Pflumm, M., Luchins, J., & Beychok, S. (1986) *Methods Enzymol. 130*, 519-534.
Ptitsyn, O. B. (1987) *J. Protein Chem. 6*, 273-293.
Rao, S. T., & Rossmann, M. G. (1973) *J. Mol. Biol. 76*, 241-256.
Richardson, J. S. (1977) *Nature 268*, 495-500.
Roder, H., Elöve, G. A., & Englander, S. W. (1988) *Nature 335*, 700-704.
Rossmann, M. G., & Argos, P. (1981) *Annu. Rev. Biochem. 50*, 497-532.
Sears, D. W., & Beychok, S. (1973) in *Physical Principles and Techniques of Protein Chemistry, Part C* (Leach, S. J., Ed.) pp 445-593, Academic Press, New York.
Sugawara, T., Kuwajima, K., & Sugai, S. (1991) *Biochemistry 30*, 2698-2706.
Touchette, N. A., Perry, K. M., & Matthews, C. R. (1986) *Biochemistry 25*, 5445-5452.
Udgaonkar, J. B., & Baldwin, R. L. (1988) *Nature 335*, 694-699.
Woody, R. W. (1977) *J. Polymer Sci.: Macromol. Rev. 12*, 181-321.
Woody, R. W. (1978) *Biopolymers 17*, 1451-1467.
Yang, J. T., Wu, C.-S., & Martinez, H. M. (1986) *Methods Enzymol. 130*, 208-269.

Self-Association of Human and Porcine Relaxin As Assessed by Analytical Ultracentrifugation and Circular Dichroism

Steven J. Shire,*,‡ Leslie A. Holladay,§ and Ernst Rinderknecht∥

Department of Pharmaceutical Research and Development and Department of Recovery Process Research and Development, Genentech, Inc., South San Francisco, California 94080, and Department of Research, Beckman Instruments Inc., Palo Alto, California 94303-0803

Received February 8, 1991; Revised Manuscript Received May 17, 1991

ABSTRACT: The self-association properties of recombinant DNA derived human relaxin, and porcine relaxin isolated from porcine ovaries, have been studied by sedimentation equilibrium analytical ultracentrifugation and circular dichroism (CD). The human relaxin ultracentrifuge data were adequately defined by a monomer–dimer self-association model with an association constant of $\sim 6 \times 10^5\ M^{-1}$, whereas porcine relaxin was essentially monomeric in solution. An approximate 5-fold increase in weight fraction of human relaxin monomer elicited by dilution of the protein resulted in no change in the far-UV CD spectrum at 220 nm. In contrast, after the same increase in weight fraction of monomer, the near-UV circular dichroism spectra for human relaxin exhibited a significant decrease in the amplitude for the CD bands near 277 and 284 nm. These CD bands, which may be assigned to the lone tyrosine in human relaxin, are superimposed on a broad envelope that is probably due to the three disulfide chromophores. Although both the human and porcine proteins contain two tryptophan residues, the near-UV CD spectra exhibit only a broad shoulder near 295 nm rather than the strong CD bands often found for tryptophan. Moreover, there is little change in this broad band after dilution of human relaxin to concentrations that resulted in a 4-fold increase in monomer weight fraction. These data suggest that dissociation of the human relaxin dimer to monomer is not accompanied by large overall changes in secondary structure or alteration in the average tryptophan environment, whereas there is a significant change in the tyrosine environment. This tyrosine is replaced by an arginine in porcine relaxin, and this substitution may affect the self-association properties of this protein.

Relaxin is a protein hormone that plays a major role in the reproductive biology of various species (Sherwood, 1988; Bryant-Greenwood, 1982). This protein appears to modulate the restructuring of connective tissues in target organs in order to generate the required changes in organ structure during pregnancy and parturition. In particular, relaxin regulates a number of biological responses of reproductive tissues in pregnant animals including softening of cervical tissue and inhibition of uterine myometrial contractions (Steinetz et al., 1959). Some of the important roles for relaxin as a pregnancy hormone include inhibition of premature labor and induction of cervical ripening prior to parturition.

The latter activity has been investigated in a clinical setting with use of relaxin isolated from porcine corpus lutea and ovaries (Evans et al., 1983; MacLennan et al., 1981, 1986). The preparation of porcine relaxin has been well documented (Sherwood & O'Byrne, 1974; Bullesbach & Schwabe, 1985, 1986), and since the sow ovary serves as a relatively plentiful source of relaxin, most studies on this hormone have been conducted with porcine relaxin. Identification of the human gene responsible for expression of human relaxin (Hudson et al., 1984) has resulted in the cloning, expression, and purification of recombinant DNA derived human relaxin (rhRlx) that is active in a cyclic AMP response assay using human endometrial cells (Fei et al., 1990).

Relaxin consists of two polypeptide chains that are linked by inter- and intrachain disulfide bonds. Although the amino acid sequence homology between relaxin and insulin is low (Figure 1), there are several important homologies between these protein hormones. The positions of the disulfide links are similar in both proteins, and there is a conservation of the glycine residues immediately adjacent to the cysteine residues in the B chain (Sherwood, 1988). These residues are crucial to maintaining the proper structural folding for insulin and

‡ Department of Pharmaceutical Research and Development, Genentech, Inc.
§ Department of Recovery Process Research and Development, Genentech, Inc.
∥ Department of Research, Beckman Instruments Inc.

0006-2960/91/0430-7703$02.50/0

A chain:	1	5	10	15	20	25
human relaxin:	gln leu tyr ser ala leu ala asn lys cys cys his val gly cys thr lys arg ser leu ala arg phe cys					
porcine relaxin:	arg met thr leu ser glu lys cys cys gln val gly cys ile arg lys asp ile ala arg leu cys					
porcine insulin:	gly ile val glu gln cys cys thr ser ile cys ser leu tyr gln leu glu asn tyr cys asn					

B chain:	1	5	10	15	20	25	30	35
human relaxin:	asp ser trp met glu glu val ile lys leu cys gly arg glu leu val arg ala gln ile ala ile cys gly met ser thr trp ser							
porcine relaxin:	gln ser thr asn asp phe ile lys ala cys gly arg glu leu val arg leu trp val glu ile cys gly ser val ser trp gly arg thr ala leu							
porcine insulin:	phe val asn gln his leu cys gly ser his leu val glu ala leu tyr leu val cys gly glu arg gly phe phe tyr thr pro lys ala							

FIGURE 1: Amino acid sequence of A and B chains of porcine and human relaxin and porcine insulin.

suggest the possibility that relaxin may have a similar tertiary structure. Molecular modeling techniques have been used to demonstrate that the crystallographic structure of insulin can accommodate the sequence of porcine relaxin (Bedarkar et al., 1977; Isaacs et al., 1978). Although there are these striking similarities between insulin and relaxin, the proteins exhibit very different physiological and biochemical properties. The aggregation properties of insulin are well-known (Blundell et al., 1971; Hodgkin & Mercola, 1972), and it is thought that this property is related to the way insulin is stored and regulated (Blundell et al., 1972). Unlike insulin, analytical ultracentrifugation studies of porcine relaxin suggest that this protein does not exhibit self-association behavior (Jeffrey & Coates, 1966; Goldman & Carpenter, 1974; Sherwood & O'Byrne, 1974). Although porcine and human relaxin exhibit similar biological function, the two proteins are only 40% homologous (Figure 1) as compared to porcine and human insulin, which display ~98% sequence homology. In this paper, we have compared by analytical ultracentrifugation and circular dichroism the association properties of recombinant DNA derived human relaxin with those of porcine relaxin isolated from porcine ovaries.

Materials and Methods

Preparation and Purification of Relaxin

Porcine Relaxin. Porcine relaxin was isolated from sow ovaries and was kindly supplied to us by Dr. Geoffrey Tregear of the Howard Florey Institute.

Recombinant DNA Derived Human Relaxin. The A and B chains of human relaxin were expressed in *Escherichia coli* and purified by a combination of gel permeation, ion-exchange, and reversed-phase chromatography to >90% purity as judged by analytical reversed-phase chromatography. The purified A and B chains were then combined to form human relaxin that was purified by ion-exchange and reversed-phase chromatographic techniques. The final product, which is active in a cyclic AMP response assay (Fei et al., 1990) is greater than 95% homogeneous as determined by analytical reversed-phase chromatography (inset in Figure 4B). The product was analyzed by amino acid analysis, tryptic mapping, and amino acid sequencing and was shown to have the expected primary structure. The disulfide bond arrangement was shown by tandem mass spectrometry to match that of insulin (Stults et al., 1990).

Reversed-Phase HPLC[1] *of Relaxin.* Analysis was performed by injecting 7 μL of relaxin at 1 mg/mL onto a Vydac C4-300 reversed-phase column equilibrated at 35 °C. The initial equilibration buffer was 18% acetonitrile, 0.1% TFA. After a 1-min hold at 18% acetonitrile, the elution was carried out by running a linear acetonitrile gradient to a final concentration of 45% acetonitrile after 30 min. The flow rate was held constant at 1.0 mL/min, and the absorbance was monitored at 214 nm.

Extinction Coefficient of Relaxin. The molar extinction coefficient and absorptivity for human relaxin was determined by use of a modification of the method of Bewley (1982). UV absorption spectra of relaxin from 400 to 240 nm were obtained before and after proteolysis overnight at 37 °C with a 1:1 mixture of thermolysin and trypsin (protease:relaxin at 1:300). A model UV spectrum was obtained from the amino acid composition and molecular weight for human relaxin. The model spectrum is a linear combination of experimentally determined spectra of the model compounds *N*-acetyl X amide, where X is tryptophan, tyrosine, or phenylalanine, and the disulfide spectrum from glutathione. The model extinction coefficient should be equivalent to that of a proteolytically digested sample, assuming all influence of tertiary and secondary structure on UV absorption has been eliminated by the proteolysis (Bewley, 1982). The molar extinction coefficient of the native folded protein at the wavelength of maximum absorbance, E_M, can be related to the model molar extinction coefficient, $E_{M,model}$, and the OD before and after digestion as

$$E_M = OD_{predigest}(E_{M,model})/OD_{digest}$$

The value of the absorptivity at 280 nm, 2.25 cm^{-1} $(mg/mL)^{-1}$, was computed from E_M by dividing by a molecular weight of 5965 and was in excellent agreement with a value of 2.30 cm^{-1} $(mg/mL)^{-1}$ obtained by quantitative amino acid analysis. An absorptivity of 1.88 was estimated for porcine relaxin from the model protein spectrum method.

All spectra for protein concentration determinations were obtained from 400 to 240 nm using a Uvicon 860 spectrophotometer and corrected for light scattering by the method of Englander and Epstein (1957). This method involves plotting the log of the absorbance from 400 to 330 nm vs the log of the wavelength and extrapolating the linear regression of this plot to 280 nm in order to estimate the absorbance due to scattering.

Circular Dichroism. Circular dichroism (CD) spectra were obtained for relaxin at 0.5 mg/mL from 250 to 190 nm (far UV) in an Aviv/Cary 60-DS spectropolarimeter with use of a 0.01-cm cell or at 350–240 nm (near UV) with a 1.0-cm cell. The resulting spectra were converted to mean residue weight ellipticities in the usual manner (Adler et al., 1973) with use of a mean residue weight of 113 for both human and porcine relaxin. Estimates of secondary structure were obtained from data collected in the far-UV range by use of the method of either Provencher (Provencher & Glockner, 1981) or Yang and co-workers (Chang et al., 1978). In the case of CD measurements of dilute solutions (10 or 20 μg/mL), cylindrical cells of 1 and 10 cm were used in the far- and near-UV spectral

[1] Abbreviations: CD, circular dichroism; HPLC, high-performance liquid chromatography; TFA, trifluoroacetic acid; OD, optical density.

range, respectively. Concentrations were obtained from gravimetrically diluted samples. Adsorptive losses to glass surfaces at these low concentrations were minimal as determined by UV absorption spectroscopy using a 10-cm cylindrical cuvette.

Analytical Ultracentrifugation

Experimental Methods. Human relaxin in 10 mM citrate buffer, ionic strength of 0.154 M (made up with NaCl) at pH 5, was diluted to concentrations ranging from 0.05 to 0.20 mg/mL with the citrate buffer. Porcine relaxin obtained from the Howard Florey institute as a lyophilized powder was prepared in the same buffer system and exhaustively dialyzed against the citrate buffer with use of dialysis tubing with a molecular weight cutoff of 1000 (Spectrapor 1).

Relaxin samples were loaded into charcoal-filled Epon six-channel Yphantis cells (Yphantis, 1964) and placed into an An-F four-hole analytical rotor. Centrifugation was performed in a Beckman Model E analytical ultracentrifuge at either 22000 or 32000 rpm at 19–21 °C for 18 h. In separate experiments it was determined that 18 h was sufficient time to obtain sedimentation equilibrium. The concentration gradient in the cell was determined by UV absorption at 280 nm using a photoelectric optical scanner equipped with a stepping motor. The radial position and photomultiplier tube voltage were digitized and saved as text files on a Macintosh SE computer with use of commercially available data acquisition boards and software (Strawberry Tree Inc.). The photomultiplier tube voltages at each rotor speed were correlated with actual OD values at 280 nm by using tryptophan solutions at known OD as determined on an 8451A Hewlett-Packard diode array spectrophotometer. The partial specific volumes for porcine and human relaxin, 0.732 and 0.741, respectively, were computed from the amino acid composition by use of the additivity rule and values for the individual amino acid residues (Cohn & Edsall, 1965).

Analysis of Sedimentation Data. The total concentration, C_r, of reversibly self-associating protein at any radial position, r, after attaining sedimentation equilibrium can be related to the monomer molecular weight, M_1, and the number of aggregates, n, by

$$C_r = \sum_{i=1}^{n+1} C_{i_0} e^{L(i)[AM_1(r^2-r_0^2)-BM_1(C_r-C_{r0})]} \quad (1)$$

where $L(i)$ is the degree of association for the ith aggregate at concentration C_{i_0} at radial reference position r_0, C_{r_0} is the total protein concentration at r_0, and B is the second virial coefficient (Haschemeyer & Bowers, 1970; Holladay & Sophianopoulos, 1972). The parameter A is defined as

$$A = (1 - \bar{V}\rho)\omega^2/2RT \quad (2)$$

where R is the gas constant, T is the absolute temperature, $\bar{V}$ is the partial specific volume, ρ is the solution density, and ω is the angular velocity. All concentrations in these and subsequent expressions are on a mass per volume scale. It is important to note that no a priori constraint exists as to the choice of the reference radial distance r_0.

A series of association constants, k_i, can be defined, where k_1 is unity and

$$k_i = C_i/(C_1)^{L(i)} \quad (3)$$

Two self-association schemes were used in the data analysis.

Indefinite Self-Association. In the special case where the free energy change is the same for the addition of a monomer to any aggregate, an intrinsic association constant, k, may be defined by use of the relationship (Adams & Lewis, 1968)

$$C_i = ik^{(i-1)}(C_1)^i \quad (4)$$

Provided that kC_1 is less than one, the total concentration is then related to the monomer concentration by

$$C = C_1/(1 - kC_1)^2 \quad (5)$$

where

$$C_1 = C_{1_0} e^{\{AM_1(r^2-r_0^2)-B'B'M_1(C_r-C_{r0})\}} \quad (6)$$

In order to constrain the value of B to be nonnegative, a substitute fitting parameter, B', was used where B' is the square root of B.

In computing the total concentration for the indefinite association case, the total concentration at the reference distance r_0 is computed as

$$C_{r_0} = C_{1_0}/(1 - kC_{1_0})^2 \quad (7)$$

Then the monomer concentration at any radial distance is computed from eq 6 and the experimental value for C_r. As the fitting process moves the computed C_r closer to the experimental C_r, the differences between these two quantities become small and the replacement of the theoretical C_r by the experimental C_r in the right side of eq 6 removes the necessity to do an iterative computation of C_r for each change in the fitting parameters (Holladay & Sophianopoulos, 1972). From the monomer concentration and the current estimate for k, the total protein concentration is computed with eq 5. The three parameters sought from single-cell data are the monomer concentration at the reference radial position, the intrinsic association constant k, and the square root of B, B'.

Discrete Self-Association. In the case of a discrete self-association model, the natural logarithms of the equilibrium constants are used as the fitting parameters rather than the equilibrium constants themselves. An examination of eq 1 shows two reasons to fit for the natural logarithms of the equilibrium constants (Johnson et al., 1981). First, a fit for the natural logarithms of the association constants is appropriate since the parameter of real interest is the free energy change on association. Second, fitting for the ln k_i values removes a need to constrain the C_{i_0} values to be nonnegative. The sought parameters, *ln* k_i values are free to be either positive or negative without need for constraint. For a model involving a monomer and n aggregates, there are n different ln k_i values, one monomer concentration C_{1_0} per cell/compartment pair, and the value of B as fitting parameters.

The parameter C_r is implicit in eq 1 in that C_r appears in the exponential term. In the fitting process, the experimental value for C_r is used in the right side of eq 1 and the value for C_{r_0} used to compute the difference term, $C_r - C_{r_0}$, is determined from the equation

$$C_{r_0} = \sum_{i=1}^{n+1} e^{[\ln(k_i)+L(i)\ln(C_{10})]} \quad (8)$$

The difference in total protein concentration, $C_r - C_{r_0}$, at radial position r and reference position r_0 is computed from the experimental C_r, current estimates for the association constants, and the monomer concentration at reference distance r_0. The fitted function C_r is then computed as

$$C_r = \sum_{i=1}^{n+1} e^{\{\ln(k_i)+L(i)[\ln(C_{10})+AM_1(r^2-r_0^2)-B'B'M_1(C_r-C_{r0})]\}} \quad (9)$$

Note this since k_1 is equal to 1, ln k_1 becomes 0, and the first term in the sum is simply the monomer concentration.

Fitting Algorithm. For multichannel sedimentation equilibrium experiments, each solution channel is loaded with relaxin at a different protein concentration. Equations such as eq 9 may then be written for each channel that differ only in the values for C_{1_0}, the monomer concentration at the ref-

Table I: Analysis of Sedimentation Equilibrium Data Using the Simplex Algorithm

model	association constants	rmslf
rhRlx in 10 mM citrate, 0.13 M NaCl, pH 5.0		
	32 000 rpm	
monomer–dimer	$K_2 = 100 \pm 4$ (g/L)$^{-1}$	0.0040 ± 0.0003
monomer–dimer–trimer	$K_2 = 104 \pm 9$ (g/L)$^{-1}$, $K_3 = 1.09 \pm 0.03$ (g/L)$^{-2}$	0.0040 ± 0.0002
monomer–dimer–tetramer	$K_2 = 170 \pm 52$ (g/L)$^{-1}$, $K_4 = 0.1 \pm 0.1$ (g/L)$^{-3}$	0.0040 ± 0.0001
isodesic	$K = 3.0 \pm 0.03$ (g/L)$^{-1}$	0.007 ± 0.001
	22 000 rpm	
monomer–dimer	$K_2 = 77 \pm 19$ (g/L)$^{-1}$	0.005 ± 0.0003
porcine relaxin in 10 mM citrate, 0.13 M NaCl, pH 5.0		
model	association constants	rmslf
	32 000 rpm	
monomer–dimer	$K_2 = 0.7 \pm 0.6$ (g/L)$^{-1}$	0.0043

erence distance. The data for all three channels were simultaneously fitted to yield three estimates for C_{1_0} (one for each channel solvent-solution pair), k for the indefinite association case or n values of ln k_i for the discrete association model, and B', which is squared to yield the estimate for B. The reference distance for the analysis was set at a radial position $^2/_3$ of the solution column height from the solution meniscus.

The data were fitted to these equations with use of the Simplex algorithm (Caceci & Cacheris, 1984; Nelder & Mead, 1965) implemented in Turbo Pascal v5.5. Convergence was considered attained when neither the fitted parameters nor the residual sum of squares changed significantly.

Results and Discussion

HPLC Analysis of Relaxin. Reversed-phase HPLC of rhRlx and porcine relaxin shows that the porcine relaxin is heterogeneous when compared to human relaxin (inset of part A and part B of Figure 4).

It has been demonstrated that different forms of porcine relaxin isolated from sow corpora lutea are equipotent but heterogeneous due to variations in length at the C terminus of the B chain from 28 to 31 amino acids (Niall et al., 1980). The peaks labeled 1–3 have been identified by amino acid analysis and sequencing (Florey Institute). The relaxin in peaks 1 and 2 have full-length A chains of 22 amino acids and B-chain lengths of 31 and 30 amino acids, respectively, whereas relaxin in peak 3 has a B chain of 29 amino acids and an A chain with an additional phenylalanine at the N terminus. It has been postulated that the N-terminal extension of the A chain is an intermediate in prorelaxin–relaxin conversion (Bullesbach & Schwabe, 1985).

Analytical Ultracentrifugation. Concentration gradients in three sectors of a Yphantis analytical ultracentrifuge cell after sedimentation equilibrium was attained are shown for porcine and human relaxin in Figures 2 and 3, respectively. Porcine relaxin data were analyzed as nonideal monomer–dimer association and human relaxin data as nonideal monomer–dimer, monomer–dimer–trimer, monomer–dimer–tetramer, and isodesmic self-association. The resulting parameters from the fits of two separate experiments are summarized as the mean values of the association constants and deviation from the mean in Table I. The root mean square lack of fit (rmslf, Table I) is also presented as a mean value and deviation from the mean. As an example, Figures 2 and 3 also show as a solid line the fit of the data to a nonideal monomer–dimer self-association model as well as the residuals to the fit. The fitted values for the second virial coefficients

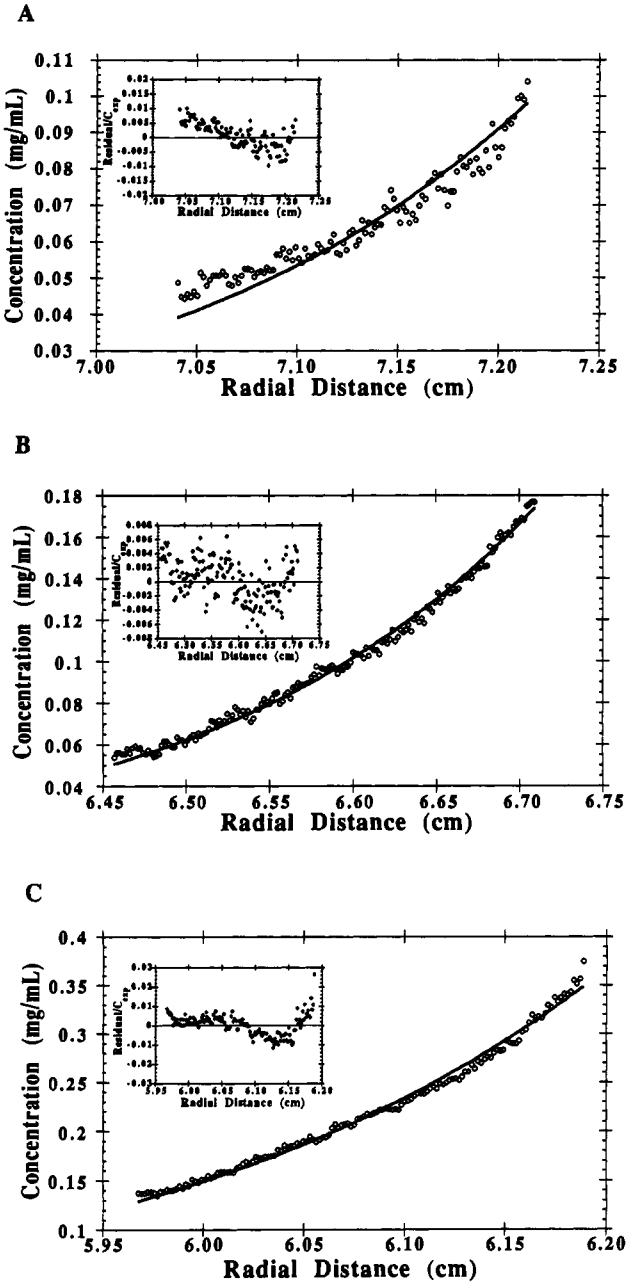


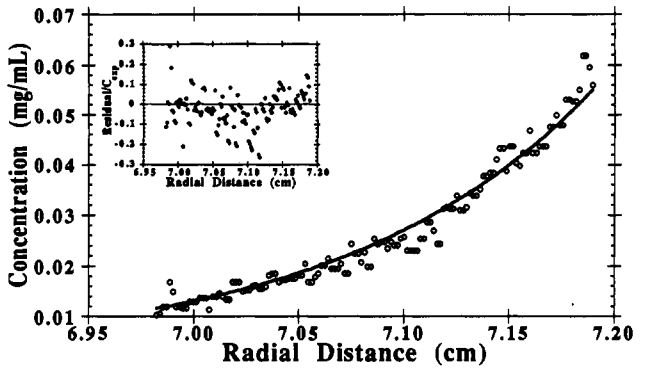
FIGURE 2: Concentration distribution of porcine relaxin during sedimentation equilibrium at 32 000 rpm in an An-F analytical ultracentrifuge rotor. The open circles are measured values in three channel pairs of a Yphantis ultracentrifuge cell. The loading concentrations in channels A, B, and C were 0.05, 0.1, and 0.2 mg/mL, respectively. The solid lines are a result of fitting all three solvent–solution pairs of data simultaneously with a simplex algorithm to a monomer–dimer model. The inset figures show the residuals to the fit, and the resulting parameters are given in Table I.

that were essentially zero (from 5×10^{-5} to 2×10^{-7}) demonstrate that nonideality was not significant during relaxin sedimentation. The contribution from excluded volume effects to the second virial coefficient is usually an order of magnitude less than the effect from the primary charge effect. The primary charge effect contribution to the second virial coefficient, B_{el}, in the presence of a uni-univalent salt is related to the effective charge, Z, on the macromolecule by[2]

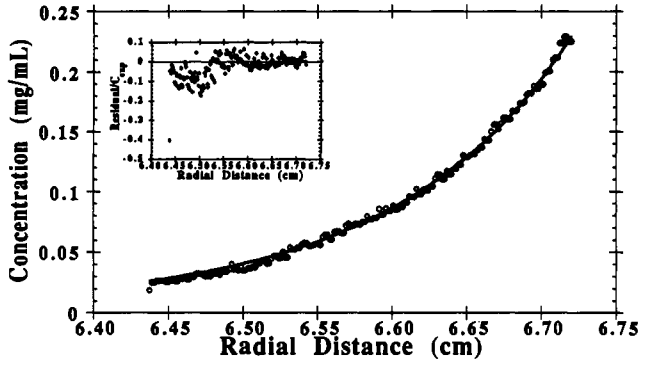
$$B_{el} = Z^2/2M_w^2m \qquad (10)$$

[2] In the fitting algorithm for the centrifuge data, the virial expansion used does not have a factor of 2 associated with the second term of the expansion. Therefore, the value of B_{el} in eq 10 is twice the value of B_{el} as defined by Roark and Yphantis (1971) for a uni-univalent salt.

A



B



C

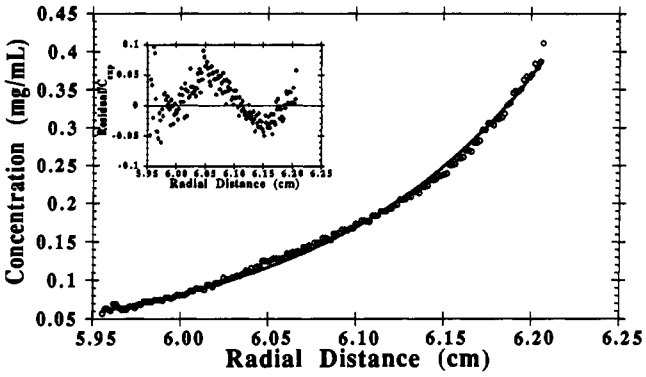


FIGURE 3: Concentration distribution of human relaxin as described in Figure 2.

where M_w is the molecular weight of the macromolecule and m is the molality of the salt (Roark & Yphantis, 1971). An estimate for Z can be made from a computation of the theoretical titration curve of relaxin with the understanding that the effective charge might be less. At pH 5, Z is computed to be 3.8 from the amino acid composition of relaxin and from intrinsic pK values for ionizable amino acid residues (Shire, 1983). Assuming the major contribution to the ionic strength in the isotonic 10 mM citrate buffer is from NaCl at ~0.13 M, B is estimated to be $\sim 1.6 \times 10^{-6}$ mol·L/g^2, which is consistent with the low values determined for B experimentally.

The simplex algorithm yields values of 1.3 and 0.06 (g/L)$^{-1}$ for the monomer–dimer association constant for porcine relaxin from two separate experiments. Figure 2 shows the result of this fit from one of the replicate experiments. The fit to the consensus model at the lowest concentration is not as good as at the higher concentrations. However, the second experiment shows a better fit at the low concentrations and a worse fit at the mid concentration range. It is likely that the variations are experimental error within one particular channel pair rather than the fact that one of the components of the heterogeneous porcine relaxin preparation associates to a different extent. This conclusion is further supported by previous sedimentation equilibrium work using three preparations of porcine relaxin (Sherwood & O'Byrne, 1974). Although these three preparations differed slightly in the length of the B chain they had essentially the same molecular weight, ~6300 ± 100, demonstrating that the different porcine relaxin components do not associate to any appreciable event. The average value from two experiments of the monomer–dimer association constant of 0.7 ± 0.7 (g/L)$^{-1}$ yields a weight average molecular weight of 7400 ± 1300 at a total porcine relaxin concentration of 0.5 mg/mL, which is in good agreement with the reported value of 6300.

Table II: Far-UV Circular Dichroism Parameters and Secondary Structure Estimations for Human Relaxin, Porcine Relaxin, and Bovine Insulin

protein	$\theta_{195}/\theta_{208}$	$\theta_{208}/\theta_{222}$	% helix	% β sheet	% remainder
porcine relaxin	1.54[a]	1.65–1.71[a,b]	29[c,f]	19[c,f]	
porcine relaxin[d]	1.16	1.86	50[e]	50[e]	0[e]
			37[f]	30[f]	32[f]
human relaxin[d]	2.44	1.18	53[e]	47[e]	0[e]
			50[f]	50[f]	0[f]
bovine insulin[g] zinc free	2.05[a]	1.38[a]	30[c,f]	15[c,f]	

[a] Schwabe and Harmon (1978). [b] Du et al. (1982). [c] Rawitch et al. (1980). [d] From present work. [e] Provencher analysis. [f] Yang analysis. [g] Wood et al. (1975).

The corresponding concentration gradient and fit to a nonideal monomer–dimer association model for human relaxin after attaining equilibrium is shown in Figure 3. The parameters from such a fit for two separate experiments are given in Table I and clearly demonstrate that human relaxin self-associates in solution. With use of the mean value for K_2 of 100 (g/L)$^{-1}$, the computed weight average molecular weight at a total human relaxin concentration of 0.5 mg/mL is 11 140. A similar result is obtained at lower rotor speeds, which demonstrates that the self-association is a reversible equilibrium dependent on total protein concentration (Teller, 1972).

The data appear to be fit equally well on the basis of the root mean square lack of fit to discrete monomer–dimer–trimer and monomer–dimer–tetramer models. However, for both higher association models, the higher aggregate association constant is at least 100-fold lower in magnitude than the monomer–dimer association constants whereas the latter are similar. The concentrations of trimers or tetramers would be far less than 0.1% of the total protein concentration, and therefore the human relaxin association is well characterized as a monomer–dimer equilibrium. Moreover, it is unlikely that human relaxin associates in an indefinite association scheme since the fit of the data to an indefinite association scheme is not as good as for the discrete association models. This is mainly reflected in a root mean square lack of fit for the indefinite association model that is almost twice that of the discrete association models (Table I).

Circular Dichroism. The far-UV CD in the spectral range of 250–190 nm for porcine and human relaxin is shown in Figure 4A,B The bold solid lines in the figures are a result of the Provencher fitting algorithm, and estimates of secondary structure from this analysis are included in Table II. Although

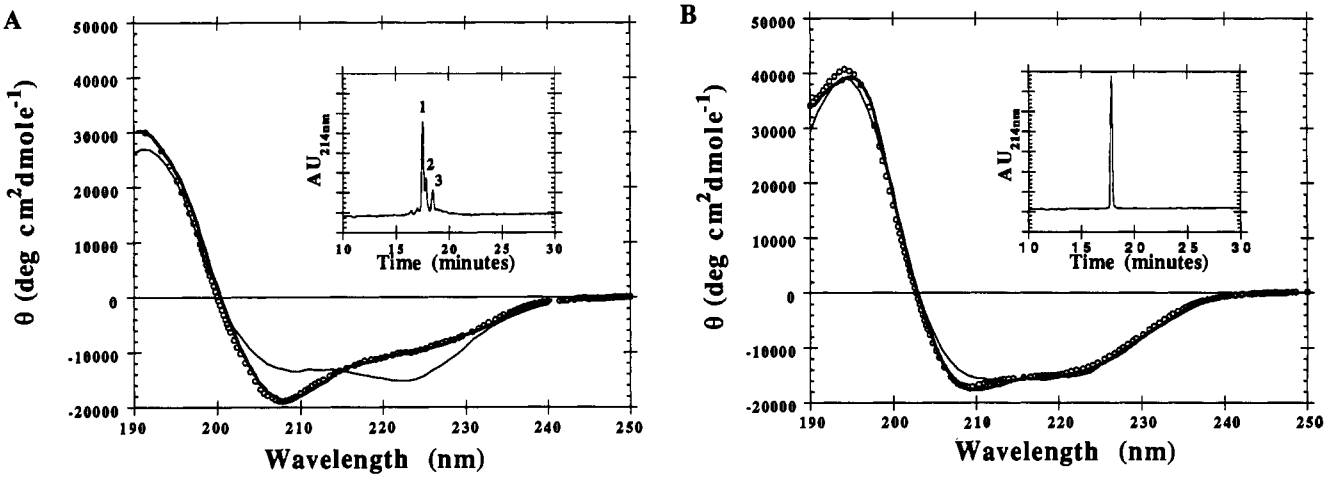


FIGURE 4: Far-UV circular dichroism of porcine (panel A) and human (panel B) relaxin at 0.5 mg/mL in a thermostated 0.01-cm cylindrical cuvette maintained at 20 °C. Open circles are the result of an average of three scans using an average time for each single data point collection of 5 s. Data were collected at 0.2-nm intervals at a spectral bandwidth of 1.5 nm. Some data points are not shown for clarity. The bold solid line is a result of a fit using Provencher's algorithm, and the thin solid line is a result of a fit using Yang's algorithm. The estimated secondary structure from these analyses is presented in Table II. The insets in panels A and B show the reversed-phase chromatograms for porcine and human relaxin, respectively. The reversed-phase chromatography was carried out with a TFA acetonitrile gradient system. The ordinate is an arbitrary scale, and each chromatogram is adjusted for full-scale expansion in order to clearly show all peaks. In panel B the human relaxin main peak is 95% of the total area and all other minor peaks are less than 1%. In panel A peaks labeled 1, 2, and 3 have porcine relaxin with different B-chain lengths as discussed in the text. The percent of total area for peaks 1–3 is 44%, 17%, and 14%, respectively.

both proteins exhibit a CD spectrum that is typical of the α-helix conformation, they differ significantly in regard to the position of the minimum near 209–211 nm and the maximum near 190–200 nm. Qualitatively, the differences between the porcine and human relaxin spectra are similar to the differences between zinc-free insulin and porcine relaxin spectra (Table II) reported by Schwabe and Harmon (1978). The overall blue shift of the CD spectrum for porcine relaxin compared to insulin is also observed when compared to human relaxin. In particular, the positive CD band near 190–200 nm for human relaxin is centered around 194.2 nm whereas in porcine relaxin it is blue-shifted to 191 nm. This shift in this CD band results in a greater value for the ratio of the ellipticities at 208 and 195 nm for human relaxin when compared to that of porcine relaxin. The CD spectra of these proteins also differ significantly in the 230–215-nm range, which leads to a lower value for the ratio of the ellipticities at 208 and 222 nm for human relaxin when compared to that of porcine relaxin. It was suggested that the differences in this wavelength region between zinc-free insulin and porcine relaxin may be attributed to the fact that zinc-free insulin is dimeric, whereas porcine relaxin is monomeric under the conditions studied (Schwabe & Harmon, 1978). At first glance the CD spectra and sedimentation studies of porcine and human relaxin appear to support this idea. However, dilution of the protein to 10 μg/mL with a concomitant increase of almost 5-fold in the weight fraction of monomer does not alter the CD spectrum of human relaxin (Figure 5). Since these differences in the far-UV spectra are not related to the state of association of the relaxin, it is tempting to ascribe these spectral differences instead to differences in secondary structure between the two molecules. The analysis using Provencher's algorithm suggests that there is little difference between the human and porcine relaxin secondary structures, whereas an analysis using the algorithm of Yang suggests that human relaxin has a 50% α-helix content compared to 37% for porcine (Table II). The values we obtain for the latter analysis are consistent with that reported by previous workers for porcine relaxin using the Yang algorithm (Rawitch et al., 1980). However, this result should be interpreted with caution. The fit of the data using the Yang algorithm is shown as the solid lines in Figure 4A,B.

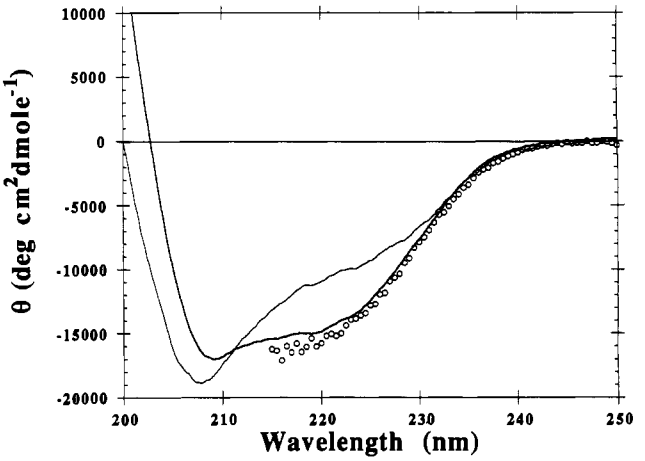


FIGURE 5: Far-UV circular dichroism of human relaxin at 0.5 mg/mL (bold solid line) and 10 μg/mL (open circles). The 0.5 mg/mL data was collected as described in Figure 4. The CD data for relaxin at 10 μg/mL were obtained in a 1-cm thermostated cylindrical cuvette and are the result of an average of 10 scans using an average time for each single data point collection of 10 s. The weight fraction of monomer estimated from the determined association constant of 100 $(g/L)^{-1}$ (Table I) is 0.13 at 0.5 mg/mL and 0.62 at 10 μg/mL. The far-UV CD spectrum for porcine relaxin is also shown as the thin solid line.

Although the fit to the human data is very close to that provided by the Provencher analysis, it is not nearly as good for the analysis of the porcine data.

An additional complication in the interpretation of the far-UV CD spectra for both relaxins is the potential contribution from the aromatic amino acid residues tryptophan and tyrosine, which have strong absorption bands in the far-UV region of 230–200 nm (Wetlaufer, 1962). If these residues contribute in a different manner to the far-UV CD for both relaxins, it is possible that the observed differences in the far-UV region are simply due to different aromatic chromophore environments. An examination of the near-UV CD spectra for porcine and human relaxin (Figure 6) suggests that this might be the case. The near-UV CD spectra of both proteins from 320 to 240 nm have some striking features. Although both proteins have two tryptophans, there is a very

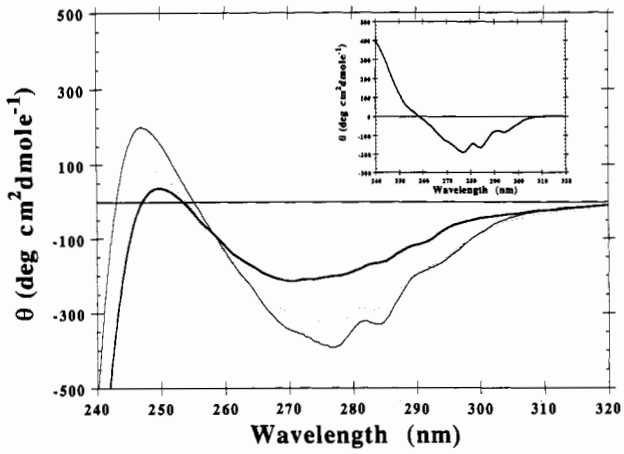


FIGURE 6: Near-UV circular dichroism of porcine relaxin at 0.5 mg/mL (bold solid line) and human relaxin at 0.5 mg/mL (solid line), and 20 μg/mL (dotted line). Relaxin at 0.5 mg/mL was thermostated at 20 °C in a 1-cm cell, whereas relaxin at 20 μg/mL was in an unthermostated 10-cm cylindrical cuvette. The temperature in the sample compartment was ~27 °C during the data collection process. The CD data were collected at 0.25-nm intervals at a spectral bandwidth of 0.5 nm and are the result of an average of three scans using an average time for each single data point collection of 5 s for the 0.5 mg/mL samples and the result of an average of ten scans using an average time for each single data point collection of 10 s for the 20 μg/mL sample. The inset shows the result of subtracting the porcine relaxin CD from the human relaxin CD at 0.5 mg/mL. The weight fraction of human relaxin monomer estimated from the determined association constant of 100 $(g/L)^{-1}$ (Table I) is 0.13 at 0.5 mg/mL and 0.50 at 20 μg/mL.

small CD signal for the ^1L_b band of tryptophan (Strickland et al., 1970). This CD band appears as a broad shoulder centered around 295 nm in the human relaxin spectrum and is practically undetectable in the porcine spectrum. The relatively low amplitude of the tryptophan CD band may be due to the canceling of negative and positive CD signals from the two tryptophans. It has been suggested on statistical grounds that the average CD per aromatic residue will tend to decrease as the number of aromatic residues increases (Strickland et al., 1970). This type of behavior has been demonstrated in aspartate transcarbamylase, which shows extensive loss of CD signal above 270 nm (Griffin et al., 1972). An alternative explanation is that the tryptophan residues in both proteins are very exposed to solvent and quite mobile. This would also result in an averaging out of the chiral environment around the tryptophan, as well as a broadening of the CD bands due to interaction of the chromophore with the polar solvent (Strickland, 1974). The latter explanation is more likely since the cancellation would be expected to be more prominent for higher molecular weight proteins with numerous aromatic amino acid residues such as aspartate transcarbamylase. However, the two proposed models of porcine relaxin structure based on insulin coordinates suggests that one of the tryptophan residues in porcine relaxin is buried and that the other is exposed (Bedarkar et al., 1977; Issacs et al., 1978). The proposed models are further supported by chemical modification and fluorescence experiments of Schwabe and Braddon (1976), who found that one of the tryptophans in porcine relaxin was not readily amenable to chemical modification. Further work will be required to determine if in fact the reduced tryptophan CD band is due to canceling of negative and positive CD bands or to the exposure to solvent and increased flexibility of both chromophores.

The intense ^1L_b CD bands for tyrosine normally occur near 277 and 283 nm, whereas for tryptophan this band usually lies between 288 and 292 nm (Strickland, 1974). The negative dichroic band in the human relaxin spectrum at ~277 nm is from the lone tyrosine residue, whereas the negative band at ~284 nm contains contributions from both tyrosine and the ^1L_b tryptophan CD band. The spacing between the two bands is characteristic of the tyrosine ^1L_b transition (Strickland, 1974), and therefore these bands probably arise mainly from tyrosine. Porcine relaxin shows a broad CD band in the region where tyrosine normally appears, and most of this is probably due to the disulfide CD contribution. The inset of Figure 6 shows the difference CD spectrum for these proteins. If the environments of the tryptophans and the disulfides in both proteins were similar, then the resulting difference spectrum should be representative of the lone tyrosine in porcine relaxin. The difference CD spectrum has bands near 278 and 283 nm, most of which probably arise from tyrosine (Strickland, 1974), as well as a broad band near 293 nm. This broad band may be due to the tryptophan chromophores, which are likely the result of different environments for these residues in the two relaxins. These differences in environment that contribute to differences in the near-UV CD may also manifest themselves as differences in the far-UV CD spectrum of porcine and human relaxin.

It has been demonstrated that decreases in the negative CD band for tyrosine at 275 nm accompany dissociation of the insulin hexamer (Strickland & Mercola, 1976). We have already demonstrated that dilution of human relaxin to 10 μg/mL results in no observed change in the far-UV CD. A similar experiment was carried out in the near-UV with use of a 10-cm path length cylindrical cell. The resulting CD spectrum (Figure 6) after dilution to 20 μg/mL shows a clear decrease in the CD signal at 277 and 284 nm, whereas there is little change in the broad band near 295 nm.

Summary and Conclusions. Porcine and human relaxin exhibit very different self-association properties in solution as determined by analytical ultracentrifuge analysis. The molecular weights determined for porcine relaxin in isotonic 10 mM citrate, pH 5, are in very good agreement with previous work performed on relaxin preparations buffered with 20 mM potassium phosphate, pH 7.4, containing 0.2 M sodium chloride (Sherwood & O'Byrne, 1974). In contrast, human relaxin formulated in the isotonic citrate buffer reversibly self-associates to a dimer. This difference in association behavior between these two relaxins is not strongly pH-dependent (unpublished observations). It is unlikely that the association of human relaxin is due to the particular method of producing this protein since human relaxin made by combining A and B chains that were synthesized by the solid-phase synthesis method of Merrifield (Canova-Davis et al., 1990) also self-associates to dimers (Shire et al., 1988).

It is interesting that dissociation of the dimer does not alter the far-UV CD, whereas the near-UV CD does show a significant change in the region usually assigned to tyrosine residues. This suggests that, upon dissociation to the monomer, the lone tyrosine environment is altered sufficiently to allow for greater mobility of the tyrosine residue and hence a decrease in CD signal strength. In contrast, the fact that the far-UV signal does not change suggests that the dissociation of the human relaxin dimer is accompanied by little overall change in secondary structure. Furthermore, the fact that there are no changes in the broad band near 295 nm suggests that the environment of the tryptophans does not change greatly upon dissociation of the dimer.

The aggregation properties of insulin, with and without zinc, are well characterized (Blundell et al., 1971; Hodgkin & Mercola, 1972) as well as the physiological role for this as-

sembly process. Evaluation of analytical ultracentrifuge data (Goldman & Carpenter, 1974; Jeffrey & Coates, 1966) agrees nicely with the organization found in various crystal forms of insulin (Blundell et al., 1972). It is also believed that, for the most part, the molecular conformations of monomer and dimers in the insulin crystals are preserved in solution (Strickland & Mercola, 1976). We have demonstrated that human relaxin in solution self-associates, and this self-association is adequately modeled as a monomer–dimer equilibrium. It may be significant that at higher concentrations we have been able to crystallize human relaxin (unpublished observations) and preliminary diffraction studies show a dimeric structure (C. Eigenbrott, personal communication). These studies eventually will reveal the molecular details of the interactions responsible for dimer formation. It will be of particular interest to understand what differences in the porcine and human relaxin sequences are responsible for the different behaviors of these related molecules. The interactions that govern insulin association may be very different from those in human relaxin because of the relative lack of sequence homology between human insulin and relaxin (~21%). The near-UV CD experiments at 20 μg/mL (Figure 6) suggest that there is a major alteration of environment of tyrosine in human relaxin upon dissociation to monomer. This tyrosine is replaced by an arginine in porcine relaxin (Figure 1), and this substitution may affect the self-association properties of the protein. Whether there is a physiological requirement for association of human relxin either at the level of a storage mechanism or receptor interaction will await further studies.

Acknowledgments

We thank Dr. Geoffrey Tregear for kindley supplying us with purified porcine relaxin (prep. no. 300). We also gratefully acknowledge the critical reading of the manuscript and helpful comments by Dr. Tom Bewley. We also thank Dr. Stephen Provencher for kindly supplying us with his computer program Contin.

Appendix: Initial Estimates, Movement, and Convergence Criteria for the Simplex Algorithm

The initial estimate for the dimerization equilibrium constant was obtained from the slope (K_{est}/M_1) of plots of $(M_w - M_1)/(2M_1 - M_w)^2$ vs total weight concentration. As explained in the text, the natural logs of the equilibrium constants are the fitted parameters. We chose to allow the program to fit the monomer–dimer constants using initial estimates of ln K_2 of 0.5, which were 10% of the value of ln K_{est}. This initial estimate was used in fitting the sedimentation data for porcine and human relaxin to a monomer–dimer association scheme. In the case of higher association models, we used the same estimates for ln K_2 and a value of 0.2 for the higher order association constant, and in the case of the isodesmic fit we used the same initial value of 0.5 for the association constant k. The initial estimate for all models for the second virial coefficient was 1×10^{-6}. The simplex formed from this set of initial estimates is a geometric figure with the number of vertices equal to the number of sought parameters plus one. The initial step size used to generate the other vertices of the starting simplex was 5% of the initial parameter estimates. The simplex algorithm seeks to reduce the residual sum of squares at each iteration by deleting from the simplex the vertex with the highest value for the residual sum of squares. The four movements used to replace the worst vertex are (1) reflection away from the worst vertex if the new vertex is better than the worst but no better than the best, (2) expansion of that reflection should the reflected vertex be lower on the residual sum of squares surface than all the current vertices, (3) contraction of the worst vertex toward the center of the simplex if the reflected vertex is higher on the response surface, and, if all of these alternatives fail, (4) shrink all the vertices toward the best vertex. An expansion coefficient of 2, a contraction coefficient of 0.5, and a reflection coefficient of 1 were used. In practice, the number of iterations required does not depend greatly on the exact values for these coefficients. The simplex thus moves steadily downhill along the residual sum of squares surface until all attempts at improving the response fail as the simplex reaches the global minimum of the surface. At this point, the vertices contract, and with successive iterations neither the parameter estimates nor the residual sum of squares changes appreciably. Convergence was considered attained when the absolute relative difference among vertices for each parameter was less than 1×10^{-5} and the maximum relative difference for the residual sum of squares was less than 1×10^{-6}.

References

Adams, E. T., & Lewis, M. S. (1968) *Biochemistry 7*, 1044–1053.

Adler, A. J., Greenfield, N. J., & Fasman, G. D. (1973) *Methods Enzymol. 27*, 675–735.

Bedarkar, S., Turnell, W. G., Bundell, T. L., & Schwabe, C. (1977) *Nature 270*, 449–451.

Bewley, T. A. (1982) *Anal. Biochem. 123*, 55–65.

Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Hodgkin, D. C., Mercola, D. A., & Vijayan, M. (1971) *Nature (London) 231*, 506–511.

Blundell, T. L., Dodson, G., Hodgkin, D., & Mercola, D. A. (1972) *Adv. Protein Chem. 26*, 279–394.

Bryant-Greenwood, G. D. (1982) *Endocr. Rev. 3*, 62–90.

Bullesbach, E. E., & Schwabe, C. (1985) *Biochemistry 24*, 7722–7728.

Bullesbach, E. E., & Schwabe, C. (1986) *Biochemistry 25*, 5998–6004.

Caceci, M. S., & Cacheris, W. P. (1984) *Byte 9*, 340–362.

Canova-Davis, E., Baldonado, I. P., & Teshima, G. M. (1990) *J. Chromatogr. 508*, 81–96.

Chang, C. T., Wu, C. C., & Yang, J. T. (1978) *Anal. Biochem. 91*, 13–31.

Cohn, E. J., & Edsall, J. T. (1965) *Proteins, Amino Acids, and Peptides as Ions and Dipolar Ions*, Hafner, New York.

Du, Y-C., Minasian, E., Tregear, G. W., & Leach, S. J. (1982) *Int. J. Pept. Protein Res. 20*, 47–55.

Englander, S. W., & Epstein, H. T. (1957) *Arch. Biochem. Biophys. 68*, 144–149.

Evans, M. I., Dougan, M.-B., Moawad, A. H., Evans, W. J., Bryant-Greenwood, G. D., & Greenwood, F. C. (1983) *Am. J. Obstet. Gynecol. 147*, 410–414.

Fei, D. T. W., Gross, M. C., Lofgren, J. L., Mora-Worms, M., & Chen, A. B. (1990) *Biochem. Biophys. Res. Commun. 170*, 214–222.

Goldman, J., & Carpenter, F. H. (1974) *Biochemistry 13*, 4566–4574.

Griffin, J. H., Rosenbusch, J. P., Weber, K. K., & Blout, E. R. (1972) *J. Biol. Chem. 247*, 6482–6490.

Haschemeyer, R. H., & Bowers, W. F. (1970) *Biochemistry 9*, 435–445.

Hodgkin, D. C., & Mercola, D. A. (1972) in *Handbook of Physiology I* (Steiner, D., Ed.) Vol. I, pp 139–157, American Physiological Society, Washington, DC.

Holladay, L. A., & Sophianopoulos, A. J. (1972) *J. Biol. Chem. 247*, 427–439.

Hudson, P., John, M., Crawford, R., Haralambidis, J., Scanlon, D., Gorman, J., Tregear, G., Shine, J., & Niall, H. (1984) *EMBO J. 3*, 2333–2339.
Issacs, N., James, R., Niall, H., Bryant-Greenwood, G., Dodson, G., Evans, A., & North, A. C. T. (1978) *Nature (London) 271*, 278–281.
Jeffrey, P. D., & Coates, J. H. (1966) *Biochemistry 5*, 3820–3824.
Johnson, M. L., Correia, J. J., Yphantis, D. A., & Halvorson, H. R. (1981) *Biophys. J. 36*, 575–588.
MacLennan, A., Green, R. C., Bryant-Greenwood, G. D., Greenwood, F. C., & Seamark, R. F. (1981) *Obstet. Gynecol. 58*, 601–604.
MacLennan, A. H., Green, R. C., Grant, P., & Nicolson, R. (1986) *Obstet. Gynecol. 68*, 598–601.
Nelder, J. A., & Mead, R. (1965) *Comput. J. 7*, 308–313.
Niall, H. D., M. John, R. James, S. Kwok, R. Mercado, G. Bryant-Greenwood, R. A. Bradshaw, M. G., & Boime, I. (1980) *Insulin: Chemistry, Structure, and Function of Insulin and Related Hormones*, de Gruyter, Berlin.
Provencher, S. W., & Glockner, J. (1981) *Biochemistry 20*, 33–37.
Rawitch, A. B., Moore, W. V., & Frieden, E. H. (1980) *Int. J. Biochem. 11*, 357–362.
Roark, D. E., & Yphantis, D. Y. (1971) *Biochemistry 10*, 3241–3249.
Schwabe, C., & Braddon, E. N. (1976) *Biochem. Biophys. Res. Commun. 68*, 1126–1132.
Schwabe, C., & Harmon, S. J. (1978) *BBRC 84*, 374–380.
Sherwood, C. D., & O'Byrne, E. M. (1974) *Arch. Biochem. Biophys. 160*, 185–196.
Sherwood, O. D. (1988) *The Physiology of Reproduction*, Raven Press, New York.
Shire, S. J. (1983) *Biochemistry 22*, 2664–2671.
Shire, S. J., Foster, D. L., & Rinderknecht, E. (1988) *Biophys. J. 53*, 72a.
Steinetz, B. G., Beach, V. L., & Kroc, R. L. (1959) *Recent Progress in the Endricrinology of Reproduction*, Academic Press, New York.
Strickland, E. H. (1974) *CRC Crit. Rev. Biochem. 2*, 113–175.
Strickland, E. H., & Mercola, D. (1976) *Biochemistry 15*, 3875–3884.
Strickland, E. H., Kay, E., & Shannon, L. M. (1970) *J. Biol. Chem. 245*, 1233–1238.
Stults, J. T., Bourell, J. H., Canova-Davis, E., Ling, V. T., Laramee, G. R., Winslow, J. W., Griffin, P. R., Rinderknecht, E., & Vandlen, R. L. (1990) *Biomed. Environ. Mass Spectrom. 19*, 655–664.
Teller, D. C. (1972) *Methods Enzymol. 27*, 346–441.
Yphantis, D. A. (1964) *Biochemistry 3*, 297–317.
Wetlaufer, D. B. (1962) *Adv. Protein Chem. 17*, 303–390.

^1H and ^{15}N NMR Resonance Assignments and Preliminary Structural Characterization of *Escherichia coli* Apocytochrome b_{562}[†]

Yiqing Feng,[‡] A. Joshua Wand,*[,‡,§,∥] and Stephen G. Sligar*[,‡]
Department of Biochemistry, University of Illinois at Urbana-Champaign, Illinois 61801, and Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111
Received April 18, 1991; Revised Manuscript Received June 12, 1991

ABSTRACT: The ^1H and ^{15}N resonances of uniformly enriched apocytochrome b_{562} (106 residues) have been assigned. The assignment work began with the identification of the majority of H^N–H^α–H^β subspin systems in two-dimensional DQF-COSY and TOCSY spectra of unlabeled protein in D_2O and in 95% H_2O/5% D_2O buffer. Intraresidue and interresidue NOE connectivities were then searched for in two-dimensional homonuclear NOESY spectra recorded on unlabeled protein and in the three-dimensional NOESY-HMQC spectrum recorded on uniformly ^{15}N-enriched protein. Those data, combined with the main-chain-directed assignment strategy (MCD), led to the assignment of the main-chain and many side-chain resonances of 103 of the 106 residues. Qualitatively, the helical conformation is found to be the dominant secondary structure in apocytochrome b_{562} as it is in holocytochrome b_{562}. The helical segments in apocytochrome b_{562} overlap extensively with the helical regions defined in the crystal structure of ferricytochrome b_{562}. In addition, a number of tertiary NOEs have been identified which indicate that the global fold of the apoprotein at least partially resembles the four-helix bundle of the holoprotein. The results presented here, together with the evidence obtained with other methods [Feng and Sligar (1991) *Biochemistry* (submitted)], support the notion that the interior of the protein is fluid and may correspond to a molten globule state.

Cytochrome b_{562} found in *Escherichia coli* is a small soluble protein consisting of 106 amino acid residues (Itagaki & Hager, 1966; Lederer et al., 1981). It is classified as an electron transfer protein, although its biological partners in *E. coli* remain elusive. X-ray crystallographic studies of ferricytochrome b_{562} indicate that the molecule is composed of four nearly antiparallel helices packed to form a bundle (Mathews et al., 1979; Lederer et al., 1981). A type b heme is ligated to the polypeptide chain through methionine 7 on the N-terminal helix and histidine 102 on the C-terminal helix. This relatively simple structure makes cytochrome b_{562} an attractive model for protein-folding studies. An important task in protein-folding studies is to determine the structure of partly

[†]Supported by NIH research Grants GM 31756 (S.G.S.) and GM 35940 (A.J.W.), by NIH Instrument Grant RR-03385, by NIH Grants CA-06927 and RR-05539, by grants from the Pew Memorial Trust and the F. Ripple Foundation, and by an appropriation from the Commonwealth of Pennsylvania awarded to the Institute for Cancer Research.
* To whom correspondence should be addressed.
[‡] University of Illinois at Urbana-Champaign.
[§] Fox Chase Cancer Center.
[∥] Present address: Department of Biochemistry, University of Illinois at Urbana-Champaign, IL 61801.

0006-2960/91/0430-7711$02.50/0