Circular Dichroism Studies of Myoglobin and Leghemoglobin†

N. A. Nicola,‡ E. Minasian, C. A. Appleby, and S. J. Leach*

ABSTRACT: The circular dichroism spectra of leghemoglobin a from the root nodules of soybean have been compared with those for sperm whale myoglobin in the far- and near-ultraviolet and the Soret and visible regions of the spectrum. Circular dichroism spectra in the far-ultraviolet show that the leghemoglobins all have a high α-helix content (soybean leghemoglobin a, 55%) regardless of the nature of bound ligands and oxidation or spin state of the heme iron. The known sequence homologies with mammalian hemoglobins may therefore be reflected in conformational homologies as suggested by the x-ray studies of Vainshtein et al. ((1975) *Nature (London) 254,* 163–164) on lupin leghemoglobin. Removal of the heme moiety decreases helicity by only 9% for leghemoglobins, compared with 23% for myoglobin. This, the much smaller heme contribution to the near-ultraviolet circular dichroism than in myoglobin, and the greater accessibility of the heme moiety to aqueous solvent (Nicola et al. (1974), *Proc. Aust. Biochem. Soc. 7,* 21) suggest that the association between heme and protein is much weaker in leghemoglobins than in myoglobin. The aromatic Soret and visible circular dichroism spectra for all derivatives of leghemoglobin are opposite in sense to those for myoglobin, showing that the patterns of protein side chain contacts with the heme are different in the two classes of heme proteins. There is strong evidence that one of the two tryptophans whose identity and structural role in myoglobin is known, is present also in plant leghemoglobins, hydrogen-bonded and in a similar nonpolar environment whether heme is present or not. The above findings help to explain the remarkably high oxygen affinity and some other ligand-binding properties of leghemoglobins which differ from those of myoglobin.

In their structure, leghemoglobins (Lb's)[1] resemble myoglobin (Mb) rather than hemoglobin (Hb). They consist of an apoprotein of molecular weight 15000–20000 and one heme group per molecule (ferroprotoporphyrin IX, Ellfolk, 1960). There is no evidence for the formation of multisubunit structures. The pH dependence of the spectral properties of ferric Lb and the fact that the heme group can be removed by acid–acetone suggest that the heme is linked to the protein through a histidyl residue as it is in Mb and Hb. The sixth coordination position of the iron atom is free to bind ligands, the resulting complexes having similar spectral and magnetic properties to those formed with Mb and Hb.

The Lb's extracted from soybean root nodules can be fractionated into at least five components of similar molecular weights but different charge properties (Appleby et al., 1975). These differ in amino acid composition and so represent different gene products. They have in common low contents of histidine and no sulfur-containing amino acids. The Lb subfractions are named in order of elution from a DEAE-cellulose column as a, b, c_1, c_2, d. To date, those hemoproteins from soybean and especially soybean Lb a have been subjected to the most intensive study. The complete amino acid sequence of Lb's from soybean (Ellfolk and Sievers, 1971), kidney bean (Lehtovaara and Ellfolk, 1974), and broadbean (Richardson et al., 1975) has been elucidated and the amino acid compositions of several other species of Lb are known.

Vainshtein et al. (1975) have carried out a 5-Å resolution structure determination of lupin Lb by x-ray diffraction. Structural studies of Lb's are of interest from a number of standpoints. Firstly, because the structure and function of Mb's and Hb's are so thoroughly understood due to the work of Perutz (1970), it may be possible to relate certain features of Lb structures to their characteristic properties—such as their very high oxygen affinity, their ease of autoxidation, their magnetic properties, and their binding of special ligands. These may prove of interest in showing how variations in structure can modify function in the evolutionary process. Secondly, it is not certain where Lb's fit in the evolutionary process since they are the only known plant Hb's, yet they appear to bear closer resemblances to some higher mammalian than to lower vertebrate Hb's. Comparison of sequence and conformation should help both in finding relationships between Lb's and other globins and in investigating how amino acid sequence determines the mode of folding of globin chains. Thirdly, comparison of structural parameters of different Lb's within one plant species and between different species may help in giving support to one or other of the postulated functions of Lb in the root nodule.

The fractionation of five Lb's from soybean is described elsewhere, together with an account of their ligand-binding, spectroscopic (absorption and circular dichroic), and some tryptic peptide map relationships (Appleby et al., 1975). A study of the heme environments in Mb and Lb's by thermal perturbation spectroscopy has also been carried out (Nicola et al., 1974; N. A. Nicola et al., in preparation). The present paper presents the results of a comparative survey of the CD spectra of sperm whale Mb and soybean Lb in the far-uv (190–250 nm), near-uv (250–300 nm), Soret (300–450 nm), and visible regions (450–650 nm). The ef-

† From the Biochemistry Department, University of Melbourne, Parkville, Victoria, 3052, Australia (N.A.N., E.M., and S.J.L.), and the Division of Plant Industry, C.S.I.R.O., Canberra, A.C.T. 2601 (C.A.A.). *Received May 6, 1975.*

* Recipient of support by the Australian Research Grants Committee. Author to whom correspondence should be addressed.

‡ Supported by a C.S.I.R.O. Studentship.

[1] Abbreviations used are Hb, hemoglobin; Mb, ferric myoglobin; Lb, ferric leghemoglobin; EDTA, ethylenediaminetetraacetate.

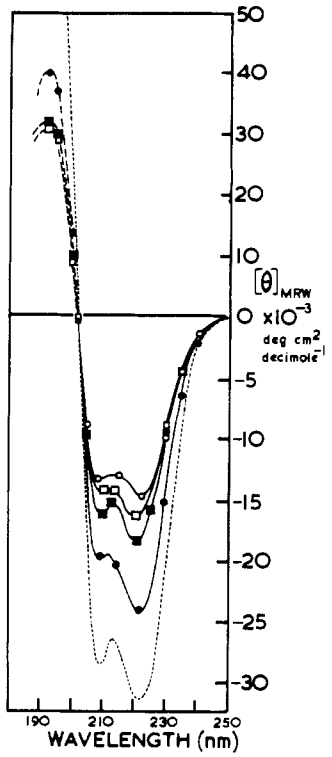

FIGURE 1: Far-uv CD spectra of sperm whale myoglobin and soybean leghemoglobin a in potassium phosphate buffer (0.02 M, pH 7.05). Concentrations 0.2–0.4%. (●) Ferrimyoglobin; (○) apomyoglobin; (■) ferrileghemoglobin a; (□) apoleghemoglobin a; (- - -) 100% α helix (Chen et al., 1971).

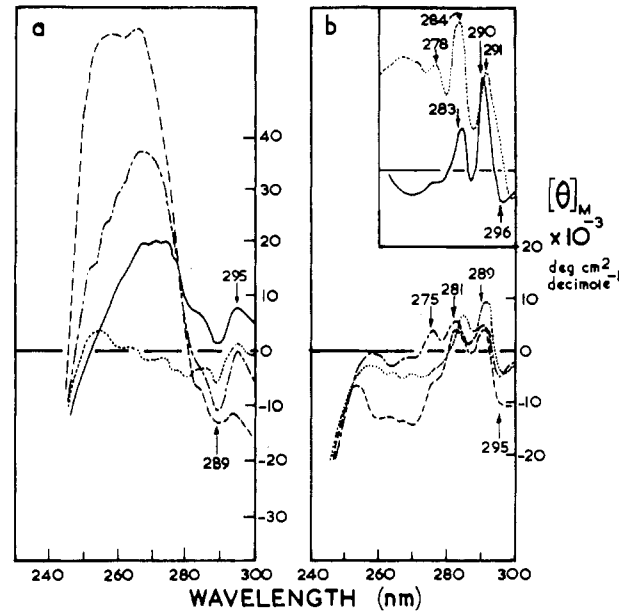

FIGURE 2: Near-uv CD spectra of (a) myoglobin (—) aquoferric complex, (···) apoprotein; (- - -) cyanoferric complex; (- · - ·) fluoroferric complex, and (b) soybean leghemoglobin a (for ferric see curve 2 of Figure 3a) (···) apo; (- - -) nicotinoferric; (- · - ·) fluoroferric. Concentrations 0.2–0.8% in phosphate buffer (0.02 M, pH 7.05). Inset: CD spectra of N-acetyl-L-tryptophanamide at 77 K (arbitrary ellipticity scale); (—) in MeOH; (···) in tetrahydrofuran diglyme (from Strickland et al., 1969).

fect of the spin and liganded states of the heme group on the absorption and CD spectra of these proteins has been given special attention. In the ferric form the high-spin fluoro complex and the low-spin cyano and nicotino complexes have been investigated. In the ferrous form, the high-spin ferrous and the low-spin oxyferrous forms were examined. The interest in the nicotinoferric form stems from the finding that nicotinate occurs in the plant nodules and that it may act as an effector in reducing the high O_2 affinity of Lb (Appleby et al., 1973).

Materials and Methods

A.R. reagents and glass-distilled water were used throughout. Acetone (A.R.) was redistilled twice from $KMnO_4$.

Soybean Lb's were extracted, oxidized to the ferric Lb form with ferricyanide and, after removing reagents by Sephadex (G-15) chromatography, fractionated to give the five component Lb's a, b, c_1, c_2, and d by chromatography on Whatman DE-52 cellulose using an acetate buffer gradient at pH 5.2 as described elsewhere (Appleby et al., 1975). The components were essentially pure by gel electrophoresis.

Concentrations of heme proteins were estimated by a variant of the pyridine hemochrome method of de Duve (1948) in which 1 ml of a solution containing equal volumes of pyridine, water, and 1 M NaOH were added to 5–200 μl of the heme protein solution, dithionite was then added to form the unstable reduced pyridine hemochromogen and the spectrum immediately measured. For complete reaction it was assumed that $A_{556}/A_{538} > 3.5$ with a peak at exactly 556 nm. A value of ϵ_M^{556} 3.46 × 10^4 was used (Paul et al., 1953). Heme estimations on sperm whale Mb from Miles Seravac gave 89.5% purity and using this value in conjunction with the absorption spectrum of the untreated Mb provided values of $\epsilon_M^{408} = 16.5 \times 10^4$ and $\epsilon_M^{280} = 3.19 \times 10^4$, in good agreement with literature values for ferric Mb. Ferrous heme proteins were made by adding the minimum amount of dithionite to the ferric form under N_2. The oxyferrous form was produced by passing ferric hemoprotein through a column of Sephadex G-10 in phosphate buffer (0.02 M, pH 7) after a layer of sodium dithionite (1–10 mg/ml) had started to pass through the column. As the hemoprotein passed through the dithionite band it was converted to the ferrous form but thereafter took up O_2. In this way it was possible to convert 97% of the Lb to Lb O_2 with minimal contact with dithionite. The cyano complex was converted to the ferric form by a similar procedure except that after passing through the dithionite band, the hemoprotein passed through a band of potassium ferricyanide (10 molar excess). The cyano, nicotino, and fluoro complexes were made from the ferric protein by titration with the appropriate ligands. The percentage of oxyferrous and ferric forms was assayed by comparing the absorbance of the sample saturated with CO with that of the pure ferric form (sample + potassium ferricyanide (about 10 equiv)) and the pure oxyferrous form (sample + dithionite + CO). That is, CO was used to replace the O_2.

Apoproteins of Mb and Lb's were made from the heme proteins by the acetone–HCl method of Rossi-Fanelli et al. (1958). The products usually had about 1% of the original heme content and were recovered in about 70% yield. Their concentrations were determined by the Lowry method using bovine serum albumin as the standard or by differential refractometry using a specific refractive increment of 0.188 ml/g. With this value, the molar absorbancies were determined as $\epsilon_M^{278} = 1.62 \times 10^4$ for apo-Mb which agrees fairly well with the literature value of 1.58 × 10^4 (Stryer,

Table I: Mean Residue Ellipticities $[\theta]_{220-222}$ for Derivatives of Myoglobin and Leghemoglobin.[a]

Species of Heme Protein	Ferric				Ferrous		
	Met	Cyano	Fluoro	Nicotino	Oxy	Deoxy	Apo
Sperm-whale Mb	−24100	−21200	−24800	−22300		−24500	−14900
Soybean Lb a pH 7.05	−18800	−17900	−17800	−18600	−19300	−19200	−16300
pH 5.4	−17400	−18300		−17400			

[a] Samples were in potassium phosphate buffer (0.02 *M*, pH 7.05) except where otherwise stated.

1965) and 1.79×10^4 for apo-Lb a of soybean.

Absorption spectra were measured on Millipored solutions at 20° with a Cary 14 spectrophotometer and CD spectra with a Jasco circular dichrometer which had been converted from a J5 to a J20 configuration (maximum sensitivity 2×10^{-5} ΔA unit/cm). The dichrometer was normally used at a sensitivity of 1×10^{-4} ΔA unit/cm to maximize the signal-to-noise ratio. In the far-uv, protein concentrations of 0.2–0.4% were used with cylindrical cells of path length 0.1 mm. Concentrations for all other wavelengths were 0.2–0.8% with either 0.5- and 2-mm cells (near-uv and visible) or 0.2-mm cells (Soret). Path lengths and CD response were calibrated with *d*-10-camphorsulfonic acid. Absorbancies of samples were kept below 2 and within the range where the relation between concentration and ellipticity was linear. Scanning speeds were 50 nm/hr or less. Mean residue ellipticities $[\theta]_{MRW}$ were calculated from raw CD data in the far-uv from $3300(A_L - A_R)MRW/lc$ where $(A_L - A_R)$ is the measured CD, MRW is the mean residue weight (taken as 116 for all of the heme proteins), l is the path length in cm, and c is their concentration in g/l. In quoting α-helix contents the procedure of Chen et al. (1972) was used in which $f_H = -([\theta]_{222} + 2340)/30300$. CD data in the near-uv, Soret, and visible regions are expressed in molar ellipticities $[\theta]_M$, since the CD in these regions is not averaged over all residues but originates in more specific regions of the molecule. None of the ellipticity values quoted are "reduced" by refractive index corrections.

Results

Far-Uv CD Spectra. The CD spectra for soybean Lb a in the holo ferric and the apo form are compared with those for sperm whale Mb in Figure 1. All are characteristic for proteins of high α-helical content. Woody and Tinoco (1967) have shown that values of $[\theta]_{222}$ are less sensitive to helix geometry and length than are $[\theta]_{209}$ and using the former with the equation of Chen et al. (1972) for calculating α-helix content, values of 55, 46, 72.5, and 42% are obtained for ferric Lb a, apo-Lb a, ferric Mb, and apo-Mb, respectively. The apo-Mb result agrees well with that reported by Breslow et al. (1965). Curve fitting the experimental curves with those of Chen et al. (1972) gave very good fits over the entire spectral range when it was assumed that the proteins consisted only of α helix and unordered structures with no β structure. There were minor discrepancies at 208 and 192 nm where the two experimental extrema were slightly smaller and the main Lb a trough was located at 220.5 rather than 222 nm. Such behavior has been attributed to short or distorted helical segments compared with segments in a long and perfect helical array (Schellman and Lowe, 1968; Chen et al. (1974)).

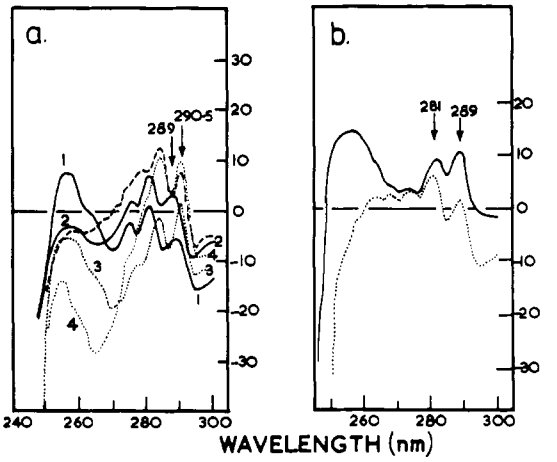

FIGURE 3: Near-uv CD spectra of soybean leghemoglobin a. Concentrations 0.3–0.4%. (a) (2) Ferric; (1) cyanoferric in phosphate buffer (0.01 *M*, pH 7.03); (- - -) ferric; (3) cyanoferric; (4) nicotinoferric, in phosphate buffer (0.02 *M*, pH 5.4). (b) (—) Oxyferrous; (· · ·) deoxyferrous.

The far-uv CD data for various complexes of the hemoproteins are summarized in Table I. Lb's from the nodules of snakebean, lupin and serradella give very similar values to those of soybean (J. G. R. Hurrell et al, in preparation). The various complexes of Lb show small variations in $[\theta]_{220}$ to $[\theta]_{222}$ and the shapes of the CD spectra are very similar. Only when the heme moiety is removed is there a major change in helix content and the effect is much greater on Mb than on Lb.

Near-Uv CD Spectra. The near-uv CD spectra for sperm whale Mb and soybean Lb a derivatives are compared in Figure 2. There are distinct differences between the two proteins, with the Mb derivatives displaying a large positive CD envelope and the Lb derivatives a smaller negative CD envelope, centered around 260 nm. Both disappear on removing the heme. The aromatic fine structure bands are superimposed on these broad envelopes but appear as extrema of opposite sense in the two sets of heme proteins.

The effect of decreasing the pH from 7.03 to 5.4 (under which condition, nicotinate is bound most strongly) on the Lb spectrum is to red-shift the main aromatic bands by 1.5–2.0 nm and this is true regardless of whether the liganded state is ferric, cyanoferric, or nicotinoferric (Figure 3a). In the oxyferrous derivatives, however, the heme bands may be weakly positive (Figure 3b).

Soret-Region CD and Absorption Spectra. Figure 4a shows the CD spectra for the Soret band in some derivatives of Mb. The major Soret CD band is positive in all cases and varies in position from 408 (ferric and fluoroferric) to 437 nm (ferrous). There are also several minor bands. In con-

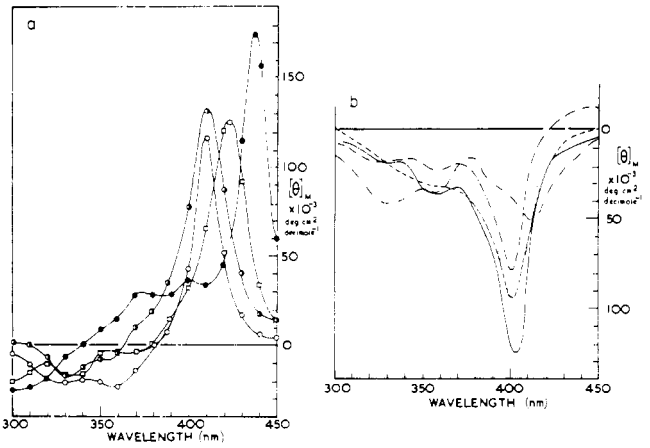

FIGURE 4: Soret-region CD spectra in phosphate buffer (0.02 *M*, pH 7.05) except where otherwise stated, and protein concentrations of 0.2–0.8%. (a) Sperm whale myoglobin: (○) ferric; (◑) fluoroferric; (□) cyanoferric; (●) ferrous. (b) Soybean leghemoglobin a: (—) ferric; (···) ferric in phosphate buffer (0.02 *M*, pH 5.4); (- - -) fluoroferric; (- · · ·) nicotinoferric; (- · - · -) cyanoferric.

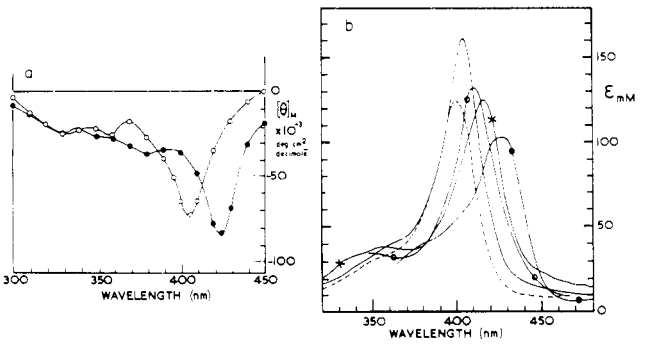

FIGURE 5: Soret-region spectra of soybean leghemoglobin a in phosphate buffer (0.02 *M*, pH 7.05). (a) CD spectra at protein concentrations of 0.2–0.4%: (○) oxyferrous; (●) ferrous. (b) Absorption spectra: (—) ferric; (- - -) fluoroferric; (-×-) cyanoferric; (-○-) oxyferrous; (-●-) ferrous.

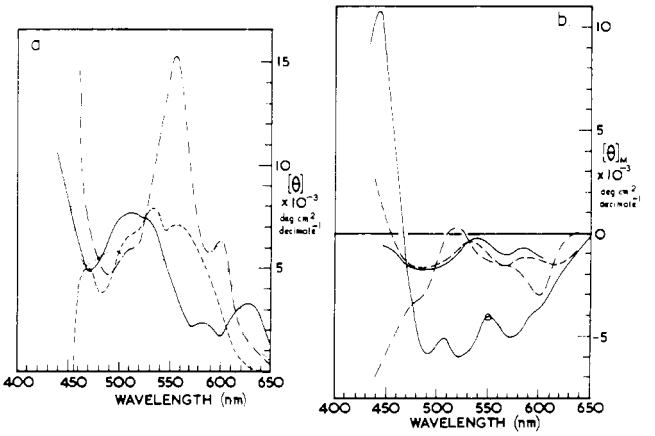

FIGURE 6: Visible-region CD spectra in phosphate buffer (0.02 *M*, pH 7.05) at protein concentrations of 0.2–0.7%. (a) Sperm whale myoglobin: (—) ferric; (···) fluoroferric; (- - -) cyanoferric; (- · -) ferrous. (b) Soybean leghemoglobin a: (- - -) ferric (pH 7.05); (—) ferric (pH 5.4); (- · -) fluoroferric (pH 7.05); (-○-) nicotinoferric.

trast, the Soret CD spectrum for each soybean Lb a derivative (Figure 4b) displays a major negative extremum which is usually at 400–402 nm (cyanoferric Lb alone has a peak at 413 nm).

The Soret CD spectra for Lb's from other plant species are not only qualitatively but also quantitatively very similar despite very large differences in amino acid sequence (J. G. R. Hurrell et al., in preparation). The heme environment reflected by these spectra may be responsible for the very high oxygen affinities of this class of hemoglobins.

The oxyferrous form of soybean Lb shows Soret CD extrema which are similarly located to the ferric form (404–405 nm) but about half their intensity (Figure 5a). Removal of O_2 to give the ferrous forms shifts the extrema to 424 nm with little change in intensity. The locations of the main CD maxima of the various Lb derivatives are similar to but not identical with the wavelengths of maximal absorption (Figure 5b) and are not so dependent upon the liganded state as are the absorption spectra. However, the 20 nm red-shift on deoxygenation is seen in both the CD and absorption spectra (compare Figure 5a and b).

Visible-Region CD and Absorption Spectra. Sperm whale Mb shows a CD spectrum in the visible region which contains several positive bands (see Figure 6a) whose positions and magnitudes vary with the oxidation state of the iron (ferric or ferrous) and the nature of the ligand (fluoro or cyanoferric). The CD spectra for soybean ferric Lb derivatives in the same region are negative (especially the nicotinates), and again show evidence of three or four component bands (Figure 6b), with little pH dependence, while the oxyferrous and ferrous forms of the Lb have visible CD spectra which have both positive and negative bands (Figure 7a). The absorption spectra of the ferrous and oxyferrous derivatives of soybean Lb a are shown in Figure 7b, and those of the ferric form and its complexes in Figure 8. The cyanoferric (low spin) complex has a single peak at 539 nm. In the nicotinoferric complex, however, this splits into a peak and shoulder, while the fluoroferric and aquoferric proteins (high spin) show four additional smaller components. On oxygenation of ferrous Lb a (Figure 7b), the single peak at 558 nm splits into two peaks of nearly equal intensity at 537 and 572 nm.

Discussion

The types of structural information obtainable from CD spectra depend on the wavelength region studied. While the far-uv CD provides information about the conformation of the backbone peptide chain, the CD characteristics in the near-uv (250–300 nm) depend upon the environment and rigidity of the chromophoric protein side chains. For hemoproteins the CD at wavelengths above 300 nm and particularly in the Soret region (300–450 nm) represents the optical activity of the heme group. Since heme itself is optically inactive the CD pattern obtained for it in the protein gives information about its interactions with its asymmetric environment.

The Far-Uv. CD spectra in the far-uv show that the plant Lb's have similar high α-helix contents ($\simeq$50%)—considerably higher than most globular proteins and approaching that of sperm whale Mb (72.5%).

It will be pointed out later in discussing Soret CD spectra that conventional estimates of α-helix content of heme proteins arrived at as above may well be subject to correction for changes in far-uv ellipticities of aromatic residues due to electronic coupling with heme residues. Such effects would make estimates of helix content low in Lb's and high in Mb's.

X-ray studies on lupin Lb at 5-Å resolution (Vainshtein et al., 1975) have shown a very similar tertiary structure to

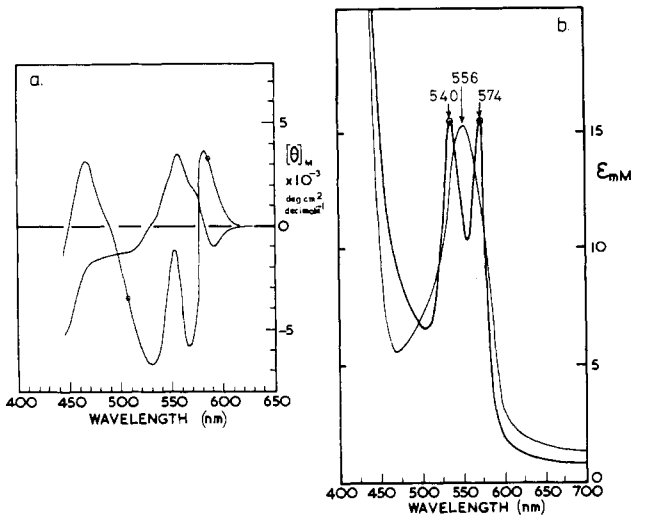

FIGURE 7: (a) Visible-region CD spectra of the oxyferrous and ferrous derivatives of soybean leghemoglobin a in phosphate buffer (0.02 *M*, pH 7.05). Concentrations, 0.2–0.4%. (—) Ferrous; (-O-) oxyferrous. (b) Absorption spectra of soybean leghemoglobin a in the visible at pH 7.05: (—) ferrous; (-O-) oxyferrous. The peak wavelengths shown are corrected for calibration of the spectrophotometer and do not therefore correspond exactly with the abscissa values.

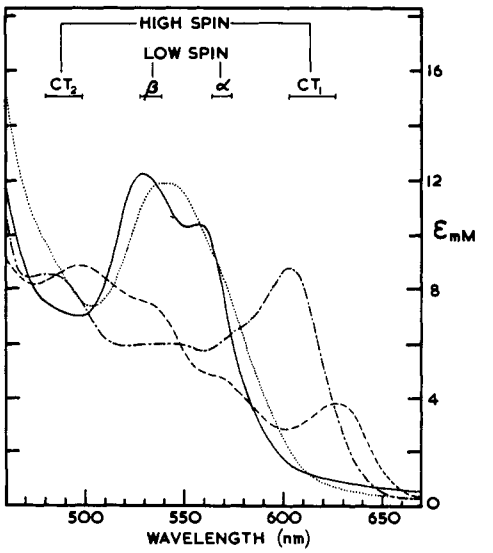

FIGURE 8: Absorption spectra of soybean leghemoglobin a derivatives in the visible at pH 7.05: (- - -) ferric; (- · -) fluoroferric; (—) nicotinoferric; (· · ·) cyanoferric. The bands labeled CT_1 and CT_2 occur only in ferric high spin complexes and are charge-transfer bands between the iron and the heme.

that of sperm whale Mb, with eight helical segments having similar relative orientations in the two proteins and almost the same number of residues in helical conformations. This is despite the considerable evolutionary distance and amino acid sequence differences between them.

The removal of the heme moiety from soybean Lb a decreases the α-helix content but is not so disruptive as the effect of removing heme from Mb. If one hypothesizes[2] that the heme pocket in Mb is more hydrophobic and the fit is tighter than in Lb's, one would expect that heme removal from Mb would be more disruptive than from Lb. Heme transfer experiments also suggest that the heme is relatively "loose" in Lb's (Ellfolk et al., 1973).

The Near-Uv. The large positive CD band centered around 260 nm in the near-uv CD spectra of Mb (Figure 2a) varies in size with the different Mb derivatives. As it is absent when the heme is removed it probably arises from the induced optical activity of the heme. Urry (1967) has reported a similar positive CD band in other hemoproteins and in the heme undecapeptide of cytochrome *c* which contains no aromatic residues. Because of this background CD in Mb it is difficult to decide whether the aromatic fine structure superimposed on it is positive or negative. However, if we assume that it is negative (Urry, 1967) then we see (Figure 2a) that all the Mb derivatives have a negative peak at 289 nm. This peak very probably corresponds to the 0–0 transition of the ^1L_b band of tryptophan which is red-shifted to 287 nm in a nonpolar solvent and can shift as far as 293 nm for a hydrogen-bonded tryptophan in such an environment (Strickland et al., 1972). All the Mb derivatives (including the apoprotein) show also a positive band at 295 nm. This is probably the ^1L_a 0–0 transition of the hydrogen-bonded tryptophan. Since this is a separate transition from the ^1L_b there is no reason why it cannot be opposite in sign. Indeed this transition is the only possible source of such a band in the apoprotein. The corresponding band in ferricytochrome *c* is at 293 nm and the single tryptophan in this protein is known to be hydrogen bonded to a propionic carboxyl group in a nonpolar region of the molecule (Strickland et al., 1971; Takano et al., 1973). In Mb where one of the tryptophans is hydrogen bonded to a glutamine the ^1L_a 0–0 band is predicted to occur between 295 and 298 nm from model studies (Strickland et al., 1972) since this band is more sensitive to hydrogen bonding and is broader than the ^1L_b band. These data are all in accord with the x-ray structure of Mb (Kendrew et al., 1960; Kendrew, 1962).

It is not possible at this stage to decide whether any changes in aromatic side-chain conformation occur with the different complexes. This may be possible if the heme background can be estimated and subtracted out. The positioning of the fine structure and its relative size suggest that any changes would be minor ones.

The near-uv CD spectra for soybean Lb a and its derivatives (Figures 2b, 3a, and 3b) appear to be superimposed on a negative CD baseline of about -5 to -15×10^3 deg cm^2 dmol^{-1}. This baseline is the residual CD from the low-wavelength tail of the negative N band (see later under "The Soret Region"). In contrast to Mb, the positive heme band centered around 260 nm is comparatively weak for all Lb spectra.

The fine structure bands due to the aromatics in the Lb spectra are consequently more clearly defined and, if the assignments for Mb are correct, show an interesting reversal relative to Mb, with a positive peak at 289 nm and a negative one at 295 nm in all the derivatives including the apoprotein. These bands would appear to have a similar origin to those in Mb, their positions again suggesting that a hydrogen-bonded tryptophan in a nonpolar environment is giving rise to the peaks. This interpretation accords very well with model compound studies of *N*-acetyltryptophanamide in MeOH at 77 K (Strickland et al., 1969). The inset in Figure 2b shows that in MeOH the ^1L_a 0–0 band occurs at 296 nm and is negative. The ^1L_b 0–0 band is sharp and positive at 290 nm. The positive band at 283 nm probably arises from the ^1L_b (0 + 850) band. The remaining bands

[2] This hypothesis is borne out convincingly by thermal perturbation difference spectra (Nicola et al., 1974; N. A. Nicola et al., in preparation).

are ^1L_a transitions. As pointed out earlier, the ^1L_a transitions are known to be more sensitive to hydrogen bonding and to solvent than are the ^1L_b transitions, so their separation and sharpness in Lb is indicative of a fairly rigid hydrogen-bonded tryptophan.

Similar types of CD spectra have been recorded for chymotrypsinogen A (Strickland et al., 1969), lysozyme (Teichberg et al., 1970), cytochrome *c* (Myer, 1968), and the hemoglobin III component from *Chironomus thummi thummi* (Wollmer and Buse, 1971). In all these cases there is a broad negative ^1L_a 0–0 band at about 297 nm and a sharper positive ^1L_b 0–0 band at about 290 nm and they are attributed to rigid tryptophans, probably hydrogen bonded in a nonpolar environment. Only one of the two tryptophans in Lb appears to be primarily responsible for the CD pattern since the one freely accessible to aqueous environment (Nicola et al., 1974; Nicola et al., in preparation) would have bands in different positions and would be broader (Strickland et al., 1969). It is not possible to say anything about tyrosine or phenyalanine CD bands in Lb. These appear to be less important than those for tryptophan although phenylalanine fine structure can be detected in the spectra.

In the soybean apoprotein the leading near-uv CD bands appear to have moved 1–1.5 nm to longer wavelengths on removing heme. Comparison with the similar spectra for the fluoroferric complex of the holoprotein indicates that this shift does not necessarily reflect the loss of heme background CD but rather a local conformational change when the heme is removed, possibly the further burial of the responsible tryptophan residue. The apoprotein CD spectrum also proves that the fine structure bands do not result from tryptophan–heme interactions. In Mb at least, Hsu and Woody (1971) have shown that the tryptophans and the heme group do not interact to produce any CD bands.[3]

The effect of decreasing the pH from 7 to 5.4 on the near-uv CD patterns for the soybean protein is to move the two major bands at 281 and 289 nm to 283 and 290.5 nm, respectively (Figure 3a). This may reflect a small conformational alteration involving mainly the tryptophan and its local environment similar to that which occurs on removal of heme. The positions of these bands and their shift with pH are not affected by the nature of the ligand in the sixth coordination position.

The Soret Region. The heme Soret absorption band (B band) is the resultant of two nearly degenerate electronic transitions polarized perpendicularly to each other and directed along the methine bridges of the heme group in Mb. The symmetric and otherwise optically inactive heme group acquires optical activity when bound in an asymmetric environment, as it is in the heme proteins, probably by a coupled oscillator interaction between the two electronic transitions of the heme and the allowed π–π^* transitions in nearby histidine, phenylalanine, and tyrosine side chains (Hsu and Woody, 1971). The resultant rotational strengths depend on the coupling of many residues and it is therefore difficult to relate the resultant heme CD to the detailed heme environment. For the same reason, however, heme CD measurements make a very sensitive probe for differences or changes in that environment.

In Mb the two Soret transitions acquire rotational strengths which are opposite in sign and whose relative intensities depend on the orientations chosen (within the heme plane and mutually perpendicular) for the two transitions. The net rotational strength of the Soret transition does not vary with the choice of orientation but the degree of band splitting does. Hsu and Woody found the best fit to the experimental myoglobin spectrum with the two Soret transitions aligned along the methine carbons rather than the pyrrole nitrogens. It is clear, however, that a change in the orientation of the transition moments and a different pattern of amino acid contacts can produce wavelength shifts relative to the absorption spectrum and also cause the negative CD component to predominate over the positive. The L band of the heme group at about 260 nm also acquires positive rotational strength of about half the size of the Soret band, from the same aromatic coupling mechanism, in agreement with the experimental results in Figure 2a.

The Soret-band CD spectra for Mb derivatives do not agree well with those of Nagai et al. (1969) but are in very good agreement with those of Willick et al. (1969). In all the derivatives the major Soret CD band is positive but there seems to be no correlation between the size of the peak in the absorption spectrum (Willick et al., 1969) and its size in the CD spectrum, unless it is an inverse one. There is clear evidence for several component bands in the Soret region as also noted by Willick et al. (1969) who proposed a minimum of five components to resolve the curves for several derivatives of Mb. Clearly, some of these bands are negative and some positive. The N band for the heme group in Mb was placed by Willick et al. at about 340 nm and theoretical predictions indicate that it should be negative (Hsu and Woody, 1971). Figure 4a shows that this is indeed the case. The band at 370 nm is probably a charge-transfer band (Willick et al., 1969).

It is difficult to determine from the Soret CD data whether any conformational changes occur on binding the different ligands to Mb. This would require calculation of the net rotational strength in each case and comparison with the dipole strength. Inspection of the spectra is not enough because different ligands will alter the orientation of the B_x and B_y Soret transitions and this will cause a change in shape of the CD spectrum without implying a conformational change. The parameter which should not change in the absence of conformational changes is the ratio of the summed rotational strength for the bands to the summed absorption intensities. A simple test is to record the results as $[\theta]_\lambda/\epsilon_\lambda$ against wavelength. Such a curve can indicate the presence of hidden bands when these have a different anisotropy to the main bands (see Figure 9 for fluoroferric Mb).

Any departure from a constant value for the ratio indicates the presence of a band which has a rotational strength disproportionate to its absorption intensity relative to the other bands. A proper analysis would require simultaneous resolution of both absorption and CD spectra into say Gaussian components. It is apparent, however, from Figure 9 that for fluoroferric Mb the anisotropy ratio varies considerably while that for fluoroferric Lb is relatively constant. If one takes an average value of about +1 for the Mb

[3] A referee has pointed out that recent calculations by Hsu and Woody show that both tryptophans, but especially tryptophan-14, *do* contribute to the heme Soret transition optical activity. Since the tryptophans are, however, too far away from the heme for direct contact, the heme cannot be producing the environment discussed here. Furthermore, since the near-uv peaks for the apoprotein are nearly as intense as for the heme complexes, it would seem that at least the *near*-uv bands of tryptophan do not acquire their intensity through coupling with heme transitions.

complex and −1 for the Lb complex this may indicate that the B_x and B_y components of the Soret CD band acquire rotational strengths of opposite sense for Mb as predicted by Hsu and Woody, whereas both components acquire negative strengths in Lb. Both hemoproteins acquire negative rotational strengths for the N band near 325 nm, the Mb result agreeing with the prediction of Hsu and Woody (1971).

The major feature of the Soret CD spectra for soybean Lb a derivatives (Figure 4b) is their opposite sense to those of Mb. Except perhaps for a band at higher wavelengths, all bands are negative. In addition, the relative sizes of the CD bands for the various complexes are different in the two proteins. The Soret band for the ferric form of Lb a is more intense than for other derivatives of Lb a.

Since the Soret Cotton effect depends on the arrangement of many of the aromatic residues around the heme and because the net CD curve can be the resultant of two much larger curves of opposite sign (Hsu and Woody, 1971) such large apparent differences between Mb and Lb do not necessarily imply a large difference in the environment of the heme in the two proteins. It is surprising, in fact, that the magnitudes of the Cotton effects in both proteins are of the same order even though of opposite sense. Negative Cotton effects for the Soret band occur also in cytochrome *c* peroxidase (Willick et al., 1969) and in Hb III from *Chironomus thummi thummi* (Gersonde et al., 1972). Lamprey Hb shows a negative Soret Cotton effect and this can become positive when nitrosobenzene or alkyl isocyanides are bound. All of these curves are of about the same order of magnitude (Lampe et al., 1972). It is apparent therefore that fairly subtle changes can cause the negative B_y component to be larger than the positive B_x component.

Bearing this in mind, a comparison of the Soret Cotton effects in some other species of Lb's shows very striking similarities in both shape and size (J. G. R. Hurrell et al., in preparation) indicating a very similar heme environment for all these proteins. Further analysis must await curve resolution; preliminary attempts to fit common component bands to both absorption and CD spectra have failed to obtain unique solutions.

An interesting consequence of the proposed mechanism by which the heme acquires its optical activity in the Soret region is that any rotational strength gained by the heme group must be accompanied by an equal but opposite-in-sign rotational strength gained by the responsible aromatic chromophore. However, since the dipole strengths of aromatic transitions in the near-ultraviolet are small, the contribution of heme coupling to their CD must also be small (Hsu and Woody, 1971). Certainly, the experimental evidence indicates that the area under the near-uv CD bands is not equal to the area under the Soret CD (compare the scales in Figures 2 and 4a). The major source of the Soret CD band must thus be coupling with the far-uv aromatic transitions which have much greater dipole strengths. Since these transitions occur in the same region as the peptide bond CD used for the estimation of secondary structure, changes in the nature of the sixth heme ligand or removal of the heme must be accompanied by small changes in the far-uv CD without *necessarily* signifying a conformational change in the backbone.

As pointed out earlier, conventional estimates of α-helix content based upon the magnitude of $[\theta]_{220-222}$ should be corrected for possible contributions from changes in aromatic ellipticity in the far-uv caused by coupling with the heme. The simplest correction would assume that all such changes occur near 220 nm, although Hsu and Woody (1971) have shown that only a few of the responsible transitions are in this region. The positive Soret ellipticity of Mb would be offset by an equal but negative contribution in the far-uv CD, exaggerating the magnitude of $[\theta]_{222}$ by up to 5%. On the other hand, the negative Soret ellipticity of the Lb's could reduce the magnitude of their $[\theta]_{220-222}$ by a similar amount. This procedure clearly represents an over correction.

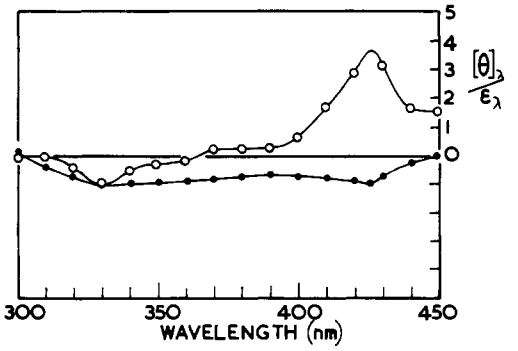

FIGURE 9: The anisotropy ratio (ellipticity divided by absorption at each wavelength) for the fluoroferric complexes of: (○) myoglobin and (●) soybean leghemoglobin a.

There are three consequences of the difference in *sense* of contribution. First, the difference in estimated helix contents between Mb and the Lb's may not be as great as $[\theta]_{220-222}$ values suggest. Second, the effects of removing heme from the two types of protein may be exaggerated when changes in $[\theta]_{220-222}$ are measured without correcting for the loss of heme-coupled aromatic ellipticity. Finally, changes in $[\theta]_{220-222}$ observed when the sixth ligand is substituted will also include a small component due to changes in the Soret region ellipticity.

We might finally mention that the mechanism proposed by Hsu and Woody for Soret optical activity, although it does explain the main features mentioned above is not the only possible one. Distortion of the heme plane when it is bound in the hemoprotein could generate optical activity and, although there is no experimental evidence for this in hemoproteins, there is x-ray evidence for formation of a domed structure under environmental changes (Hoard, 1968).

The Visible Region. The visible bands (Q transitions) of low-spin ferric hemes arise from transitions from the a_u π orbitals to the doubly degenerate e_g π^* orbitals. Because of the approximate D_{4h} square symmetry, the Q transitions are doubly degenerate and point in mutually perpendicular directions. If symmetry is reduced to D_{2h} then this degeneracy is released and four visible bands may be seen: $Q_{a_u}{}^x \rightarrow e_g$ 0–0, 0–ν, $Q_{a_u}{}^y \rightarrow e_g$ 0–0, 0–ν where ν represents higher vibrational levels of the e_g state. In most hemes, however, one sees two bands, the Q 0–0 (α band) and the Q 0–ν (β band). In high-spin ferric hemes, additional charge-transfer bands (porphyrin to metal), which borrow intensity from the α and β bands, can be seen at longer and shorter wavelengths (see Figure 3b).

The visible CD patterns for Mb derivatives (Figure 6a) all show only positive bands as predicted by Hsu and Woody (1971) and are of the predicted sizes. The CD patterns reveal bands which are not clearly defined in the absorption spectra (compare the cyanoferric and ferrous CD

spectra for example with the absorption spectra in Figure 8. These are for Lb but Mb spectra are nearly identical). In each case there is evidence for at least three and possibly four bands showing the decrease in symmetry of the heme group on combination with the protein. There is quite a good relationship between the size of each component absorption band and its corresponding CD band. This is to be expected if the direction of the transition moment for each band is the same.

There is also a reasonable correlation between the ratios of ellipticities to absorption intensities from one complex to another except for deoxy Mb in both the Soret and visible regions (see Conclusions).

As in the Soret region, the visible CD patterns for Lb's are mainly negative, particularly for their ferric derivatives (Figure 6b). Their oxyferrous and ferrous derivatives, however, have both positive and negative bands (Figure 7a). From the absorption spectra for oxyferrous and ferrous Lb's (Figure 7b) it appears that the x and y components of the Q transitions (540–570 nm) are very nearly degenerate since only two absorption bands are clearly seen. The type of CD spectra obtained for these complexes in Lb might then be explained if the x and y components each gained rotational strengths of opposite sign (as for the Soret band). The separation in position of the positive and negative components does not appear to occur in most of the other Lb complexes. This might be explained by small conformational differences between the complexes or merely changes in direction of the Q transition moments caused by the ligand and/or any change in environment as discussed for the Soret transitions.

Conclusions

The CD studies described in this paper show that the overall folding of the Mb and Lb chains is similar; the aromatic residues also exhibit similar CD spectra. The major spectral differences between the two heme proteins stem from the heme moiety and the way in which it interacts with protein side chains in the heme pocket. These differences show up especially in the near-uv around 260 nm and in the Soret region around 400 nm.

Ruckpaul et al. (1970) have postulated that the magnitude of the heme CD envelope centered around 260 nm reflects the strength of interaction of the heme with the apoprotein. Hsu and Woody (1971) have shown that the heme derives its optical activity by a coupled oscillator mechanism with the far-uv transitions of nearby aromatic residues. Since this mechanism is quite sensitive to the distance between the heme and protein side chains, it is not surprising that the 260 nm CD intensity should reflect the packing around the heme group. In fact the *difference* CD intensity at 260 nm between deoxy and oxy complexes should give an indication of the extent to which the interactions between the protein side chain and the heme change upon oxygenation. A survey of a number of hemoproteins indeed shows a rough correlation between oxygen affinity and the magnitude of the CD change at 260 nm. Thus, Figures 2 and 3 show that the size of this band is spin-state dependent, being more positive with the low-spin derivatives (e.g., cyanoferric and oxyferrous) than with the high-spin derivatives (e.g., fluoroferric and deoxyferrous). This may be correlated with the movement of the proximal histidine relative to the heme group when the iron atom changes from the high- to the low-spin state (Hoard, 1968; Perutz, 1970) leading to a more intimate contact between the heme and adjacent protein residues. If one compares the magnitude of the 260-nm heme band for Mb and Lb's in the same spin state, or the magnitude of the change when there is a change in spin state, then it would appear that the heme is more strongly bound and in closer contact with the protein side chains in Mb than in Lb. The structural reorientations of side chains occurring as a consequence of spin-state changes or oxygen binding are therefore greater for M than for Lb (cf. oxy vs. deoxy complexes) and this may be a reason for the higher oxygen affinities of Lb's (see also the discussion under "Far-uv"). For deoxy Mb but not for deoxy Lb the ratio of ellipticity to absorption intensity is considerably higher in both the Soret and visible regions than for other complexes. This implies a greater asymmetry in deoxy Mb than in deoxy Lb, a situation which might be expected if the iron atom were further out of the heme plane in the former case. Perutz et al. (1974 a–c) has suggested that any intramolecular forces which hold the iron out of plane tend to decrease oxygen affinity by opposing the movement of the iron atom back into the plane when oxygen is bound. This would be consistent with the finding that the heme moiety is more loosely held in Lb's than in Mb and that Lb's have the higher oxygen affinity. The greater flexibility of the heme pocket in Lb than in Mb is also consistent with the capacity of Lb to bind bulky ligands like long-chain fatty acids (Ellfolk, 1961) and nicotinic acid (Appleby et al., 1973) which Mb is unable to do and also the greater accessibility of the Lb heme moiety to bulk water (Nicola et al., 1974; N. A. Nicola et al., in preparation).

References

Appleby, C. A., Nicola, N. A., Hurrell, J. G. R., and Leach, S. J. (1975), *Biochemistry* (in press).

Appleby, C. A., Wittenberg, B. A., and Wittenberg, J. B. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 564–568.

Breslow, E., Beychok, S., Hardman, K. D., and Gurd, F. R. N. (1965), *J. Biol. Chem. 240,* 304–309.

Chen, Y.-H., Yang, J. T., and Chau, K. H. (1974), *Biochemistry 13,* 3350–3359.

Chen, Y.-H., Yang, J. T., and Martinez, H. M. (1972), *Biochemistry 11,* 4121.

de Duve, Ch. (1948), *Acta Chem. Scand. 2,* 264–289.

Ellfolk, N. (1960), *Acta Chem. Scand. 14,* 1819–1827.

Ellfolk, N. (1961), *Acta Chem. Scand. 15,* 975–984.

Ellfolk, N., Pertila, U., and Sievers, G. (1973), *Acta Chem. Scand. 27,* 3601–3602.

Ellfolk, N., and Sievers, G. (1971), *Acta Chem. Scand. 25,* 3532–3534.

Gersonde, K., Sick, H., Wollmer, A., and Buse, G. (1972), *Eur. J. Biochem. 25,* 181–189.

Hoard, J. L. (1968), in Structural Chemistry and Molecular Biology, Rich, A., and Davidson, N., Ed., San Francisco, Calif., W. H. Freeman, pp 573–594.

Hsu, M.-C., and Woody, R. W. (1971), *J. Am. Chem. Soc. 93,* 3515–3525.

Kendrew, J. C. (1962), *Brookhaven Symp. Biol. 15,* 216–228.

Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C., and Shore, V. C. (1960), *Nature (London) 185,* 422–427.

Lampe, J., Rein, H., and Scheller, W. (1972), *FEBS Lett. 23,* 282–284.

Lehtovaara, P., and Ellfolk, N. (1974), *FEBS Lett. 43,* 239–240.

Myer, Y. P. (1968), *J. Biol. Chem. 243,* 2115–2122.
Nagai, M., Sugita, Y., and Yoneyama, Y. (1969), *J. Biol. Chem. 244,* 1651–1653.
Nicola, N. A., Appleby, C. A., Leach, S. J., and Minasian, E. (1974), *Proc. Aust. Biochem. Soc. 7,* 21.
Paul, K. G., Theorell, H., and Åkesson, Å. (1953), *Acta Chem. Scand. 7,* 1284–1287.
Perutz, M. F. (1970), *Nature (London) 228,* 726–734.
Perutz, M. F., Fersht, A. R., Simon, S. R., and Roberts, G. C. K. (1974b), *Biochemistry 13,* 2174–2186.
Perutz, M. F., Heidner, J., Ladner, J. E., Beetlestone, J. G., Ho, C., and Slade, E. F., (1974c), *Biochemistry 13,* 2187–2200.
Perutz, M. F., Ladner, J. E., Simon, S. R., and Ho, C. (1974a), *Biochemistry 13,* 2163–2173.
Richardson, M., Dilworth, M. J., and Scawen, M. D. (1975), *FEBS Lett. 51,* 33–37.
Rossi-Fanelli, A., Antonini, E., and Caputo, A. (1958), *Biochim. Biophys. Acta 30,* 608–615.
Ruckpaul, K., Rein, H., and Jung, F. (1970), *Acta Biol. Med. Germ. 24,* 445–454.
Schellman, J. A., and Lowe, M. J. (1968), *J. Am. Chem. Soc. 90,* 1070–1072.
Strickland, E. H., Billups, C., and Kaye, F. (1972), *Biochemistry 11,* 3657–3662.
Strickland, E. H., Horwitz, J., and Billups, C. (1969), *Biochemistry 8,* 3205–3213.
Strickland, E. H., Horwitz, J., Kay, E., Shannon, L. M., Wilchek, M., and Billups, C. (1971), *Biochemistry 10,* 2631–2638.
Stryer, L. (1965), *J. Mol. Biol. 13,* 482–495.
Takano, T., Kallai, O. B., Swanson, R., and Dickerson, R. E. (1973), *J. Biol. Chem. 248,* 5234–5255.
Teichberg, V. J., Kay, C. M., and Sharon, N. (1970), *Eur. J. Biochem. 16,* 55–59.
Urry, D. W. (1967), *J. Biol. Chem. 242,* 4441–4448.
Vainshtein, B. K., Harutyunyan, E. H., Kuranova, I. P., Borisov, V. V., Sosfenov, N. I., Pavlovsky, A. G., Grebenko, A. I., and Konareva, N. V. (1975), *Nature (London) 254,* 163–164.
Willick, G. E., Schonbaum, G. R., and Kay, C. M. (1969), *Biochemistry 8,* 3729–3734.
Wollmer, A., and Buse, G. (1971), *FEBS Lett. 16,* 307–310.
Woody, R. W., and Tinoco, I., Jr. (1967), *J. Chem. Phys. 46,* 4927–4945.

Amino Acid Sequence Studies on the α Chain of Human Fibrinogen. Location of Four Plasmin Attack Points and a Covalent Cross-Linking Site†

Takashi Takagi‡ and Russell F. Doolittle*

ABSTRACT: The amino acid sequence of a 38-residue midsection piece of the α chain of human fibrinogen has been determined using a combination of plasmin-derived peptides and cyanogen bromide fragments. The segment contains several important features, including four early plasmin attack points, one of the two α-chain cross-linking acceptor sites, and a peptide homologous to one isolated from plasmin digests of bovine fibrinogen, and reported to have anticoagulant activity. The segment is sequentially adjacent to and overlapping with a large molecular weight (20000–25000) fragment released during plasminolysis. This latter material is very rich in glycine and serine and deficient in nonpolar amino acids. It also contains the other α-chain cross-linking acceptor site.

During the course of structural studies aimed toward elucidating the entire covalent structure of human fibrinogen, we have deduced the amino acid sequence of a 38-residue segment from the central portion of the α chain. As it happens, this region of the molecule contains a number of important features—including four early plasmin cleavage points, one of the two α-chain cross-linking acceptor sites, and a peptide homologous to one isolated from plasmin digests of bovine fibrinogen and reported to have anticoagulant activity. The segment occupies a position in the α chain corresponding approximately to residues 200–250.

The sequence was determined using data from several different starting materials. First, the progressive plasminolysis of human fibrinogen was studied over a time course ranging from 10 min to 15 hr, the digestion stages being terminated by heat precipitating the incubation mixtures. This report deals with the material which remains in solution upon such treatment; we have previously reported observations dealing with the precipitated material (Takagi and Doolittle, 1975). Secondly, we have purified the same peptides from plasmin digests of fibrin containing an incorporated substitute cross-linking donor. Finally, we isolated the corresponding fragments from cyanogen bromide digestions of purified α chains. The combination of plasmin-derived peptides and cyanogen bromide fragments allowed us to find the necessary overlaps for proving the peptide ar-

† From the Department of Chemistry, University of California—San Diego, La Jolla, California 92037. *Received June 3, 1975.* A preliminary report of these findings was presented Nov 21, 1974 at the American Heart Association Meetings, Dallas, Texas. This work was supported by U.S. Public Health Service Grants HE-12,759 and GM-17,702.

‡ Present address: Biological Institute, Tohoku University, Sendai, Japan.