoprotein hormones that share a common α-subunit of 92 amino acids with target-specific β-subunits of 111–145 amino acids. Furthermore, hCG is the hormone necessary for the maintenance of pregnancy and is found at high concentrations in the urine of pregnant women, as well as in the urine of many cancer patients (O'Connor et al., 1994). The 3D structure of hCG has been pursued unsuccessfully for over 20 years since the early 1970s, when its primary structure was determined (Bellisario et al., 1973; Carlsen et al., 1973; Morgan et al., 1973,1975). It is only recently that we have succeeded in the determination of its crystal structure, in collaboration with two other groups (Lapthorn et al., 1994; Wu et al., 1994), by the crystallization of HF-treated, partially deglycosylated hCG (Lustbader et al., 1995). The crystal structure of this hydrogen-fluoride-treated form of hCG (HF-treated hCG) indicated that the previously proposed disulfide pairings for hCG were mostly in error (Mise and Bahl, 1980,1981). The structure showed that this hormone, and presumably the other three homologous hormones, FSH, LH and TSH, are members of a larger group of proteins sharing a structural motif known as a cysteine knot. Other known members of this family are three growth factors: NGF (McDonald et al., 1991), TGF-β2 (Daopin et al., 1992; Schlunegger and Grutter, 1992) and PDGF-BB (Oefner et al., 1992).

It has also not been possible to crystallize the subunits of hCG, even after chemical deglycosylation. It became therefore important to solve the structure of the native, glycosylated hCG. At present, NMR is the only available method to determine the structure of the isolated hCG subunits, as well as that of the native, glycosylated form of hCG.

Using algal sources of ^{13}C- and ^{15}N-labeled amino acids, we formulated a special cell-culture medium, which was developed by manipulation of the concentrations of

the labeled amino acids to achieve high levels of hormone expression by CHO cells. These cells are the type most commonly used for the production of eukaryotic proteins (Keene et al., 1989). The medium supports the growth of CHO mammalian-cell uniformly labeled hCG after only two passages in the medium. Use of this commercially available growth medium should allow widespread application of this labeling technology and a rapid growth in the number of glycoproteins that can be solved by the NMR approach.

Materials and Methods

Recombinant construct

Chinese hamster ovary cells, which had been transfected with the DNA of the α- and β-subunits of hCG (Matzuk et al., 1989), were generously donated to our laboratory by Dr. Irving Boime (Washington University, St. Louis, MO). The cells were first grown to determine the expression levels of the hormone. Originally, the cells produced approximately 1–2 mg/l of hCG. The cells were recloned by using a technique of limiting dilution (1 cell/well). After a number of trials we were able to recover a clone that expressed hCG at 20 mg/l (Lustbader et al., 1995).

Optimal media composition

A summary of the media formulations with the most significant effects on hormone expression is given in Table 1. Test media formulations, which included glucose and pyruvate as noted in Table 1, were unlabeled and were added to CHO-S-SFM-I (Chinese hamster ovary serum-free medium), which was obtained from the supplier (Life Technologies, Grand Island, NY) with all amino acids and carbohydrate omitted. For expression of labeled glycoprotein, the final ^{13}C,^{15}N medium formulation as well as ^{13}C-glucose (3.9 g/l) and ^{13}C-sodium pyruvate (70 mg/l) were added to the CHO-S-SFM-I minus amino acids and carbohydrate (Life Technologies). Concentrations of amino acids other than proline were determined by reversed-phase chromatography (YMC, 1994) of the corresponding *ortho*-phthalaldehyde derivatives using a 15-cm Supelcosil column (Supelco Chromatography Products, Bellefonte, PA) with a mobile-phase rate of 0.7 ml/min and the absorbance detector set at 340 nm. Proline was determined as previously described (Hancock, 1977). The resulting medium was sterilized by filtration through a 0.22 μm filter. The ^{13}C,^{15}N-labeled medium is available from Martek Biosciences Corp. (Columbia, MD) as Celtone M 54999 (Table 1, formulation 5C).

Incorporation of NMR isotopes: r-hCG and ^{13}C,^{15}N-labeled hCG cell culture

Cultures were maintained in tissue culture flasks in Ham's F-12 (JRH Biosciences, Lenexa, KS), 2.5% bovine calf serum, 2.5% fetal calf serum and 400 μg/ml G418 (Life Technologies) in a humidified 37 °C incubator with 5% CO_2. For screening purposes, 24-well tissue culture plates were seeded with the hCG-secreting CHO cells (usually 10^5 cells/ml) using the various media formulations. The cell supernatants were assayed for hCG production over several days using the B107 radioimmunoassay (monoclonal antibody specific for dimer hCG). The expression levels reported in Table 1 are those in the plate culture. For optimal recombinant hCG production, the cells were grown in suspension in biological stirrers (Techne, Princeton, NJ). No adaptation period was necessary; the cells were treated with trypsin to detach them from the tissue culture flasks and seeded at 10^5 cells/ml in the stirrer bottles. The serum-free medium used was the original formula of CHO-S-SFM (Life Technologies) with addition of 400 μg/ml G418 (Life Technologies). Cells were grown for 3 days, reaching a density of approximately 10^6 cells/ml. For the production of unlabeled hCG, cells were collected by centrifugation and resuspended in fresh CHO-S-SFM containing G418. The bottles were incubated for an additional 3–4 days and the hCG-containing medium was harvested by centrifugation and filtration through a 0.45 μm membrane. For the production of ^{13}C- and ^{15}N-labeled hCG, the cells were seeded and grown for 3 days as described above and then harvested by centrifugation, washed with PBS and resuspended in CHO-S-SFM with G418, containing only ^{13}C- and ^{15}N-labeled amino acids and ^{13}C-glucose and ^{13}C-pyruvate. After an overnight incubation during which unlabeled amino acids were exhausted in the growing cells, the cells were transferred to fresh ^{13}C- and ^{15}N-labeling medium, including ^{13}C-glucose and ^{13}C-pyruvate (the overnight incubation medium being discarded). Cells were grown in the labeling medium for 4 days and the medium was harvested as above.

Mass spectrometry

The mass spectra of the proteolytic peptide fragments of hCG α-subunit were obtained with an electrospray ionization quadrupole mass spectrometer constructed at Rockefeller University (Chowdhury et al., 1990). All spectra were obtained in the positive ion mode. The analyte solution was electrosprayed through a stainless steel needle that was maintained at 4 kV relative to a metal capillary tube through which the ions entered the mass spectrometer. A syringe pump (Harvard model 2400-001) maintained a flow rate of 0.5 μl/min of the analyte solution through the spray needle. Analyte concentrations of the electrospray solutions (water/methanol/acetic acid 47.5:47.5:5.0% (v/v/v)) were in the range of 15–20 μM. The mass spectra were acquired using a commercial data system (Teknivent Vector II). The total data collection time was 2 min per spectrum. The mass spectra of the products of total acid hydrolysis were obtained with a

Finnigan TSQ 700 triple quadruple mass spectrometer. The solution (10 μM total analyte in water/methanol/acetic acid 48.5:48.5:3 (v/v/v)) was electrosprayed at a rate of 3 μl/min.

Affinity purification of r-hCG

Affinity columns were prepared by coupling the monoclonal antibody B107 (specific for the dimer hCG) to CNBr-activated Sepharose 4B (Pharmacia, Piscataway, NJ) at a concentration of 5 mg B107 per ml of Sepharose. The hCG-containing cell supernatants were passed through the 25-ml bed-volume column and circulated overnight at 4 °C, followed by 15–20 wash volumes of phosphate-buffered saline (PBS) and a final wash with deionized water. Bound hCG was eluted with 1M acetic acid and the eluant was immediately brought to a pH of approximately 5 with ammonium hydroxide and dialyzed against deionized water.

Subunit preparation

The dimeric hormone was incubated in 1 ml of 6M guanidine hydrochloride, 0.1M Tris-acetate, pH 4.0, for 30–60 min at 37 °C. The solution was injected onto a Vydac C_4 semipreparative column (1 cm × 22 cm) and run at 2.0 ml/min with a mobile phase of 0.1% TFA (trifluoroacetic acid) (Buffer A) and 0.1% TFA in 100% acetonitrile (Buffer B). The gradient program was: 5 min at 0% B, 65 min to 35% B, 70 min to 90% B, then recycle to start. The column elution profile was monitored at 220 and 280 nm and each subunit peak was collected manually to obtain the highest purity. Each subunit peak was separated to base-line resolution. Solvent was removed by evaporation in a Savant Speed Vac (Savant, Farmingdale, NY).

Receptor binding and biological activities

CHO cells, engineered to express the rat LH receptor on their surface, were developed by Dr. William Moyle (Rutgers University, Piscataway, NJ) and generously donated for these studies. These cells were designated CHO-LR-HHH by Dr. Moyle. A competitive receptor assay was constructed as follows. The receptor-binding-assay mixture contained 100 μl of competing hCG (or buffer) in 0.1% BSA/0.9% NaCl, 100 μl of ^{125}I-hCG (50–100 000 cpm) in 0.1% BSA/0.9% NaCl and 100 μl CHO-LR-HHH (2×10^5 cells in phosphate-buffered saline). The cells were removed from the flask surface by adding versene only. The mixture was incubated for 1 h at 37 °C with occasional stirring, followed by centrifugation for 10 min at 750 × g. The supernatant was aspirated and the cell pellet was then counted.

Biological activity was determined on the basis of cAMP production. Amounts of 2×10^4 to 3×10^4 CHO-LR-HHH cells were stimulated with varying concentrations of hormone in 200 μl of 20 mM Hepes, 0.1% BSA, 4.75 mM KCl, 1 mM EDTA, 1.2 mM KH_2PO_4, 1.2 mM $MgSO_4$, 150 mM NaCl, pH 7.4. Cells were incubated for 15 min at 37 °C in a shaking water bath and then for 3 min in a 75–80 °C water bath. Tubes were spun at 750 × g for 10 min and then 100 μl of the supernatant was assayed for cAMP content using a commercial kit (DuPont/New England Nuclear, Boston, MA) according to manufacturer's instructions.

The reference preparations used in the bioassay study have all been described earlier. Preparation hCG CR127 is the currently widely distributed reference preparation for hCG (Birken et al., 1991). Nick-free hCG preparation 728 was prepared by hydrophobic chromatography from hCG CR127 as described earlier, and was shown to have an intact β-subunit by amino acid sequencing (Birken et al., 1993).

Amino acid sequencing and analysis

Protein sequencing was performed on an ABI model 470A sequencer according to manufacturer's directions. Amino acid analyses were performed on a Beckmann 6300 amino acid analyzer.

Electrophoresis

SDS–polyacrylamide gel electrophoresis was performed according to the procedure described by Laemmli (1970). The sample buffer contained 125 mM Tris-HCl, pH 6.8, 10% glycerol, 2% SDS and 0.01% bromophenol blue (reduced samples contained 5% 2-mercaptoethanol). The gel was silver-stained according to Wray et al. (1981).

Radioimmunoassays

Dose–response curves of urinary, recombinant wild-type and NMR-labeled α-subunit were obtained using a radioimmunoassay with polyclonal antisera as previously described (Birken et al., 1982). R111 (Birken and Canfield, 1982), which binds native α-subunit and hCG dimer, was used at an initial dilution of 1:12 000. R116 (Birken et al., 1986) has a high affinity for reduced carboxymethylated α-subunit but also binds native α-subunit and dimer to a lesser degree, and was used in this study at an initial dilution of 1:1000. The assay mixture contained 50 μl of the various α-subunit samples, 150 μl ^{125}I-hCG (20 000–40 000 cpm) and 50 μl of antisera, all diluted in PBS with 0.05% sodium azide and 1% normal rabbit serum. Samples were incubated for 1–3 h at 37 °C and then overnight at room temperature. Sheep anti-rabbit serum (0.5 ml of a 1:10 dilution) was added and the samples were incubated for 20 min at 37 °C and for 1 h at 4 °C, followed by centrifugation at 2000 × g for 15 min. The supernatants were aspirated and the pellet was counted in a Packard gamma counter.

Sialic acid analysis

Of each sample, 50 pmole was pipetted into 1.7 ml

o-ring cap tubes (Sarstedt, Newton, NC) along with 408 pmole of rhamnose internal standard. Tubes were dried in a speed-vac (Savant, Farmingdale, NY). Hydrolysis was carried out with 0.1N TFA for 1 h at 80 °C. The tubes were then removed, brought to room temperature, and dried. Chromatography was carried out using isocratic conditions, 150 mM NaOH, 100 mM NaOAc as the mobile phase using a Carbopac PA1 4 × 250 mm column and ED40 electrochemical detector (both from Dionex, Sunnyvale, CA).

Subunit combination

Urinary or isotopically labeled recombinant-hCG α-subunit (500 pmole) were incubated with a 10-fold molar excess of urinary β-subunit for 7 h at 37 °C in a total volume of 10 µl 0.2M ammonium bicarbonate. Aliquots were taken for gel electrophoresis and then the samples were diluted for the biological activity assay.

NMR spectrometry

The NMR spectra of the ^{13}C,^{15}N-labeled α-subunit were acquired on a Varian Unity Plus Spectrometer operating at a proton resonance frequency of 500 MHz. Lyophilized sample was dissolved in 700 µl of PBS and the pH was adjusted to 5.5 with HCl.

Results

Development of labeling medium

Table 1 summarizes landmark changes in the labeling medium, which resulted in major changes in the expression levels of hCG by the CHO cells. Development of an appropriate labeling medium required manipulation of the amino acid and sugar profiles obtainable from isotopically labeled algal biomass. A variety of unlabeled algal protein hydrolysates were tested, as shown in Table 1. Initial experiments (1A,B) utilized Celtone, a bacterial medium commercially available in isotopically labeled form (Martek Biosciences, Inc.). These formulations A and B represented partial hydrolysates similar to media used for bacterial cultures. Very little hCG production was observed. Use of complete methanesulfonic acid hydrolysates of various algal protein fractions (2A–5C), followed by removal of inhibitory components by ion exchange, resulted in significant increases in hCG expression. All amino acids, except cysteine, glutamine and asparagine, were stable under these conditions. The amino acids from some hydrolysates were isolated in order to be added as supplemental amino acids to various test formulations. Supplemental amino acids, as listed in Table 1, were always added in addition to the amino acids already present in the medium, which were derived from the algal hydrolysates. Aspartic acid and asparagine were found to be nonessential amino acids for these cells (5A–5C) and were omitted. Earlier formulations contained aspartic acid from the algal hydrolysis. The cells presumably synthesize these amino acids from glutamic acid and/or glucose. The total amount of amino acids was increased to 1.4 g/l for 3A and 3B and then doubled for all other formulations. For the production of the ^{13}C,^{15}N-labeled 5A–5C media, protecting group chemistries to protect cystine residues during hydrolysis were devised (details will be described elsewhere). Labeled cystine was cleaved to cysteine and this amino acid was added to the other labeled amino acids in the quantities indicated in Table 1. Arginine, aspartic acid and glutamic acid were each separately isolated from the amino acid mixture, the latter for subsequent conversion to glutamine as required. Replacement of glutamic acid with glutamine resulted in improved expression and this formulation was taken as the final medium (5C).

Table 2 shows the complete amino acid composition of the most successful media (5A and 5C). For the isolation of expressed material, the production was switched from the media-testing format (24-well plates) to the spinner bottles used in the preparation format. Our current production level of ^{13}C- and ^{15}N-labeled hCG is approximately 10 mg/l.

TABLE 2
AMINO ACID ANALYSIS OF MEDIUM 5A AND 5C

Amino acid	Amino acid content (mg/l)	
	Medium 5A	Medium 5C
Alanine	380	380
Arginine-HCl	240[a]	240
Asparagine	0[b]	0
Aspartic acid	0[b]	0
Cysteine	160[a]	160
Glutamic acid	330[a]	0
Glutamine	0	330[c]
Glycine	130	130
Histidine	60	60
Isoleucine	90	90
Leucine	330	330
Lysine	400	400
Methionine	80	80
Phenylalanine	180	180
Proline	180	180
Serine	200	200
Threonine	220	220
Tryptophan	50	50
Tyrosine	160	160
Valine	200	200

For medium 5A and 5C the contents of the final versions are presented.

[a] Arginine, cysteine and glutamic acid are separated from the rest of the amino acids during purification of the algal protein hydrolysate. Each was added back to the pool at the concentrations indicated.

[b] Aspartic acid and asparagine were found not to be essential for growth. Since these amino acids were separated during purification procedures, they were not added back to the amino acid pool.

[c] Glutamic acid was converted to glutamine and added back to the media. The glutamine was still uniformly isotopically labeled.

TABLE 3
TIME COURSE OF RELEASE OF SIALIC ACID DURING INCUBATION WITH TRIFLOROACETIC ACID

Protein	Temperature	Released sialic acid (%)				
		0 min	10 min	60 min	300 min	1400 min
Urinary hCG	20 °C	0.9	1.3	1.4	3.6	12.3
Urinary hCG	37 °C	0	1.7	4.7	17.1	68.8
Recombinant hCG	37 °C	0	0	4.8	25.0	64.5
Urinary hCGα	20 °C	–	–	–	–	12.7
Urinary hCGβ	20 °C	0.8	1.5	2.2	4.2	14.7

Numbers indicate the percentage of free sialic acid released (moles free sialic acid/moles total sialic acid).

Purification of hCG and its subunits

The hCG was purified from the medium by immunoaffinity chromatography, as in earlier studies (Lustbader et al., 1995). Very little hCG was dissociated during this procedure. Gel electrophoresis studies and amino-terminal sequence studies indicated that the hCG was homogeneous and did not contain any peptide bond cleavages. The ^{13}C- and ^{15}N-labeled hCG appeared identical to unlabeled hCG in both of these analyses.

The reversed-phase method of subunit separation after incubation in guanidine resulted in quantitative recovery of each subunit with negligible complementary subunit contamination, as determined by SDS gel electrophoresis analysis. It was necessary to preincubate the hCG in guanidine prior to reversed-phase HPLC in TFA in order to achieve complete dissociation. It was found that incubation in TFA starting buffer only did not result in full subunit dissociation. After learning that the dried, purified subunits undergoing NMR analysis contained significant quantities of asialo glycoprotein and free sialic acid, we investigated release of sialic acid under the chromatography conditions used for subunit separation. TFA has been used for many years for glycoprotein subunit purification and was thought to be fairly innocuous to glycoproteins (Sairam et al., 1990). We investigated a time course of release of sialic acid after incubation in the TFA buffer at room temperature and at 37 °C to determine when and how the sialic acid glycosidic linkage had been hydrolyzed. We found that incubation of hCG in this acidic buffer leads to slow release of sialic acid from hCG and its free subunits (Table 3). Incubation under the conditions used during actual chromatography (1 h, room temperature) would be expected to release less than 2% sialic acid. However, after 5 h at room temperature 4% was released and as much as 15% was released after one day of incubation. After one day at 37 °C, as much as 69% of the sialic acid had been freed from the glycoprotein. It became clear that much of the sialic acid had been cleaved free of the peptide chain not during chromatography but at a later stage within the dried, lyophilized glycoprotein.

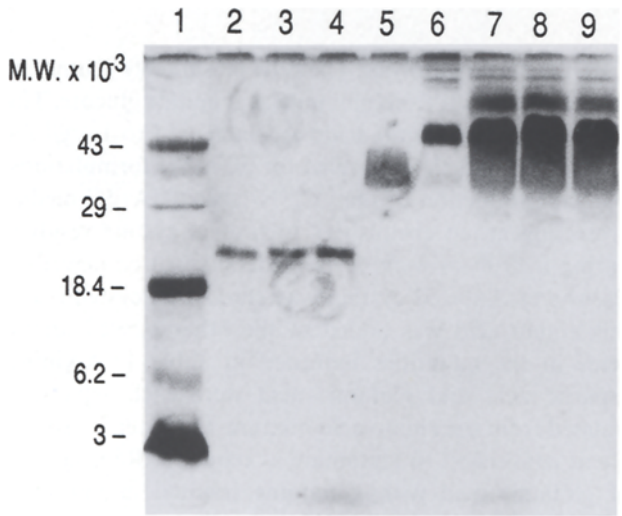

Fig. 1. Gel electrophoresis of the subunit combination mixtures. The α-subunits were recombined with urinary-β subunit and run on a 15% acrylamide, nonreduced Laemmli gel and silver-stained. Lanes 2–6 contain 25 pmole of subunit or dimer while lanes 7–9 represent combination mixtures of 25 pmole α-subunit and 250 pmole β-subunit. Lane 1: molecular-weight markers; lane 2: urinary α-CR123; lane 3: wild-type recombinant α-subunit; lane 4: ^{13}C,^{15}N-labeled recombinant α-subunit; lane 5: β-CR129; lane 6: hCG CR121; lane 7: recombined urinary-α-subunit; lane 8: recombined wild-type recombinant α-subunit; lane 9: recombined ^{13}C,^{15}N-α-subunit.

The isolated α-subunit was examined by two different methods to ascertain that it retained its native structure. Firstly it was examined immunologically by performing radioimmunoassays, using a rabbit antiserum directed to native urinary α-subunit as well as an antiserum directed to reduced, carboxymethylated urinary α-subunit. No significant differences in immunopotencies were observed among urinary α-, recombinant α- or ^{13}C- and ^{15}N-labeled α-subunit with either of the antisera. Again in this case, the ED_{50} values were calculated by the ALLFIT program (DeLean et al., 1978) and their ratio was found to be close to 1. Further evidence that the labeled recombinant α-subunit retains its native structure was its capability to combine with the urinary β-subunit (Fig. 1) and to display the same biological potency as in a simultaneous combination study using urinary α-subunit. Table 4 illustrates this study.

Biological activity

The isotopically labeled hCG was compared with the

TABLE 4
BIOLOGICAL ACTIVITY OF SUBUNIT COMBINATION

Produced dimer (fmole)[a]	Added α-subunit (fmole)[b]	Added β-subunit (fmole)[b]	Biological activity (fmole cAMP)			
			Urinary α-subunit + urinary β-sununit	Recombinant α-subunit + urinary β-subunit	$^{13}C,^{15}N$-α-subunit + urinary β-subunit	Urinary β-subunit[c]
1	1.1	11	236[c]	200[c]	268[c]	< 200[d]
3	3.3	33	376	510	554	
10	11.1	111	1380	1610	1920	< 200
30	33.3	333	2580	2870	3250	
100	111	1110	4700	4660	5130	296[d]
300	333	3330	4720	6070	5240	

[a] Theoretical quantity of dimer produced assuming a 90% combination.
[b] Quantity of subunit added to combination solution.
[c] Generated cAMP (fmole) following stimulation by combined, bioactive hormone (zero hormone equals less than 200 fmoles cAMP).
[d] Bioactivity of the β-subunit is attributed to contamination of the α-subunit in that preparation. This preparation of the β-subunit is known to have this extent of contamination of hCG.

current standard urinary hCG reference preparation (CR127), a new nick-free further purified CR127, and recombinant expressed hCG in both receptor-binding assays and signal transduction (Fig. 2). The receptor binding and biological activity of all four forms were similar, although both of the expressed hormones bound better than either of the urinary hCG preparations. When the data were fitted with the ALLFIT logistic curve-fitting routine (DeLean et al., 1978), the following ED_{50} values (fmole of amino-acid-analyzed calibrated

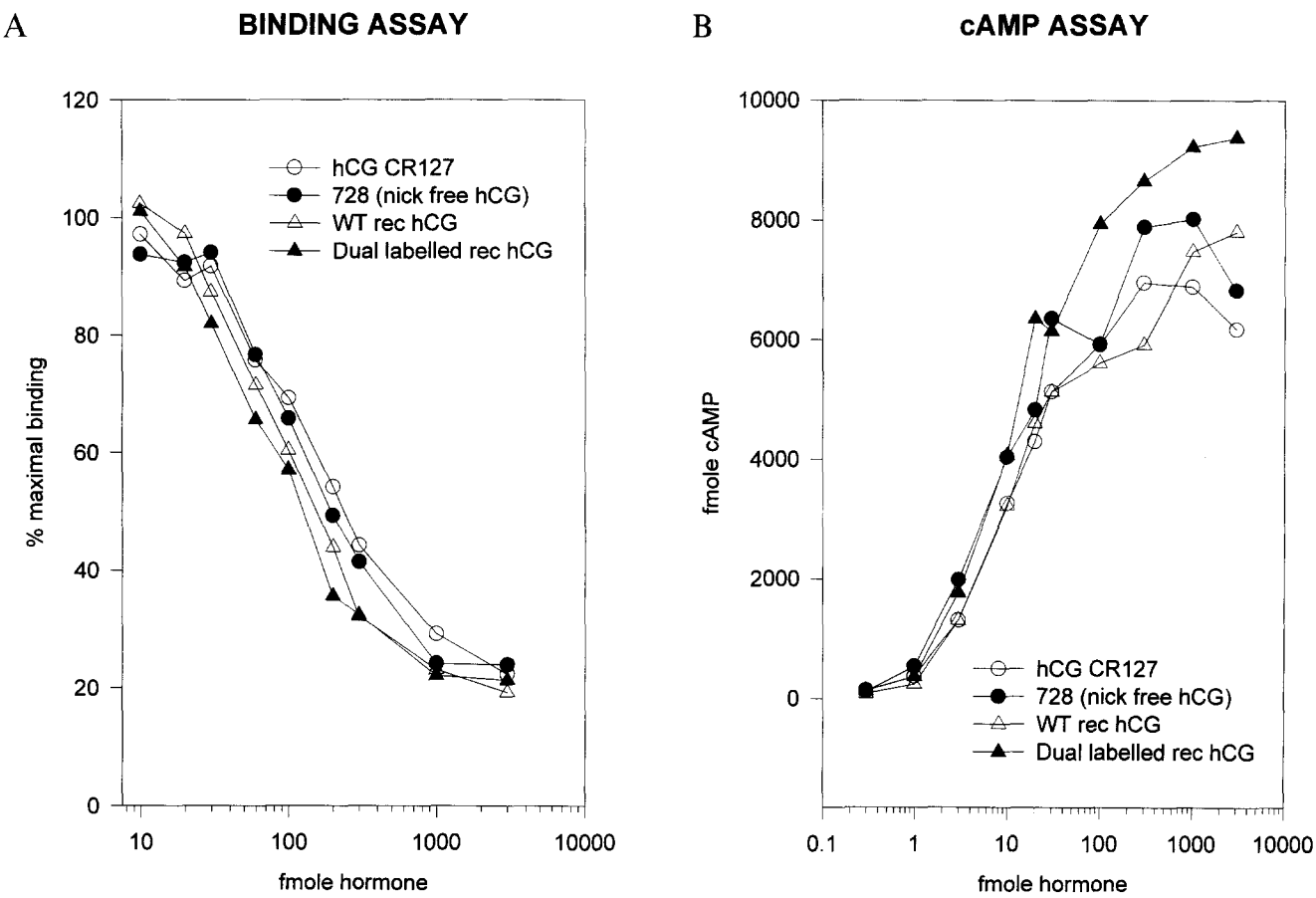

Fig. 2. The panel on the left (A) is a receptor-binding assay using CHO cells expressing the rat LH receptor. The hormones were incubated with 2×10^5 cells and 50–100 000 cpm of ^{125}I-hCG in a total volume of 300 µl. Maximal binding is the cpm bound without the addition of unlabeled hormone. The panel on the right (B) is a biological activity assay using the same CHO cells; 2×10^4 cells were incubated with varying amounts of hormone in a total volume of 200 µl. After this stimulation period, total production of cAMP was measured using a commercial kit. Hormones tested were: reference preparation hCG CR127, nick-free CR127, preparation 728, expressed hCG-designated wild-type and isotope-labeled expressed hCG (dual-labeled rec-hCG). Quantities on the X-axis are from amino-acid-analyzed calibrated solutions of hormones.

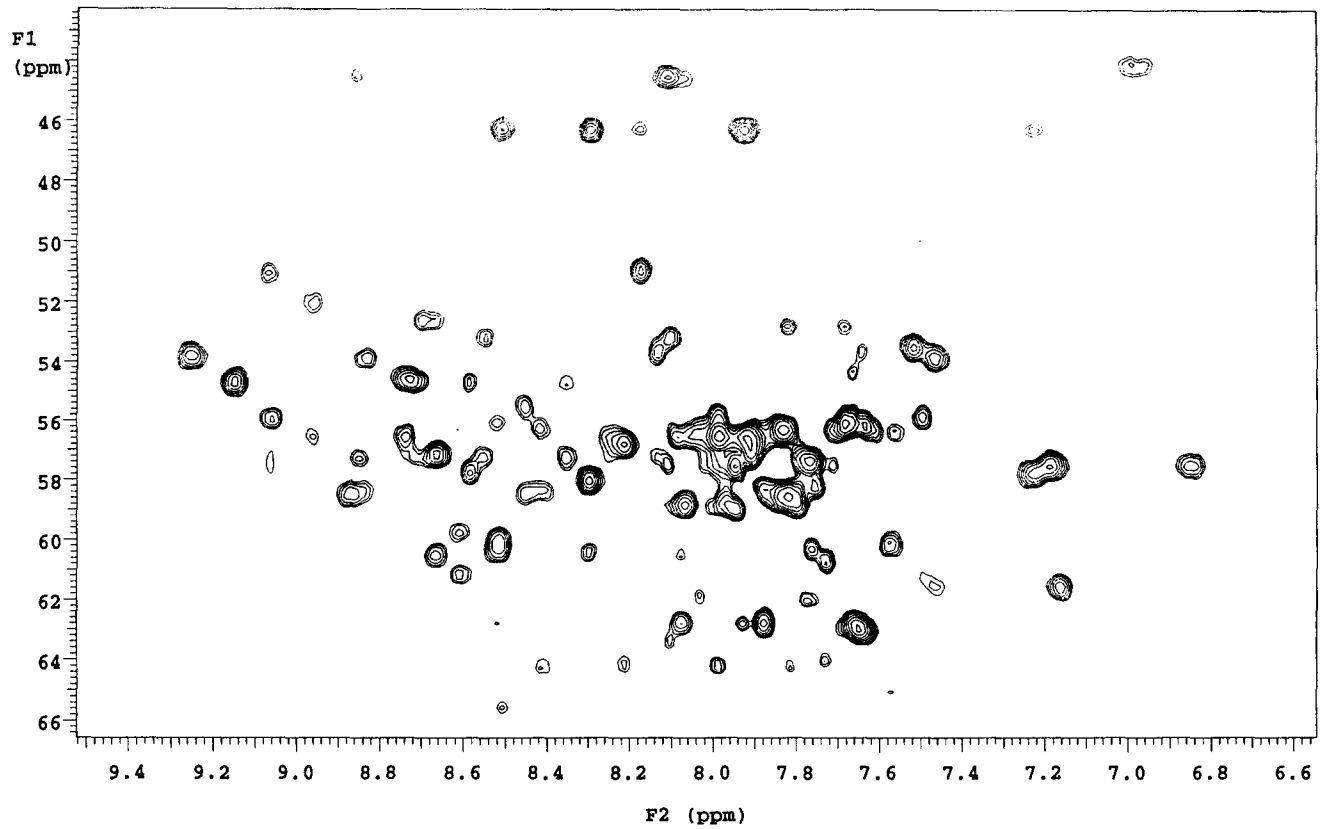

Fig. 3. A two-dimensional analogue of the three-dimensional constant-time HNCA experiment was recorded according to Grzesiek and Bax (1992) with the N15 (t_1) incremental period set to zero. The sample examined in this study contained 5.7 mg of ^{13}C,^{15}N-labeled α-subunit in 0.7 ml (0.55 mM protein concentration) of H_2O/D_2O (90:10), at pH 5.5 (512 scans/t_1 increment). Data were acquired with spectral widths of 3000 Hz in both the $^{13}C^{\alpha}$ (F1) and ^{1}H (F2) dimensions, using 128 and 1024 complex points, respectively. Prior to two-dimensional Fourier transformation, data were apodised with squared Gaussian functions in each dimension; about 95% of the expected intraresidue connectivities were actually observed.

standard solutions) were calculated for the receptor-binding assay (CV = 100 × standard error/parameter value): CR127 146.2 ± 17.3 (%CV = 12.3), nick-free CR127 130.8 ± 11.6 (% CV = 8.9), recombinant-hCG 84.1 ± 9.5 (%CV = 11.2) and NMR-labeled hCG 64.6 ± 8.7 (%CV = 13.5).

Mass spectrometry

Mass spectrometric analysis of the unfractionated hydrolysis products of the unlabeled and labeled hCG α-subunit produced complex mass spectra. The incorporation of ^{13}C and ^{15}N could be unambiguously assigned for only two amino acids: proline and serine. The average incorporation of ^{13}C and ^{15}N in proline and serine was 96 ± 10% and 95 ± 10%, respectively.

The incorporation of ^{13}C and ^{15}N in the hCG α-subunit was also determined by mass spectrometric analysis of three unlabeled and labeled contiguous non-glycosylated tryptic fragments comprising more than half of the α-subunit, α_{1-35}, α_{36-42} and α_{45-51}. The first fragment, α_{1-35}, represents one-third of the mass of the α-subunit. The analysis yielded ^{13}C and ^{15}N incorporation levels of 94 ± 2% for this large fragment α_{1-35}; 75 ± 8% for fragment α_{36-42} and 93 ± 3% for fragment α_{45-51}. Since the unexpectedly low incorporation of the ^{13}C and ^{15}N in the α_{36-42} peptide could not be explained, the study was repeated with another preparation of labeled hormone. In the second study, ^{13}C and ^{15}N incorporation for these three peptides were 90 ± 3% for α_{1-35}, 95.9 ± 2% for α_{36-42} and 98.8 ± 2% for α_{45-51}. We conclude that the hormone was at least 90% uniformly labeled.

NMR analysis

Preliminary data on the labeled hCG α-subunit are shown in Fig. 3, which illustrates a two-dimensional HNCA spectrum, correlating the intraresidue H^{α} and C^{α} shifts for each residue, via the N-C^{α} scalar coupling. In addition, a number of correlations between H^{α} and C^{α} of the (n-1)th residue were observed via the two-bond N-(CO)-C^{α} scalar coupling. Because of the relatively small one-bond ^{15}N-C^{α} scalar coupling (8–12 Hz), this experiment is quite insensitive for proteins of moderate molecular mass (Grzesiek and Bax, 1992). However, in the three-dimensional version of this experiment, over 95% of the expected connectivities were observed (data not shown). The level of isotopic enrichment is uniform and adequate for a full structural analysis of the solution structure of this protein.

Discussion

The NMR technique for solution of protein structures has evolved rapidly in recent years and now is capable of solving the structures of ever-larger proteins. It may soon be possible to solve the structures of proteins in excess of 30 000 MW with this technique in conjunction with NMR isotopic labeling (Fogh et al., 1994; Remerowski et al., 1994). Such labeling is an absolute requirement for solution of the structures of glycoproteins, since the sugar resonances prevent identification of many of the amino acid side chains.

The protein described in this report, hCG, is a glycoprotein of which the chemically modified structure has previously been solved by X-ray crystallography. That structural solution raised a number of important questions which can only be answered by NMR structural analysis. For example, the crystallographic solution of the structure of HF-treated hCG disagreed with most of the disulfide-bridge pairings determined by chemical means (Mise and Bahl, 1980,1981; Lapthorn et al., 1994; Wu et al., 1994; Lustbader et al., 1995), raising the possibility that some disulfide bridges may have rearranged during the crystallization process. In fact, crystal growth is accelerated in the presence of a reducing agent, making such disulfide interchange a distinct possibility (Lustbader et al., 1995). HF-treated hCG may differ in structure from natural hCG: it lacks biological activity (Chen et al., 1982; Manjunath and Sairam 1983; Sairam and Manjunath, 1983); its subunits have highly increased binding affinity for one other (Manjunath and Sairam 1983; Sairam and Manjunath, 1983); and some antibodies detect subtle conformational alterations in this form of hCG (Sairam and Manjunath, 1982; Manjunath and Sairam 1983; Keutmann et al., 1985; Sairam et al., 1988,1990). Other open questions include: (i) is the crystallographic solution of the structure of HF-treated hCG the same as the NMR solution of the structure of native hCG? If different, the structure of HF-hCG would also be examined by NMR as a potential template of an antagonist structure; (ii) what is the orientation of the carbohydrate groups on hCG, and do they interact with the polypeptide chain; (iii) what are the structures of the isolated hormone subunits which could not be crystallized? The two noncovalently bound subunits of hCG undergo conformational alterations upon association and dissociation (Strickland and Puett, 1981).

Development of the labeling medium required manipulation of the amino acid and sugar content obtainable from isotopically labeled algal biomass. It was found that replacement of glutamic acid with glutamine and the removal of aspartic acid significantly increased expression of hCG. Despite the fact that the cells then had to synthesize aspartic acid and asparagine de novo from glutamic acid and/or glucose, expression improved in the absence of aspartic acid. Thus, biosynthesis does not seem to be a rate-limiting step, since after two media passages mass spectrometry demonstrated that peptide α_{1-35} containing two aspartic acids, was 95% labeled. Although the labeling medium was developed using unlabeled amino acids, switching to ^{13}C,^{15}N-labeled amino acids, glucose and pyruvate did not decrease expression of hCG by the CHO cells, indicating that the cells could incorporate the heavy isotopes without a decrease in biosynthetic rate.

Mass spectrometric studies were used to assess the extent of incorporation of the isotopic labels. We purified three peptides, comprising more than half of the mass of the α-subunit, from both unlabeled and labeled hormone: α_{1-35}, α_{36-42} and α_{45-51}. These studies indicated that over 90% isotopic substitution had taken place. Preliminary NMR studies support the conclusion of such uniform labeling of the protein (S. Homans, personal communication).

The subunits produced in this study from the expressed, isotopically labeled hCG were made by dissociation in guanidine followed by reversed-phase HPLC using the commonly employed TFA/acetonitrile system. We found that, while only a small amount of sialic acid was released from hCG during HPLC chromatography, residual TFA in the dried material continued to hydrolyze the glycosidic linkage between sialic acid and galactose. Thus, it is desirable to conduct the acid chromatography step as rapidly as possible and then neutralize and desalt the material to prevent a significant loss of sialic acid from the glycoprotein. Alternatively, a nonvolatile buffer can be used for chromatography, such as that used by Renwick and colleagues for sugar studies (Weisshaar et al., 1991), followed by desalting on gel filtration.

Immunochemical analysis for binding to both anti-native α- and anti-reduced, carboxymethylated α-rabbit sera was used to test for differences between the urinary hCG α-subunit produced earlier by the established urea dissociation and ion-exchange separation procedure (Birken et al., 1990) and the α-subunits prepared in the current study by reversed-phase HPLC in acid buffer. The immunological potencies of both the expressed, reversed-phase separated α-subunit and the urinary α-subunit reference preparations were the same with both types of antisera. Finally, a biological-activity-recovery–subunit-combination study, comparing the urinary and isotope-labeled recombinant α-subunit preparations, also showed that the two preparations were indistinguishable. Both α-subunits were found to combine with the urinary hCG β-subunit to the same extent under similar conditions, and both exhibited identical biological activity after combination. Therefore, the reversed-phase purification of the subunits did not detectably effect their structures and we conclude that the isotope-labeled α-subunit, and most likely the β-subunit as well, is native in its three-dimensional structure. The isotope-labeled α-subunit will be the first of the hCG molecules of which the structure will be solved using the labeled material described here.

Conclusions

We have developed an isotopically labeled medium and used it to label the glycoprotein hormone hCG during its synthesis by CHO cells. Preliminary NMR information, already accumulated on the purified labeled α-subunit, demonstrates that the labeling of the protein is adequate and that the structure can be solved by the NMR technique. This labeling procedure represents an advance in NMR labeling technology for mammalian posttranslationally modified proteins. Indeed, by this method it should be possible to determine the conformation of the carbohydrate groups on the surface of hCG and assess any interactions with the protein backbone. The labeling medium is commercially available and the procedure for its use in protein expression by CHO cells is presented in this report. It should now prove possible to solve the structures of many medium-sized glycoproteins which cannot be crystallized in their natural state, such as the glycoprotein described in this report.

Acknowledgements

We thank Dr. Steven Homans of St. Andrews University, who generously performed the NMR spectral study included in this report. We are especially grateful to Martek Biosciences for providing the labeled amino acids necessary for our study. This work was supported in part by NIH Grants HD15454 (to J.W.L., S.B., S.P. and R.E.C.) and RR00862 (to B.T.C.).

References

Bellisario, R., Carlsen, R.B. and Bahl, O.P. (1973) *J. Biol. Chem.*, **248**, 6796–6809.

Birken, S. and Canfield, R.E. (1982a) In *Tumor Imaging: The Radioimmunological Detection of Cancer* (Eds, Burchiel, S.W. and Rhodes, B.A.), Masson Publishing USA, Inc., New York, NY, pp. 41–52.

Birken, S., Canfield, R.E., Agosto, G. and Lewis, J. (1982) *Endocrinology*, **110**, 1555–1563.

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

Birken, S., Krichevsky, A., O'Connor, J., Lustbader, J. and Canfield, R. (1990) In *Glycoprotein Hormones* (Eds, Chin, W.W. and Boime, I.), Serono Symposia, Norwell, MA, pp. 45–61.

Birken, S., Gawinowicz, M.A., Kardana, A. and Cole, L.A. (1991) *Endocrinology*, **129**, 1551–1558.

Birken, S., Chen, Y., Gawinowicz, M.A., Lustbader, J.W., Pollak, S., Agosto, G., Buck, R. and O'Connor, J. (1993) *Endocrinology*, **133**, 1390–1397.

Carlsen, R.B., Bahl, O.P. and Swaminathan, N. (1973) *J. Biol. Chem.*, **248**, 6810–6827.

Chen, H.-C., Shimohigashi, Y., Dufau, M.L. and Catt, K.J. (1982) *J. Biol. Chem.*, **257**, 14446–14452.

Chowdhury, S.K., Katta, V. and Chait, B.T. (1990) *Rapid Commun. Mass Spectrom.*, **4**, 81–87.

Daopin, S., Piez, K.A., Ogawa, Y. and Davies, D.R. (1992) *Science*, **257**, 369–373.

DeLean, A., Munson, P.J. and Rodbard, D. (1978) *Am. J. Physiol.*, **235**, E97–102.

Fogh, R.H., Schipper, D., Boelens, B. and Kaptein, R. (1994) *J. Biomol. NMR*, **4**, 123–128.

Grzesiek, S. and Bax, A. (1992) *J. Magn. Reson.*, **96**, 432–440.

Hancock, W.S. (Ed.) (1977) *Specifications of Amino Acids*, 4th ed., Ajinomoto, Japan, p. 70.

Hansen, A.P., Petros, A.M., Mazar, A.P., Pederson, T.M., Rueter, A. and Fesik, S.W. (1992) *Biochemistry*, **31**, 12713–12718.

Keene, J.L., Matzuk, M.M., Otani, T., Fauser, B.C.J.M., Galway, A.B., Hseuh, A.J.W. and Boime, I. (1989) *J. Biol. Chem.*, **264**, 4769–4775.

Keutmann, H.T., Johnson, L. and Ryan, R.J. (1985) *Febs Lett.*, **185**, 333–338.

Laemmli, U.K. (1970) *Nature*, **227**, 680–685.

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.

Lustbader, J.W., Wu, H., Birken, S., Pollak, S., Gawinowicz Kolks, M.A., Pound, A.M., Austen, D., Hendrickson, W.A. and Canfield, R.E. (1995) *Endocrinology*, **136**, 640–650.

Matzuk, M., Keene, J.L. and Boime, I. (1989) *J. Biol. Chem.*, **264**, 2409–2414.

McDonald, N.Q., Lapatto, R., Murray-Rust, J., Gunning, J., Wlodawer, A., Blundell, T.L. (1991) *Nature*, **354**, 411–414.

Mise, T. and Bahl, O.P. (1980) *J. Biol. Chem.*, **255**, 8516–8522.

Mise, T. and Bahl, O.P. (1981) *J. Biol. Chem.*, **256**, 6587–6592.

Morgan, F.J., Birken, S. and Canfield, R.E. (1973) *Mol. Cell. Biochem.*, **2**, 97–99.

Morgan, F.J., Birken, S. and Canfield, R.E. (1975) *J. Biol. Chem.*, **250**, 5247–5258.

Manjunath, P. and Sairam, M.R. (1983) *J. Biol. Chem.*, **258**, 3554–3558.

O'Connor, J.F., Birken, S., Lustbader, J.W., Krichevsky, A., Chen, Y. and Canfield, R.E. (1994) *Endocr. Rev.*, **15**, 650–683.

Oefner, C., D'Arcy, A., Winkler, F.K., Eggimann, B. and Hosang, M. (1992) *Embo J.*, **11**, 3921–3926.

Osborn, B.L. and Abramson, F.P. (1995) *Anal. Biochem.*, **229**, 347–350.

Remerowski, M.L., Domke, T., Groenewegen, A., Pepermans, H.A.M., Hilbers, C.W. and Van de Ven, F.J.M. (1994) *J. Biomol. NMR*, **4**, 257–278.

Sairam, M.R. and Manjunath, P. (1982) *Int. J. Pept. Protein Res.*, **19**, 315–320.

Sairam, M.R. and Manjunath, P. (1983) *J. Biol. Chem.*, **258**, 445–449.

Sairam, M.R., Linggen, J. and Bhargavi, G.N. (1988) *Biosci. Rep.*, **8**, 271–278.

Sairam, M.R., Linggen, J., Sairam, J. and Bhargavi, G.N. (1990) *Biochem. Cell Biol.*, **68**, 889–893.

Schlunegger, M.P., Grutter, M.G. (1992) *Nature*, **358**, 430–434.

Strickland, T.W. and Puett, D. (1981) *Endocrinology*, **109**, 1933–1942.

Weisshaar, G., Hiyama, J. and Renwick, G.C. (1991) *Glycobiology*, **1**, 393–404.

Wold, F. (1981) *Annu. Rev. Biochem.*, **50**, 783–814.

Wold, F. and Moldave, K. (1984) *Methods Enzymol.*, **106**, 29–58.

Wray, W., Bolikas, T., Wray, V.P. and Hancock, R. (1981) *Anal. Biochem.*, **118**, 197–203.

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

YMC Application No. 10132 Technical Report (1994), YMC Inc., Wilmington, NC.