Marquardt, D. W. (1963) *J. Soc. Ind. Appl. Math. 11*, 431-441.
McKinnon, A. E., Szabo, A. G., & Miller, D. R. (1977) *J. Phys. Chem. 81*, 1564-1570.
Munro, I., Pecht, I., & Stryer, L. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 56-60.
Norrestam, R., Stensland, B., & Söderberg, E. (1972) *Acta Crystallogr., Sect. B B28*, 659-666.
O'Connor, D. V., Ware, W. R., & Andre, J. C. (1979) *J. Phys. Chem. 83*, 1333-1343.
O'Kane, D. J., & Lee, J. (1985a) *Biochemistry* (second of five papers in this issue).
O'Kane, D. J., & Lee, J. (1985b) *Biochemistry* (fourth of five papers in this issue).
O'Kane, D. J., Karle, V. A., & Lee, J. (1985) *Biochemistry* (first of five papers in this issue).
Papenhuyzen, J., & Visser, A. J. W. G. (1983) *Biophys. Chem. 17*, 57-65.
Platenkamp, R. J. (1981) Ph.D. Thesis, Chapter 6, University of Leyden, The Netherlands.
Rigler, R., & Ehrenberg, M. (1973) *Q. Rev. Biophys. 6*, 139-199.
Searle, G. F. W., van Hoek, A., & Schaafsma, T. J. (1983) in *Picosecond Chemistry and Biology* (Doust, T. A. M., & West, M. A., Eds.) pp 35-67, Science Reviews Ltd., Northwood, Middlesex, England.
Small, E. D., Koka, P., & Lee, J. (1980) *J. Biol. Chem. 255*, 8804-8810.
Small, E. W., & Isenberg, I. (1977) *Biopolymers 16*, 1907-1928.
Tao, T. (1969) *Biopolymers 8*, 609-632.
Thomas, J. C., Allison, S. A., Appellof, C. J., & Schurr, J. M. (1980) *Biophys. Chem. 12*, 177-188.
Turko, B. T., Nairn, J. A., & Sauer, K. (1983) *Rev. Sci. Instrum. 54*, 118-120.
van Hoek, A., & Visser, A. J. W. G. (1981) *Rev. Sci. Instrum. 52*, 1199-1205.
van Hoek, A., & Visser, A. J. W. G. (1983) *Rev. Sci. Instrum. 54*, 1145-1147.
van Hoek, A., Vervoort, J., & Visser, A. J. W. G. (1983) *J. Biochem. Biophys. Methods 7*, 243-254.
Vervoort, J., O'Kane, D. J., Carreira, L. A., & Lee, J. (1983) *Photochem. Photobiol. 37*, 117-119.
Visser, A. J. W. G., & van Hoek, A. (1979) *J. Biochem. Biophys. Methods 1*, 195-208.
Visser, A. J. W. G., & Lee, J. (1980) *Biochemistry 19*, 4366-4372.
Visser, A. J. W. G., & van Hoek, A. (1981) *Photochem. Photobiol. 33*, 35-40.
Visser, A. J. W. G., & Lee, J. (1982) *Biochemistry 21*, 2218-2226.
Wahl, Ph. (1977) *Chem. Phys. 22*, 245-256.
Wahl, Ph. (1979) *Biophys. Chem. 10*, 91-104.
Wahl, Ph. (1983) in *Time-Resolved Fluorescence Spectroscopy in Biochemistry and Biology* (Cundall, R. B., & Dale, R. E., Eds.) pp 497-521, Plenum Press, London.
Yguerabide, J. (1972) *Methods Enzymol. 26C*, 498-578.
Yguerabide, J., Epstein, H. F., & Stryer, L. (1970) *J. Mol. Biol. 51*, 573-590.

The Histidine-Rich Glycoprotein of Serum Has a Domain Rich in Histidine, Proline, and Glycine That Binds Heme and Metals†

William T. Morgan
Department of Biochemistry, Louisiana State University Medical Center, New Orleans, Louisiana 70112
Received July 13, 1984

ABSTRACT: Histidine-rich glycoprotein (HRG) from rabbit serum was digested with plasmin, reduced, and carboxymethylated, and the fragments produced were resolved by reverse-phase high-performance liquid chromatography. Several peptide fractions were obtained that contain unusually high contents of histidine, proline, and glycine. One His-Pro-Gly-rich peptide (apparent M_r 30 000) was obtained in sufficient yield and purity for further study. This peptide is 29 mol % histidine, 37% proline, and 16% glycine, indicating that most of these three amino acids are located in one region of HRG. The peptide contains 9% by weight carbohydrate and is devoid of tyrosine or tryptophan. The far-ultraviolet circular dichroism spectrum of the peptide has a minimum at 203 nm, indicating that the peptide contains polyproline II helical sections. The peptide represents a binding domain of HRG since it retains much of the ability of intact HRG to bind heme and metals including Zn^{2+}, Ni^{2+}, and Cu^{2+}. As with the parent HRG molecule, interaction of the peptide with heme and metals is dependent on pH and intact histidine residues.

The histidine-rich glycoprotein[1] (HRG) of serum presents an interesting subject for protein structure–function studies. This protein has been isolated and characterized from both human (Heimburger et al., 1972; Morgan, 1978) and rabbit (Morgan, 1981) serum. In both cases, the protein has a single polypeptide chain, contains a fairly high amount of carbohydrate (ca. 15%), and has an unusually high content of proline as well as histidine. As a first step in determining the structure of this protein, an examination of the arrangement of these

† This work was supported in part by a grant from the USPHS, National Institutes of Health (HD 14481).

[1] Abbreviations: HRG, histidine-rich glycoprotein; heme, iron protoporphyrin IX; mesoheme, iron mesoporphyrin IX; $NaDodSO_4$, sodium dodecyl sulfate; Me_2SO, dimethyl sulfoxide; TFA, trifluoroacetic acid; EDTA, ethylenediaminetetraacetic acid; HPLC, high-performance liquid chromatography; RCM, reduced and carboxymethylated; Hepes, *N*-(2-hydroxyethyl)piperazine-*N'*-2-ethanesulfonic acid.

0006-2960/85/0424-1496$01.50/0

two amino acids in the protein was undertaken.

Such information is needed to elucidate the mechanism of action of HRG in vivo, although the physiological function of HRG is uncertain at present. The protein has been shown to interact with heme and metals (Morgan, 1978, 1981; Guthans & Morgan, 1982), with heparin (Heimburger et al., 1972; Koide et al., 1982; Lijnen et al., 1983a; Kindness et al., 1984), with the lysine-binding site of plasmin(ogen) (Lijnen et al., 1980), and with thrombospondin (Leung et al., 1984). Moreover, HRG has been shown to have activity as an inhibitor of autorosette formation (Rylatt et al., 1981; Lijnen et al., 1983b) and to be associated with platelets (Leung et al., 1983).

In this report, the results of our initial efforts on this question are presented. These show that most of the histidine, proline, and glycine that occurs in HRG is found in a single domain of M_r 30 000–40 000. Several His-Pro-Gly-rich peptides are generated by limited proteolysis of HRG with plasmin. One isolated peptide (M_r 30 000) was shown to retain the ability of HRG to bind heme and metals but apparently not heparin.

Materials and Methods

HRG was isolated from rabbit serum (Pel-Freeze, Rogers, AR) as described (Morgan, 1981). The protein was more than 97% pure as judged by polyacrylamide gel electrophoresis in sodium dodecyl sulfate ($NaDodSO_4$) and as isolated was less than 5% degraded by plasmin. Electrophoresis was performed according to Laemmli (1970) except that a 3–20% gradient of acrylamide was used as noted. Gels were stained for protein with Coomassie Blue R-250 and for carbohydrate with a periodic acid–Shiff method (Segrest & Jackson, 1972).

Plasminogen was purified from fresh frozen human plasma (Southeast Louisiana Blood Center) by chromatography on lysine–Sepharose (Pharmacia) and elution with ϵ-aminocaproic acid (Deutsch & Mertz, 1970). The zymogen was activated with urokinase (Calbiochem-Behring) and stored with glycerol at −20 °C (Morris et al., 1981). Plasmin activity was determined with the chromogenic substrate L-valylleucyllysine-*p*-nitroanilide (S-2251, Kabi).

HRG was treated with plasmin (1:50 w/w) for up to 2 h at 25 °C in 0.1 M sodium phosphate, pH 7.1, and reduced and carboxymethylated by adding urea to 8 M, mercaptoethanol to 1.3%, and EDTA to 0.2%. The sample was flushed for 2 h with nitrogen, bromoacetic acid was then added to 0.3 M, and the reaction was allowed to proceed for an additional 2 h under nitrogen. The sample was then dialyzed exhaustively against water and lyophilized. There was no loss in ninhydrin-reactive material and no change in electrophoretic profile after dialysis, demonstrating that no major peptides were lost.

Amino acid and amino sugar analyses were run on a Glenco modular analyzer after hydrolyzing samples in constant-boiling HCl for 24, 48, or 72 h at 110 °C in vacuo. Aminoguanidinopropionic acid was added to hydrolysates as an internal standard. Cysteic acid was measured after hydrolysis in the presence of dimethyl sulfoxide (Spencer & Wold, 1969). Tryptophan was determined fluorometrically (Pajot, 1976) and after base hydrolysis (Hugli & Moore, 1972), and sialic acid was determined after hydrolysis in 0.05 M H_2SO_4 for 1 h at 80 °C (Skoza & Mohos, 1976).

HRG concentrations were estimated by using an absorbance at 280 nm of 0.53 for a 1 mg/mL solution, and concentrations of peptides were estimated by using an absorbance at 220 nm of 11.5 for a 1 mg/mL solution. Both values were determined from amino acid composition data and amino acid analyses to calibrate the absorbance. Antiserum to rabbit HRG was produced in a goat as described (Morgan, 1978) and showed only a single precipitin line of identity with HRG when tested against whole rabbit serum in double immunodiffusion.

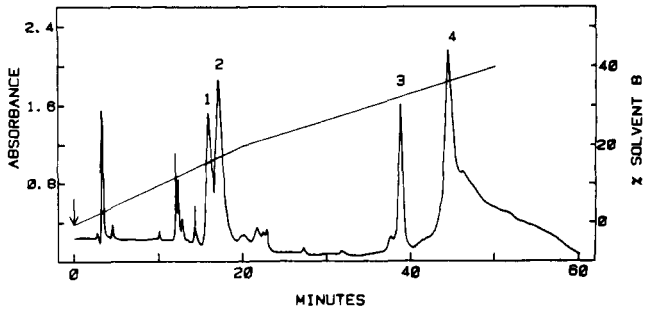

FIGURE 1: HPLC of plasmin-digested HRG. Rabbit HRG was digested with plasmin and reduced and carboxymethylated. RCM-HRG was dissolved in 0.06 M HCl and chromatographed on a reverse-phase C-18 column with 0.1% TFA and the indicated gradient of 1-propanol as solvent B. The chromatogram shown is for 150 μg of material on a 4.1 × 250 mm column at 0.8 mL/min. Peptides marked 1 and 2 are rich in histidine, proline, and glycine, and those marked 3 and 4 are not enriched in these amino acids.

Absorbance and fluorescence spectra and binding titrations were carried out and the data analyzed as described (Morgan, 1978, 1981). Absorbance spectra were obtained with a Cary 210 or Hewlett-Parkard 8450A instrument and fluorescence measurements with a Perkin-Elmer 650-40 spectrofluorometer. Iron mesoporphyrin IX was obtained from Porphyrin Products (Logan, UT). The concentration of mesoheme dissolved in Me_2SO was obtained by using an absorbance of 170 at 403 nm for a 1 mM solution. The effects of pH and diethyl pyrocarbonate on heme binding were determined by using methods applied previously to study heme–hemopexin (Morgan & Muller-Eberhard, 1972; Morgan & Muller-Eberhard, 1976) and metal–HRG interactions (Morgan, 1981). Metal binding was estimated by the ability of added metal to prevent the binding of mesoheme added after the metal (Morgan, 1978, 1981). Absorption spectra were recorded within 5 min of adding the mesoheme and 2 h later. Ethoxyformylation of histidine was monitored by using the characteristic change in absorbance at 242 nm (Miles, 1977).

Peptides were isolated from HRG after digestion with plasmin and carboxymethylation by dissolving the lyophilized, cleaved protein in 0.06 M HCl (up to 4 mg/mL) and separating the material on a Synchrom RP-P (Linden, IN) C-18 reverse-phase column (4.1 × 250 mm or 10 × 300 mm). The column was equilibrated with 0.1% trifluoroacetic acid (TFA) in water and eluted at ambient temperature with a programmed gradient (see Figure 1) on a Varian 5020 chromatograph with neat 1-propanol or 2-propanol as the second solvent (solvent B). The flow was 0.8 mL/min in analytical and 4 mL/min in preparative runs. The effluent was monitored at 220 nm. Fractions were collected, and appropriate pools were made, gently flushed with filtered air at 37 °C to remove alcohol, and lyophilized. The purity of peptides was measured by electrophoresis in $NaDodSO_4$ with Coomassie blue staining and high-performance gel permeation chromatography on a Varian TSK 3000 column (7 × 280 mm) with 0.1% TFA as the solvent and a flow rate of 0.4 mL/min.

Circular dichroism spectra were recorded on a Jasco 500 C spectropolarimeter at 25 °C with 0.1 cm path cells. The instrument was calibrated with *d*-10-camphorsulfonic acid, and all spectra were taken in 5 mM sodium phosphate, pH 7. A 2-nm bandwidth and 4-s time constant were used.

Results and Discussion

Rabbit HRG, like the human congener (Morgan & Smith, 1984), is rapidly cleaved by plasmin at two or three apparently

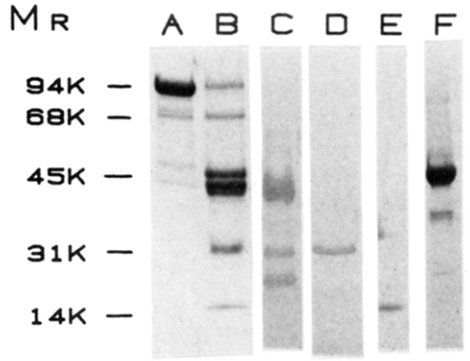

FIGURE 2: $NaDodSO_4$–PAGE of plasmin-digested HRG. Samples (15 μg of protein per lane) were electrophoresed on 3–20% acrylamide gels and stained with Coomassie Blue R-250. Shown are (lane A) intact rabbit HRG, (lane B) digested RCM-HRG, (lane C) peptide 1, (lane D) peptide 2, (lane E) peptide 3, and (lane F) peptide 4. The migration distances of proteins used as molecular weight standards are represented by lines to the left of the figure.

preferred sites as well as at a few secondary sites. After reduction and carboxymethylation the resultant peptides can be resolved by reverse-phase HPLC as shown in Figure 1. Four of the major peptide-containing peaks (labeled 1, 2, 3, and 4) were chosen for further study. Electrophoretic profiles in Figure 2 show that the peptides in these peaks account for the major products of plasmin digestion and have molecular weights ranging from 45 000 to 15 000. The major peptides produced are apparently disulfide linked since no separation by either electrophoresis in $NaDodSO_4$ or HPLC occurs without reduction of these bonds. Peaks 1 and 2 contain peptides that show strikingly high contents of histidine, proline, and glycine (total near 80 mol %) and very low contents of other amino acids (Table I). The minor peaks preceding peak 1 also are enriched in histidine, proline, and glycine but in differing molar ratios. The number of the His-Pro-Gly-rich peptides strongly suggests that they are derived by multiple proteolytic cleavage events from a single region in HRG that is unusually rich in these three amino acids. Peak 3 contains a low molecular weight peptide (M_r 15 000) that is also uncommon in that it is unusually rich in glutamic acid/glutamine, proline, and phenylalanine (Table I). Conversely, the last major peak (i.e., 4) is composed primarily of a peptide (M_r 45 000) that does not have an unusual amino acid composition (Table I).

The peptide from peak 2, termed peptide 2, was chosen for initial further study since it was available in sufficient amount in purified form and since it has one of the higher apparent molecular weights of the His-Pro-Gly-rich peptides (Figure 2; Table II). This peptide, like the other His-Pro-Gly-rich peptides, stained very strongly with silver stain and produced a definite pink or pink-purple color with Coomassie Blue R-250, as found with other proline-rich proteins (Henkin et al., 1978). Peptide 2 has an apparent M_r near 30 000 (Table II) and contains nearly 9% by weight carbohydrate (Table I). The peptide does not contain sialic acid; however, the low sialic acid content of the protein after being reduced and carboxymethylated (2.9%) and exposure of the digest and peptide to acidic conditions during isolation suggest a loss of sialic acid rather than its intrinsic absence.

Peptide 2 is devoid of tryptophan and tyrosine (Table I). This is supported by the spectra shown in Figure 3. There is essentially no absorbance in the 300–250-nm region, and the second derivative spectrum is zero from 320 to 240 nm. The second derivative spectrum of HRG (not shown) at comparable concentration contains features in the 300–250-nm region characteristic of aromatic amino acids (Levine & Federici, 1982) and also differs from peptide 2 in the 220–200-nm region. The far-ultraviolet circular dichroism spectrum of peptide 2 is compared with that of intact HRG in Figure 4. It can be seen that the spectrum of peptide 2 has a minimum near 203 nm and that of HRG near 208 nm. The spectrum of peptide 2 resembles that reported for a proline-rich glycoprotein from human parotid saliva, which has a high content of polyproline II helical structure arising from $(Pro)_n$

Table I: Amino Acid and Carbohydrate Content of Rabbit HRG and Isolated Peptides[a]

	HRG	peptide 1	peptide 2	peptide 3	peptide 4
			Mole Percent		
residue					
Asx	8.1	5.5	3.5	8.7	15.5
Thr	3.9	4.3	0.6	3.2	6.4
Ser	5.7	3.0	1.7	8.8	9.1
Glx	8.4	6.9	2.7	17.2	11.5
Pro	16.6	21.7	37.4	16.2	5.3
Gly	9.4	16.7	16.0	7.8	5.5
Ala	3.8	–	–	5.7	7.2
Cys	2.4	ND	ND	ND	ND
Val	4.4	1.7	–	4.9	6.4
Met	–	ND	ND	ND	ND
Ile	1.7	1.7	1.5	4.5	1.9
Leu	6.2	–	–	9.3	9.3
Tyr	1.9	–	–	2.3	3.3
Phe	5.7	1.1	2.8	12.8	5.5
Lys	4.5	5.5	1.2	5.1	6.0
His	11.2	23.7	29.3	5.4	3.2
Arg	5.6	4.7	1.3	0.6	5.9
Trp	0.5	ND	–	ND	ND
			Weight Percent		
carbohydrate					
total	17.5	ND	8.9	ND	ND
sialic acid	4.8	ND	–	ND	ND

[a] Amino acid and carbohydrate contents were determined for intact HRG and for peaks 1–4 from plasmin-cleaved, reduced–carboxymethylated HRG in Figure 1, denoted as peptides 1, 2, 3, and 4 below. Peptide 1 appears to be a mixture of three His-Pro-Gly-rich peptides. Peptides 2 and 3 are over 93% pure and peptide 4 is about 87% pure, as judged by electrophoretic and HPLC analyses. Results are the average of at least two determinations on two preparations expressed as mole percent (not corrected for carbohydrate) or as weight percent. The carbohydrate content of HRG is based on analysis of individual sugars, while that of the peptide is total hexose as mannose equivalents. (Carboxymethyl)cysteine was detected in peptides 1–4, but insufficient material was available for full analysis. ND, not determined; (–), not detected or trace only.

Table II: Molecular Weights and Properties of Heme Complexes of HRG and Peptide 2[a]

			wavelength maxima (nm)			
			Soret		vis	
	M_r	stoichiometry/K_d	ox.	red.	ox.	red.
HRG	94 000	(28 ± 2)/1 μM	405	413	525, 555 sh	520, 549
RCM-HRG	ND	ND	404	412	524, 554 sh	519, 548
peptide 2	30 000	(25 ± 2)/1 μM	404	410	525, 553 sh	520, 549

[a] Apparent molecular weights were estimated by using $NaDodSO_4$–PAGE profiles as shown in Figure 2. The absorption spectral properties were determined from oxidized (ox.) and reduced (red.) spectra of rabbit mesoheme–HRG and mesoheme–peptide 2 complexes as shown in Figure 5. Stoichiometry (moles of heme bound per mole) and apparent K_d values were determined from binding titrations as shown in Figure 4. RCM-HRG is plasmin-digested, reduced–carboxymethylated HRG before fractionation by reverse-phase HPLC. Fifteen equivalents of mesoheme was added to 90 μg/mL HRG or RCM-HRG and 30 μg/mL peptide 2 in 15 mM phosphate buffer, pH 7.3, and the absorption spectra were recorded before and after reduction with dithionite. Shoulders are denoted by sh; ND, not determined.

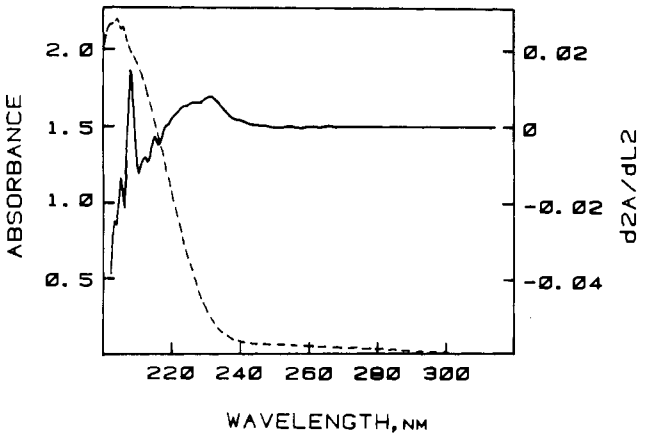

FIGURE 3: Ultraviolet absorption and second derivative spectra of peptide 2. The absorption spectum (---) of peptide 2 (9 μg/mL) in 5 mM phosphate, pH 7.3, was recorded at 22 °C and the second derivative spectrum (—) calculated and plotted by the HP 8450A spectrophotometer.

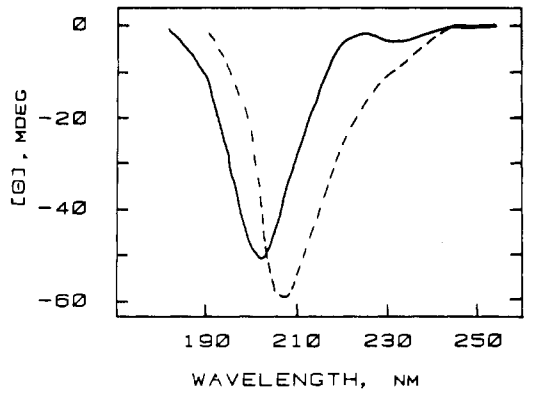

FIGURE 4: Far-ultraviolet circular dichroism spectrum of His-Pro-Gly-rich peptide 2 and of intact HRG. Spectra were recorded at 25 °C in 5 mM sodium phosphate, pH 7.1. Peptide 2 (—) and intact rabbit HRG (---) are shown. The concentration of peptide 2 was 3 μM and that of HRG 1.6 μM. Ellipticity is presented as mdeg·cm^{-1} for each concentration.

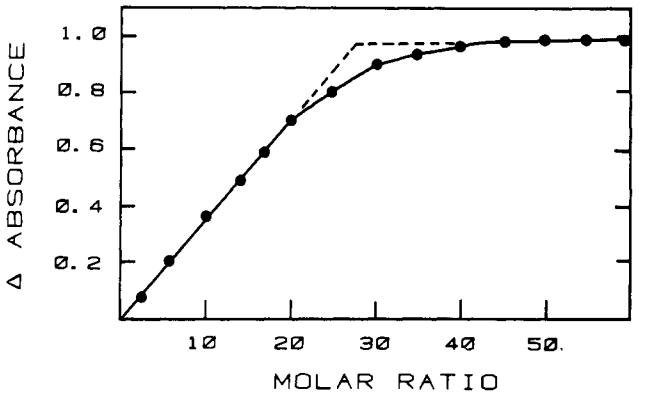

FIGURE 5: Titration of the heme-binding activity of His-Pro-Gly-rich peptide 2. Peptide 2 (30 μg/mL) in 15 mM sodium phosphate, pH 7.3, was mixed with increasing amounts of mesoheme. Similar additions of mesoheme were made to the reference cuvette containing only buffer. Difference spectra were recorded from 450 to 335 nm following each addition, and binding was assessed by change in absorbance at the wavelength of maximum difference, 405 nm.

sequences with n = 2, 3, or 4 (Aubert et al., 1982). Whether similar polyproline structures occur in peptide 2 of HRG requires further study of its sequence and secondary conformation.

Much of the characteristic ability of HRG to bind heme can be accounted for by this peptide. As shown in Figure 5,

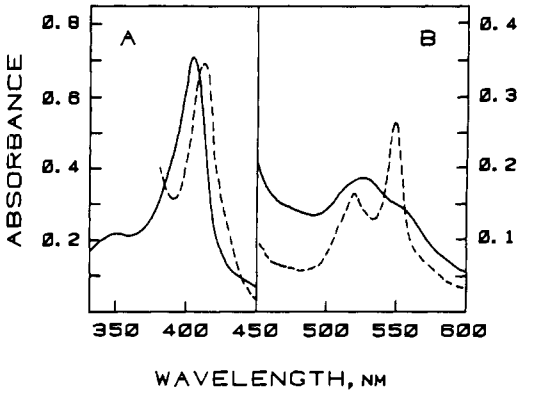

FIGURE 6: Absorption spectra of the mesoheme–peptide 2 complex. Fifteen equivalents of mesoheme was added to peptide 2 (15 μg/mL) in 15 mM phosphate buffer, pH 7.3, and the absorption spectrum was recorded before (—) and after (---) adding $Na_2S_2O_4$ to reduce the heme complex.

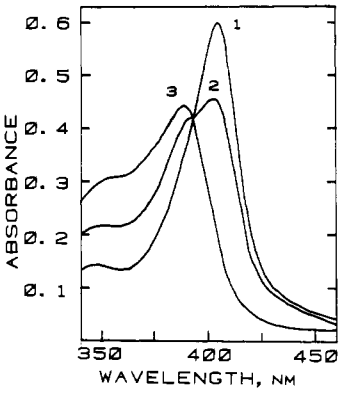

FIGURE 7: Inhibition of heme binding by metals. Binding of metals by peptide 2 was assessed by the ability of metals to inhibit the binding of 15 equiv of mesoheme added to 15 μg/mL peptide 2 in 10 mM Hepes buffer, pH 7.2, 30 min after adding 30 equiv of the metal. Spectra were recorded 5 min after adding heme and 2 h later. No significant change with time was noted. Shown are control, no metal (scan 1), Ni^{2+} (scan 2), and mesoheme in buffer (scan 3). Co^{2+} gave a spectrum like Ni^{2+}; Zn^{2+} and Cu^{2+} inhibited mesoheme binding almost completely. Heparin, Mn^{2+}, and Mg^{2+} were without effect.

binding of mesoheme is saturable, and the titration shows about 25 mol bound/mol of peptide at saturation. The average apparent dissociation constant was estimated to be near 1 μM (Table II). Both parameters are similar to those of intact HRG (Table II). The oxidized and reduced mesoheme–peptide absorption spectra in the Soret and visible regions (Figure 6) resemble intact mesoheme–HRG (Morgan, 1978, 1981) and are consistent with a low-spin, bis(imidazole) type coordination complex. Histidine involvement is also supported by the results of pH dependence and chemical modification studies below. No absorbance band at 620 nm, indicative of high-spin heme complexes (Soininen & Ellfolk, 1973), is found. Carboxymethylated, plasmin-digested HRG also binds heme (Table II) presumably due to the presence of histidine-rich peptides like peptide 2, suggesting either that binding is not strongly conformation dependent or that only local conformation is needed. Two additional sites that bind heme and Rose Bengal are located in another domain of HRG that is not His-Pro-Gly rich (M. K. Burch and W. T. Morgan, unpublished results).

Peptide 2 also interacts with metals. As shown in Figure 7, prior binding of metal prevents the binding of mesoheme added subsequently. As found with intact HRG (Morgan, 1978, 1981; Guthans & Morgan, 1982), divalent nickel, copper, zinc, and cobalt are more tightly bound than cadmium

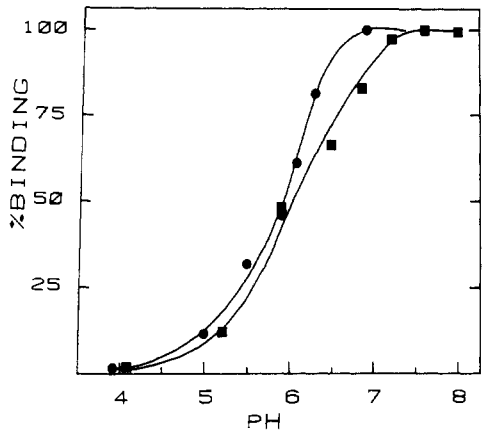

FIGURE 8: pH dependence of the heme–peptide 2 interaction. Mesoheme (20 equiv) was added to 1-mL aliquots of peptide 2 (30 μg/mL) (■) and HRG (90 μg/mL) (●) in 5 mM sodium phosphate, pH 8. The pH was lowered with HCl to the indicated values, and the absorption spectrum of each from 450 to 325 nm was recorded. Volume changes were less than 5%.

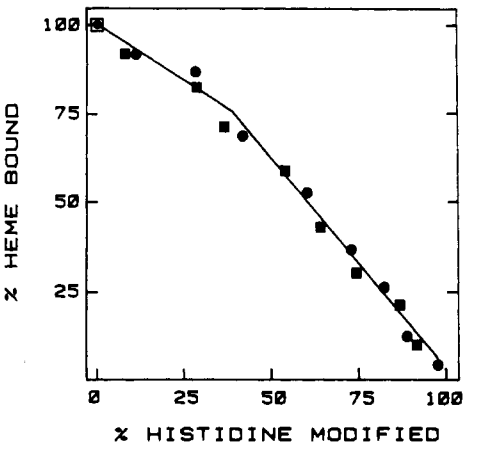

FIGURE 9: Effect of ethoxyformylation of histidine residues of peptide 2 on its ability to bind heme. Aliquots of peptide 2 (30 μg/mL) (■) and of HRG (90 μg/mL) (●) in 15 mM phosphate buffer, pH 7, were treated with increasing amounts of diethyl pyrocarbonate for 30 min on ice. The difference in absorbance at 242 nm between a treated sample and an untreated control was measured to determine the number of histidine residues modified. Then 20 equiv of mesoheme was added, and the absorption spectrum between 450 and 335 nm was recorded to assess residual heme binding.

(not shown). Calcium (3 mM), magnesium (1 mM), manganese (200 μM), and heparin (25 μg/mL) do not inhibit heme binding of either HRG or peptide 2. This suggests either that the heparin-binding site(s) of HRG resides (reside) on another domain of HRG since the apparent affinity of HRG for heparin (Lijnen et al., 1983) is higher than that for heme (Table I) or that the interaction of HRG with heparin requires additional determinants not present in peptide 2. The latter is supported by the enhanced binding of heparin by HRG in the presence of metals (Lijnen et al., 1983).

The interaction of peptide 2 with heme, like that of HRG with heme, is dependent on pH and intact histidine residues. As shown in Figure 8, heme binding decreases at pH values below 7, showing an apparent pK_a of 5.8 for both binding species. The peptide is more sensitive to pH than the protein; at pH 6.5 the peptide has lost about 20% of its heme-binding ability whereas the protein is relatively unaffected. This difference may reflect differences in pK_a values of histidine residues in different environments in the two species, consistent with the presence of two heme-binding sites in another domain of HRG (M. K. Burch and W. T. Morgan, unpublished results). Heme binding also decreases in proportion to the extent of modification of histidine residues with diethyl pyrocarbonate (Figure 9). Together with the heme absorption spectra, these results strongly indicate that histidine residues in the peptide, as in intact HRG, coordinate with heme iron and with the other metals bound. The decreases in binding with increasing histidine modification obtained with intact HRG and peptide 2 are essentially superimposable. There is a decrease in the slope of the line after about 30% of the histidine residues are modified, suggesting that some of the more accessible (and so more readily attacked) residues are not involved in heme binding in either HRG or the peptide. Since binding is abolished only after modification of nearly all histidine residues and since not all residues can be involved in binding heme, the modification of histidine residues is random and not a function of whether a residue binds heme.

In conclusion, HRG is readily cleaved by plasmin and the resultant peptides can be resolved by reverse-phase HPLC. Most of the binding sites of HRG for heme and metal are contained in a limited portion of the molecule. This domain is unusually rich in histidine, proline, and glycine and accounts for the majority of the histidine and proline in HRG. The histidine residues are active in heme binding, and the proline residues may occur in $(Pro)_n$, $n = 2$, 3, or 4, sequences that give rise to a polyproline II type secondary conformation. These results together with the observations that the N-terminal sequence of HRG is related to antithrombin III (Koide et al., 1982) and that two hydrophobic heme-binding and Rose Bengal binding sites exist in a separate region of HRG (M. K. Burch and W. T. Morgan, unpublished results) show that HRG is composed of multiple domains with distinct properties. Understanding how these domains interact and their functions will require further elucidation of the structure and properties of HRG.

ACKNOWLEDGMENTS

The expert technical assistance of J. Chang is gratefully acknowledged. Dr. Alfred Esser, University of Florida Health Center, Gainesville, FL, provided circular dichroism spectra, and Dr. Robert B. Trimble of the New York State Department of Health, Albany, NY, supplied carbohydrate analyses. W. Santa Marina and M. Williams aided in preparation of the manuscript.

Registry No. Mesoheme, 18040-04-5; Ni, 7440-02-0; Cu, 7440-50-8; Zn, 7440-66-6; Co, 7440-48-4.

REFERENCES

Aubert, J. P., Porchet, N., Boersma, A., Loucheux-Lefevore, M. H., & Degand, P. (1982) *Biochim. Biophys. Acta 704*, 361–365.

Cuatrecasas, P., Fuchs, S., & Anfinsen, C. B. (1967) *J. Biol. Chem. 242*, 4759–4767.

Deutsch, D. G., & Mertz, E. T. (1970) *Science (Washington, D.C.) 170*, 1095–1096.

Guthans, S. L., & Morgan, W. T. (1982) *Arch. Biochem. Biophys. 218*, 320–328.

Heimburger, N., Haupt, H., Kranz, T., & Baudner, S. (1972) *Hoppe-Seyler's Z. Physiol. Chem. 353*, 1133–1140.

Henkin, R. I., Lippoldt, R. E., Bilstad, J., Wolf, R. O., Lum, C. K. L., & Edelhoch, H. (1978) *J. Biol. Chem. 253*, 7556–7565.

Hugli, T. E., & Moore, S. (1972) *J. Biol. Chem. 247*, 2828–2834.

Kindness, G., Long, W. F., Vigrow, E. A., Williamson, F. B., Morgan, W. T., & Smith, A. (1984) *Thromb. Res. 33*, 139–143.

Koide, T., Odani, S., & Ono, T. (1982) *FEBS Lett. 141*, 222–224.
Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.
Leung, L. L. K., Harpel, P. C., Nachman, R. L., & Rabellino, E. M. (1983) *Blood 62*, 1016–1021.
Leung, L. L. K., Nachman, R. L., & Harpel, P. C. (1984) *J. Clin. Invest. 73*, 5–12.
Levine, R. L., & Federici, M. M. (1982) *Biochemistry 21*, 2600–2606.
Lijnen, H. R., Hoylaerts, M., & Collen, D. (1980) *J. Biol. Chem. 255*, 10214–10222.
Lijnen, H. R., Hoylaerts, M., & Collen, D. (1983a) *J. Biol. Chem. 258*, 3803–3808.
Lijnen, H. R., Rylatt, D. B., & Collen, D. (1983b) *Biochim. Biophys. Acta 742*, 109–115.
Miles, E. W. (1977) *Methods Enzymol. 47*, 431–442.
Morgan, W. T. (1978) *Biochim. Biophys. Acta* 535, 319–333.
Morgan, W. T. (1981) *Biochemistry 20*, 1054–1061.
Morgan, W. T., & Muller-Eberhard, U. (1972) *J. Biol. Chem. 247*, 7181–7187.
Morgan, W. T., & Muller-Eberhard, U. (1976) *Arch. Biochem. Biophys. 176*, 431–441.
Morgan, W. T., & Smith, A. (1984) *Protides Biol. Fluids 31*, 23–26.
Morris, J. P., Blatt, S., Powell, J. R., Strickland, D. K., & Castellino, F. S. (1981) *Biochemistry 20*, 4811–4816.
Pajot, P. (1976) *Eur. J. Biochem. 63*, 263–269.
Rylatt, D. B., Sia, D. Y., Mundy, J. P., & Parish, C. R. (1981) *Eur. J. Biochem. 119*, 641–646.
Segrest, J. P., & Jackson, R. L. (1972) *Methods Enzymol. 28*, 54–63.
Skoza, L., & Mohos, S. (1976) *Biochem. J. 159*, 457–462.
Soininen, R., & Ellfolk, N. (1973) *Acta Chem. Scand. 27*, 35–46.
Spencer, R. L., & Wold, F. (1969) *Anal. Biochem. 32*, 185–190.

Theory for the Folding and Stability of Globular Proteins†

Ken A. Dill
Departments of Pharmaceutical Chemistry and Pharmacy, University of California, San Francisco, California 94143
Received May 3, 1984

ABSTRACT: Using lattice statistical mechanics, we develop theory to account for the folding of a heteropolymer molecule such as a protein to the globular and soluble state. Folding is assumed to be driven by the association of solvophobic monomers to avoid solvent and opposed by the chain configurational entropy. Theory predicts a phase transition as a function of temperature or solvent character. Molecules that are too short or too long or that have too few solvophobic residues are predicted not to fold. Globular molecules should have a largely solvophobic core, but there is an entropic tendency for some residues to be "out of place", particularly in small molecules. For long chains, molecules comprised of globular domains are predicted to be thermodynamically more stable than spherical molecules. The number of accessible conformations in the globular state is calculated to be an exceedingly small fraction of the number available to the random coil. Previous estimates of this number, which have motivated kinetic theories of folding, err by many tens of orders of magnitude.

Few heteropolymers are both globular and soluble. Proteins are the principal exception. Globularity and solubility cannot be achieved with any random sequence of monomers; certain principles of structure and function must be obeyed (Kauzmann, 1959; Fisher, 1964; Flory, 1969; Tanford, 1968; Brandts, 1968; Edsall, 1968; Edsall & McKenzie, 1983; Lifschitz, 1968; Volkenstein, 1970; Richards, 1977; Klapper, 1971, 1973). A molecule with too many solvophilic residues will prefer solvation to globularity. Molecules with too many solvophobic residues will aggregate, as occurs with oil in water. In addition, globularity requires that the chain can pack well in the condensed state. Typical globular proteins have densities approaching those of crystalline hydrocarbons and amino acids and compressibilities a factor of 20 smaller than liquid hydrocarbons, nearly equal to those of some metals, and they contain less than 3 vol % of internal water or cavities (Richards, 1974, 1977; Klapper, 1971, 1973; Chothia, 1975; Kuntz & Kauzmann, 1974; Connolly, 1981; Gavish et al., 1983; Sturtevant, 1977; Nemethy et al., 1981). The importance of packing also follows from the fact that evolution conserves residue size and shape (Schultz & Schirmer, 1979). But high density comes at a high price; enormous configurational entropy must be overcome to achieve it. The fact that most enzymes are condensed suggests that catalytic function may require the high density state. It is a reasonable hypothesis that this is due to the requirement that the atoms of the active site have relatively invariant spatial positions during a significant fraction of the time required to attract and hold the substrate for the catalytic act. In this regard, the primary molecular mechanism for maintaining relative spatial invariance, i.e., for reducing the amplitude of out-of-phase internal thermal motion, is that of steric constraint, which is achieved, as in the solid state, through high density packing. In the solid state, incident thermal energy may be distributed in modes of motion whose spatial wavelengths are larger than the size of the active site, and thus, this thermal energy may be exchanged with the protein through relatively nondisruptive rigid body motions of the active site.

For those sequences which satisfy the requirements above, it follows that relatively few spatial conformations are available in the globular state, their number being limited by (i) the

† Acknowledgment is made to the donors of the Petroleum Research Fund, administered by the American Chemical Society, for partial support of this work and to the National Institute of General Medical Sciences.

0006-2960/85/0424-1501$01.50/0