A Structure of Sperm Whale Myoglobin at a Nitrogen Gas Pressure of 145 Atmospheres†

Robert F. Tilton, Jr.,*,‡ and Gregory A. Petsko*
Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
Received July 1, 1987; Revised Manuscript Received November 20, 1987

ABSTRACT: A structure of sperm whale metmyoglobin under a nitrogen gas pressure of 145 atm (2200 psi) has been solved by X-ray diffraction using data to 2.0-Å resolution. The perturbation of the gas pressure on the overall structure of the protein is minimal with a root mean square deviation of backbone atoms between the pressurized and unpressurized structures of 0.22 Å. Additional electron density is observed, however, in two cavities of the protein molecule. The density is interpreted as a nitrogen molecule bound in the proximal cavity and as a water molecule hydrogen bonded in a separate cavity (cavity 3). In addition, alternate conformations are observed for three internal residues (Leu-135, Phe-138, and Ile-142) that border these cavities. These alternate conformations are not observed in atmospheric pressure structures and are presumed due to the effects of pressure and/or gas binding. The appearance of these alternate conformations implies a repacking of the protein interior and produces a new distribution of cavity spaces. The profile of the Debye–Waller factors for the pressurized structure is similar to that for the room pressure except for a small increase in the distal region (residues 61–69) of the protein.

Myoglobin and hemoglobin are known to absorb relatively inert gases including the common air gases N_2 and H_2 (Mortimer & Bauer, 1960), the noble gases argon and xenon (Keyes & Lumry, 1968; Ewing & Maestas, 1970; Conn, 1961; Catchpool, 1968), and several halogenated and unhalogenated hydrocarbons (Schoenborn, 1967; Brown & Halsey, 1980; Wishnia, 1969; Featherstone et al., 1961). Binding of these gases is generally weak with measured equilibrium binding constants less than 200 L/(mol·s). Dispersion interactions dominate the binding energy, contrasting these gases with those that coordinate directy to the heme iron (CO, NO, O_2). Of the inert gases, xenon has been the best characterized and has been a useful probe of interior spaces or cavities within proteins. Among the most tightly bound of the inert ligands, with a binding constant of 200 L/(mol·s) (Ewing & Maestas, 1970), xenon has been shown to bind in several interior sites of both myoglobin and hemoglobin (Schoenborn, 1965; Shoenborn et al., 1965; Tilton et al., 1984). These sites are preformed, interior cavity spaces or "packing defects" within the protein molecule that are not directly involved with the O_2 binding cavity (Figure 1). They are characterized by being surrounded predominately by hydrophobic residues and are located when secondary structure elements pack against each other or against the heme group. The different cavities do not, however, have identical gas binding affinities. Computations suggest that the binding energy (ΔE) varies between 5 and 10 kcal/mol (Tilton et al., 1986), in reasonable correspondence with the experimentally determined average binding enthalpy (Ewing & Maestas, 1970) of 7 kcal/mol and in qualitative correspondence with the experimental occupancies determined by X-ray crystallography (Tilton et al., 1984). The most straightforward variable to control the binding of gaseous ligands to proteins is the use of gas pressure—increasing gas pressure increases fractional binding. To investigate the binding of inert gas ligands with weak binding affinities, elevated pressures must be employed. In this paper we describe the results of an X-ray crystallographic study of sperm whale metmyoglobin under a nitrogen gas pressure of ~145 atm (2200 psi).

Materials and Methods

High-Pressure Crystallographic Cell. A high-pressure X-ray crystallographic cell has been designed specifically for high-pressure gas binding studies (Figure 2). The details of the high-pressure fixture are published separately (Tilton, 1988), and only a brief description follows. The cell consists essentially of two pieces: a shroud and a base. The shroud is machined from high-purity S-94B beryllium (Brush-Wellmann). Beryllium is the only material that combines both the high tensile strength (42 000 psi) and low X-ray absorption cross section necessary for such X-ray windows. This shroud is fastened to a stainless steel block by six hex screws with a metal compression C-ring serving as a gas-tight seal. Welded to the block is a modified high-pressure valve (Whitey Co.) that is used to regulate and maintain gas pressure. The mass of the cell is 110 g and is due almost entirely to the stainless steel block and valve assembly. The dimensions of the cell are 7.5 cm high and 9.0 cm across at the widest point. This pressure cell was designed for use on standard diffractometers and can be mounted directly on a standard arcless goniometer head. This prototype cell has an upper limit to the working pressure of 6000 psi (400 atm). No gas leakage is detected over a time period of 1 month. The portability of the pressure cell is unique, making it well suited for the collection of X-ray data from proteins under high-pressure gases.

Crystal Preparation. Lyophilized sperm whale metmyoglobin was purchased from Sigma and ultrapure ammonium sulfate was purchased from Bethesda Researh Laboratories. Crystals of metmyoglobin were grown from 75% saturated ammonium sulfate, pH 6.8 at 18 °C, as described by Kendrew and Parrish (1956). Crystals are rectangular in shape with dimensions 1.0 × 0.8 × 0.8 mm.

Crystal Mounting and Pressurization. Mounting crystals in the high-pressure cell requires only slight modifications to standard crystal mounting techniques. The protein crystal is briefly immersed in mineral oil and mounted between two

† Supported by NIH Grants GM-26788, GM-38758, and AM-07339.
‡ Present address: Department of Structural Chemistry, Miles Inc., 400 Morgan Lane, West Haven, CT 06516.

0006-2960/88/0427-6574$01.50/0

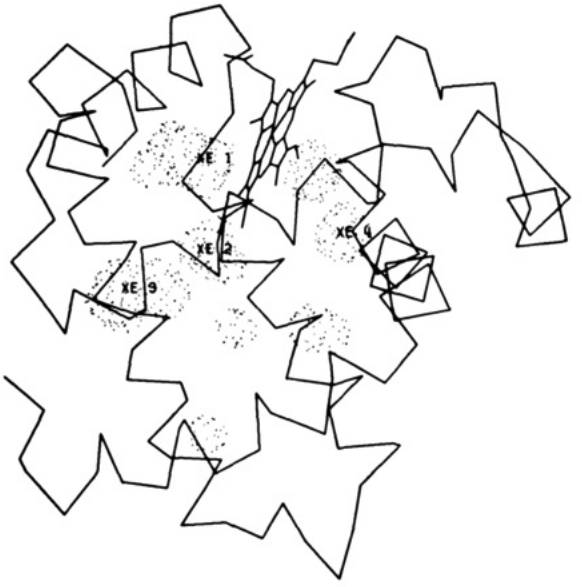

FIGURE 1: A C_α backbone tracing of sperm whale myoglobin. The heme group is encapsulated by the archetypal globin fold resulting in atom-sized packing defects or cavities. These cavities are indicated by dotted surfaces and are calculated with the Connolly algorithms using a 4.4-Å probe. Cavities 1–4 are observed to bind atoms of xenon while cavity 5, in the met form, binds a water molecule that coordinates to the iron of the heme group. Cavity 5 is directly across the heme group from cavity 1. Cavity 1 is also referred to as the proximal cavity and cavity 5 as the distal cavity.

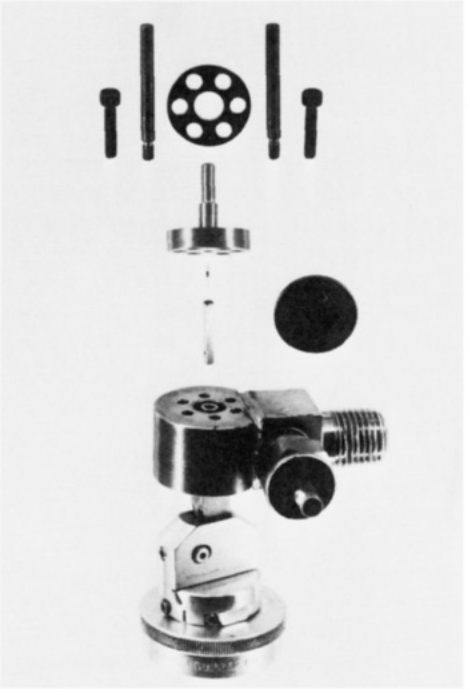

FIGURE 2: An "exploded" view of the high-pressure fixture. The hex screws, beryllium shroud, crystal mounting capillary and capillary mounting pin, valve mount, and goniometer are shown. The valve handle has been removed. Details of the fixture are published separately (Tilton, 1988).

moistened cotton plugs in a 1.0-mm quartz capillary (Charles Supper Co.). The oil causes no damage to the protein crystal and prevents the crystal from dehydrating. The cotton plugs are used to increase stability of the crystal and to keep the crystal from moving during the gas pressurization procedure. No mother liquor or oil plugs are used, and both ends of the capillary remain open. The capillary is inserted and glued into a small steel tube that fits directly into the block. The high-pressure cell and the holding tube are designed to facilitate simultaneous pressurization from both ends of the glass capillary.

Optical alignment of the crystal on the diffractometer is done before and beryllium shroud is lowered over the crystal. With the shroud secured, the cell is pressurized directly from a cylinder of N_2 gas and replaced on the goniostat. The crystal does not move significantly during this procedure, and the positioning is satisfactory to locate low-angle reflections (6–8°) and refine a low-angle matrix. The remainder of the "optical" alignment is done with X-rays by choosing two orthogonal high-angle reflections and their Friedel pairs and adjusting the goniometer slides and height adjustment until the refined two θ angles agree to within 0.04°. This manual adjustment is laborious but results in a high-angle matrix that correctly predicts the positions of reflections to at least 1.5-Å resolution.

Data Collection. Standard data collection techniques can be used with the high-pressure beryllium cell. Data collection was performed on a Nicolet P3 diffractometer using nickel-filtered Cu Kα radiation of wavelength 1.54 Å from a sealed tube source (26 mA, 40 kV). A 0.8° ω step scan was used with 24 s spent on the peak scan and 12 s spent on individual backgrounds collected 0.5° away on θ from the calculated peak position. The width at half-maximum height for the 020 reflection using the ω scan was 0.4°. Because of the protruding valve and bulky stainless steel block, data can only by collected on positive θ and positive χ of the four-circle diffractometer to avoid clipping reflections. This restricts Friedel pair collection to 20° on two θ (although positive equivalents can be measured out beyond 60°) and prevents the use of the Syntex LT-1 low-temperature device. Because of these restrictions only a unique set of reflections was measured to 1.7-Å resolution. In addition, an $h0l$ zone was measured to 1.5-Å resolution. Five standard reflections distributed through reciprocal space were collected every 300 reflections to monitor crystal stability and radiation damage. All of these data were collected at room temperature (18 °C) from a single crystal.

Data Reduction and Model Refinement. Measured intensities were adjusted for background, anisotropic absorption, Lorentz and polarization corrections (North et al., 1968), and radiation damage to generate a unique set of structure amplitudes. Refinement of these data was done by using the method of restrained least squares (PROLSQ) by Konnert and Hendrickson (1980). Target σ's used in restraining the geometry were 0.03 Å for distances involving bonds, 0.04 Å for distances involving angels, and 0.05 Å for distances involving torsional angles. The original model used in the refinement was a structure of metmyoglobin previously refined with data to 2.0-Å resolution collected at 250 K and 1 atm of pressure (R. F. Tilton, D. Ringe, and G. A. Petsko, unpublished results). No water molecules were included from the original model. Electron density maps were generated with FFT routines using syntheses with amplitudes $2F_o - F_c$ and $3F_o - 2F_c$ and phases from the previous round of refinement (Ten Eyck, 1977). These maps were displayed using the FRODO software (Jones, 1982). Solvent molecules and ligands were located visually and modeled as water molecules if they were in good hydrogen-bonding geometry or as sulfate counterions if the electron density indicated tetrahedral oxygen atoms. Alternate protein side-chain conformations were included if clear evidence existed for the two positions. The occupancy of each conformation was originally set to 0.5 and individually refined such that the sum of the occupancies was equal to 1 (Kuriyan et al., 1986). Solvent molecules were retained that had refined Debye–Waller factors less than 45 Å^2 with an occupancy fixed

at 1.0. Refinement of partial occupancy for solvent molecules was not attempted. Regions of the model that were poorly fit or weakly determined were reexamined with the aid of difference delete maps where the protein residues in question were deleted from the model coordinate list used to generate F_c and the phase angles.

Results

Data collection with the beryllium cell is routine except for the initial crystal alignment and geometric restrictions in data collection. However, diffraction data beyond 2.0-Å resolution are severely contaminated by the presence of three intense powder diffraction rings with peaks at 1.97, 1.79, and 1.73 Å. These peaks are clearly observed in an otherwise normal 2θ–ω background curve and are characteristic of the (100), (002), and the (101) reflections of beryllium microcrystals (JCPDS, 1951). While the peaks are reasonably sharp (less than 1.0° wide), high-pressure data collection beyond 2.0-Å resolution must be undertaken cautiously and these annuli of data must be discarded. In background–peak–background scans, the beryllium scatter would be observed typically as very different backgrounds on either side of the peak resulting in large negative background corrected intensities. Protein data from 1.75- to 1.5-Å resolution in the *h*0*l* zone appear normal. In this paper we use only data collected from 10.0 to 2.0 Å.

Crystals of metmyoglobin under 145 atm of nitrogen gas pressure have the symmetry of space group $P2_1$ with unit cell dimensions a = 64.57 Å, b = 30.81 Å, c = 34.84 Å, and β = 105.6°. These values are similar to those observed for crystals of metmyoglobin at one atmosphere. No change was observed in the diffracted peak width or shape before or after pressurization. Some caution is necessary, however, to ensure that the crystal does not dehydrate during the course of the experiment. Evidence for dehydration in early studies included both an abrupt change in the unit cell dimensions, but with no change in peak width, and as a slow gradual broadening of the peak width. The different behavior is presumed due to the degree or severity of the dehydration. Treatment of the crystal with a thin coating of mineral oil and the use of prehydrated gas eliminated the problem of crystal dehydration. A total of 9200 reflections were collected to 2.0-Å resolution. Of these reflections, 7114 had intensities greater than twice their standard deviation and were used in the refinement. A linear decay (0.12%/h) in the intensities of the standard reflections, independent of resolution, was observed over the 8-day collection. This is a normal decay rate for crystals of myoglobin at this temperature, indicating that pressurized nitrogen gas does not facilitate radiation-induced decay. All intensities were corrected back to zero time.

Reciprocal space refinement initially included data between 10.0 and 2.5 Å with a fixed Debye–Waller factor of 10.0 Å^2. The starting R factor was 38.1%. Higher resolution data were gradually included over 38 cycles until the R factor converged to 26.0%. Addition of individual Debye–Waller factors reduced the R factor to 23.1%. Inclusion of 152 solvent molecules, three alternate protein conformations, and two presumed bound ligands that were located by explicit examination of electron density maps resulted in a final R factor of 16.1% and a model with a final root mean square (rms) deviation in bond lengths of 0.026. A total of 83 cycles of refinement have been completed. All structural comparisons in this paper are with a model solved independently with data collected to 2.0-Å resolution at 300 K (Frauenfelder et al., 1979).

Protein Structure. The global structure of the protein is changed very little by the application of gas pressure. The rms difference between myoglobin (300 K) at 1 and 145 atm is 0.22 Å for main-chain atoms (CA, C, N) and increases to 0.53 Å with the inclusion of CB atoms. These values are typical of rms differences between refined myoglobin molecules solved with data collected from different crystals (Kuriyan, 1986). A plot of the rms deviation as a function of protein residue number is shown in Figure 3. The largest deviations are observed in the turns, loops, and carboxy and amino termini of the protein. These regions of the protein are typically the most poorly defined. The overall volume and shape of the molecule are unaffected by gas pressure. Both the radius of gyration (R_g) and the total molecular volume are nearly identical in the two models (Table I).

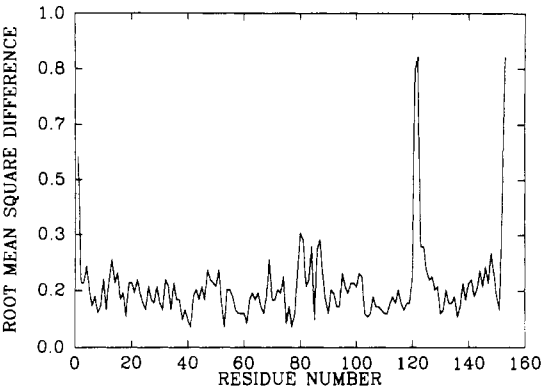

FIGURE 3: Root mean square (rms) deviation between a model of metmyoglobin at 1 atm and the model of metmyoglobin at 145 atm of N_2. The overall deviation is 0.22 Å for main-chain (C, CA, N) atoms. The largest deviation is observed in the GH corner of the protein.

Table I

	radius of gyration (Å)	
	1 atm (300 K)	145 atm (300 K)
CA backbone atoms	15.067	15.054
all atoms	15.087	15.065
	molecular volume (Å^3)[a]	
probe size (Å)	1 atm (300 K)	145 atm (300 K)
1.2	21897	21958
1.4	22111	22115
1.6	22684	22619

[a] Molecular volumes are calculated with the Connolly algorithms MS, MSSEP, and MSAV.

While the global features of the protein are maintained, local perturbations are observed under elevated N_2 gas pressures. In particular, three internal residues on the proximal side of the molecule are disordered and can be modeled with two partially occupied conformations. One of the conformations is similar to that observed in the "native" 1-atm structure while the second is in a strikingly different alternate conformation. These residues are Leu-135, Phe-138, and Ile-142 (Figure 4). All are completely buried residues and none have been reported to adopt the alternate conformations at atmospheric pressure. The dihedral angles of the native and alternate conformations as well as the refined occupancy values and average Debye–Waller factors are given in Table II. We note that the refined Debye–Waller factors for atoms in both the native and alternate conformations are similar and that the Debye–Waller factors for Phe-138 are substantially higher in the 145-atm structure than in a 1-atm structure solved independently (Frauenfelder at al., 1979).

Contacts between the native and alternate side-chain conformations with the surrounding protein indicate similiar steric interactions, although different protein residues are involved (Table III). The similarity in contact distances indicates that

Table II

residue	dihedral angle	myoglobin (145 atm of N_2) native conformation	myoglobin (145 atm of N_2) alternate conformation	myoglobin (1 atm), native conformation
Leu-135	N, CA–CB, CG	−65.7	−76.8	−75.0
	CA, CB–CG, CD1	171.1	−71.2	−177.0
	CA, CB–CG, CD2	−62.5	51.3	−51.2
	refined occupancy	(0.65)	(0.35)	(1.00)
	Debye–Waller factor	8.4	8.3	9.3
Phe-138	N, CA–CB, CG	159.4	−68.6	169.9
	CA, CB–CG, CD1	−82.3	170.8	−102.7
	refined occupancy	(0.55)	(0.45)	(1.00)
	Debey–Waller factor	15.6	16.3	10.7
Ile-142	N, CA–CB, CG1	−70.6	−64.7	−79.2
	N, CA–CB, CG2	167.7	174.0	160.8
	CA, CB–CG1, CD1	152.7	−60.1	180.0
	refined occupancy	(0.57)	(0.43)	(1.00)
	Debye–Waller factor	7.5	7.5	10.0

Table III[a]

residue	atom	distance	residue	atom
Leu-135	CG	3.21	Met-131	O
Leu-135	CBX	3.36	Met-131	O
Leu-135	CBX	3.25	Glu-136	N
Leu-135	CGX	3.47	Met-131	O
Leu-135	CD2X	2.79	Met-131	O
Leu-135	CD2X	3.49	Met-131	CG
Leu-135	CD2X	3.23	Met-131	SD
Phe-138	CB	3.11	Arg-139	N
Phe-138	CD1	3.35	Arg-139	N
Phe-138	CD1	3.40	Ile-142	CD1
Phe-138	CE1	3.05	Ile-142	CD1
Phe-138	CZ	3.15	Ile-142	CD1
Phe-138	CGX	3.38	Ala-134	O
Phe-138	CD1X	3.42	Leu-72	CD1
Phe-138	CD2X	3.50	Ala-134	O
Phe-138	CE1X	3.16	Leu-72	CD1
Phe-138	CE1X	3.19	Leu-76	CD2
Phe-138	CE2X	3.02	Ile-75	CG2
Phe-138	CE2X	3.07	Wat-166[b]	O
Phe-138	CZX	3.04	Ile-75	CG2
Phe-138	CZX	3.48	Leu-76	N
Phe-138	CZX	3.15	Leu-76	CD2
Ile-142	CB	3.07	Ala-143	N
Ile-142	CG1	3.43	Phe-138	O
Ile-142	CG2	3.39	Ala-143	N
Ile-142	CG2	3.39	Ala-143	N
Ile-142	CG2	3.46	Tyr-146	CE2
Ile-142	CD1	3.42	Phe-138	O
Ile-142	CD1	3.40	Phe-138	CD1
Ile-142	CD1	3.05	Phe-138	CE1
Ile-142	CD1	3.15	Phe-138	CZ
Ile-142	CG1X	3.31	Phe-138	O
Ile-142	CG2X	3.44	Ala-143	N
Ile-142	CG2X	3.43	Tyr-146	CE2
Ile-142	CD1X	3.31	Leu-86	CG
Ile-142	CD1X	3.23	Leu-86	CD2
Ile-142	CD1X	3.31	Ala-90	CB

[a] Atom names ending in X denote the alternate conformation. [b] Wat-166 is believed to be a water molecule bound in cavity 3 of the protein.

these alternate conformations do not sterically interfere with the surrounding native protein. The volume necessary to accommodate the alternate side-chain conformations is already present in the native packing of the protein, and no correlative motion between amino acid side chains need be hypothesized. However, we are unable to determine the relationships between the alternate conformations, i.e., to distinguish if all three residues have alternate conformations in the same protein or if some protein molecules have two and other protein molecules

Table IV

protein conformation[a]	probe size (Å)	cavity volume[b] (Å^3)	no. of cavities
135, 138, 142	1.2	382.1	10
	1.4	213.6	6
	1.6	96.9	3
135X, 138, 142	1.2	375.0	10
	1.4	215.0	6
	1.6	78.5	2
135, 138X, 142	1.2	439.9	8
	1.4	304.2	4
	1.6	183.9	2
135, 138, 142X	1.2	376.0	10
	1.4	214.3	6
	1.6	96.7	2
135X, 138X, 142	1.2	431.1	8
	1.4	294.0	4
	1.6	172.0	2
135X, 138, 142X	1.2	368.9	10
	1.4	215.7	6
	1.6	78.5	2
135, 138X, 142X	1.2	450.3	8
	1.4	318.7	4
	1.6	221.3	2
135X, 138X, 142X	1.2	441.7	8
	1.4	308.4	4
	1.6	210.5	2

[a] X denotes alternate conformation. Ligands have been removed in all calculations. [b] Cavity volumes calculated with Connolly algorithms.

Table V[a]

cavity	native conformation	alternate conformation
1	142, 104, 89, 93, 90, 138 heme	+86
2	107, 138, 108	+111, +139
3	7, 75, 72, 76, 135, 137, 138, 82, 134	−7, −76, −137, −134, −82
4	68, 25, 65, 28, 29, 69	no change
5	68, 29, 43 heme	no change
6	14, 17, 28, 115	no change

[a] Residues contributing to the surface of cavities were computed with the Connolly algorithms and a 1.4-Å probe. Native refers to the protein structure with side-chain conformations observed at 1 atm. Alternate refers to the protein structure with the alternate conformation observed at 145 atm of N_2. Only additions and deletions to the residue list of the native structure are listed for the alternate conformations. Cavities 1, 2, and 3 of the native structure merge to form one large cavity in the structure with alternate conformations.

one etc. Because of these ambiguities we must examine the eight possible conformational isomers that can be generated from the three alternate side-chain conformations. To visuallize the effects of the alternate conformations on packing defect distribution, we use the Connolly algorithms that implement the Richards definition of probe-accessible surfaces (Richards, 1977) to compute cavity features and volumes (Connolly, 1981, 1983, 1985). A summary of the number of cavities and cavity volumes computed for the eight models is given in Table IV. We note a general increase in total cavity volume and a decrease in the number of cavities for those models involving alternate conformations. The largest changes involve the respositioning of Phe-138. The cavity features for the protein with the three side chains in both the native and alternate conformations are shown in Figure 5. Changes are also found for residues bordering, and hence defining, these cavities (Table V). In particular, residues Trp-7, Leu-76, Leu-137, Ala-134, and His-82 lose their internal cavity surface while Leu-86, Ile-111, and the hydrophobic carbon side chain

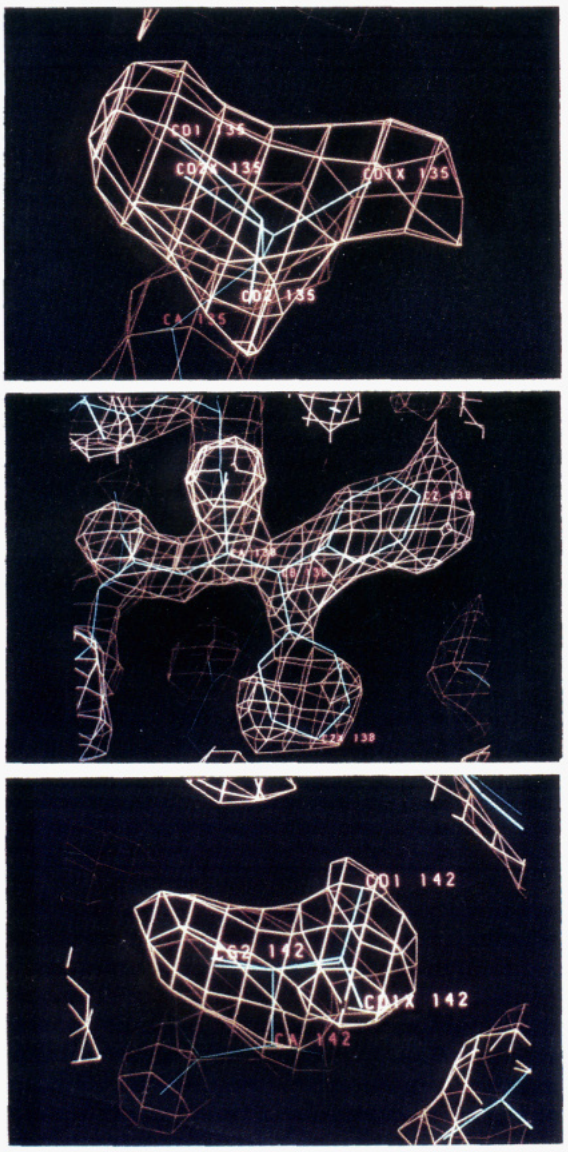

FIGURE 4: Electron density ($2F_o - F_c$ synthesis) and the modeled native and alternate conformations for three internal side chains of the protein. Atom names in the nonnative (alternate) conformation are terminated with an X. See Table II for dihedral angles and refined occupancies. (a, upper) Leucine-135, (b, middle) phenylalanine-138, and (c, lower) isoleucine-142. Pictures were generated with the program FRODO (Jones, 1982).

of Arg-139 border newly created cavity surfaces. The alternate conformation of residue Phe-138 occupies and substantially eliminates cavity 3 and allows the fusion of cavities 1 and 2 into a single large cavity.

Protein–Ligand Interactions. As a result of the high-pressure N_2 gas, two extra electron density features are observed bound in two of the native cavity features—cavities 1 and 3 (Figure 6). These electron density features were, for lack of unambiguous evidence, both modeled as water molecules. The refined isotropic Debye–Waller factor for the ligand in cavity 1 is 23 Å^2 while the Debye–Waller factor for the ligand in cavity 3 is 20 Å^2. Close contacts between the refined centers of the ligand and surrounding protein atoms are given in Table VI. We find only van der Waals contacts in the case of the ligand bound in cavity 1. The contacts are similar for both the native and alternate protein side-chain conformations. However, for the ligand bound in cavity 3, we observe contacts with the main-chain N of Gly-80 (2.65 Å) and the ND2 of-His 82 (2.45 Å). These contact distances are suggestive of hydrogen-bonding interactions. The alternate conformations of protein side chains do not cause steric overlap with the ligand. From these contact distances we make the assumption that the ligand in cavity 1 (proximal cavity) is a N_2 molecule while the ligand in cavity 3 is a single water molecule. High-resolution X-ray structures (Phillips, 1980; Kuriyan et al., 1986), as well as neutron structures (Hanson & Schoenborn, 1981), obtained at atmospheric pressure show no evidence for water bound in any of the internal cavities in myoglobin, except that bonded to the iron atom in the distal cavity of the met form of the protein.

Table VI

residue	atom	distance	residue	atom
		Ligand I[a]		
Wat	O	4.61	His-93	CB
Wat	O	4.07	His-93	CG
Wat	O	4.28	His-93	ND1
Wat	O	4.09	His-93	CD2
Wat	O	4.35	His-93	CE1
Wat	O	4.18	His-93	NE2
Wat	O	4.84	Leu-104	CG
Wat	O	3.96	Leu-104	CD1
Wat	O	4.30	Leu-104	CD2
Wat	O	4.23	Phe-138	CE1
Wat	O	4.45	Ile-142	CG1
Wat	O	3.85	Ile-142	CD1
Wat	O	4.71	Ile-142	CG1X
Wat	O	4.59	Ile-142	CD1X
Wat	O	4.41	Hem-154	CDA
Wat	O	4.48	Hem-154	ND
Wat	O	4.589	Hem-154	CD1
Wat	O	4.36	Hem-154	CD2
Wat	O	4.05	Hem-154	CD3
Wat	O	4.14	Hem-154	CD4
Wat	O	4.340	Hem-154	CD3A
		Ligand II[b]		
Wat	O	3.21	Ile-75	O
Wat	O	2.65	Gly-80	N
Wat	O	3.20	Gly-80	CA
Wat	O	2.45	His-82	ND1
Wat	O	3.23	His-82	CE1
Wat	O	3.07	Phe-138	CE2X

[a] Ligand I is believed to be a molecule of N_2, although it is modeled as a water molecule and is bound in cavity 1 (proximal). [b] Ligand II is believed to be a water molecule and is bound in cavity 3.

Debye–Waller Factors. While structural perturbations are observed in the vicinity of the two bound ligands, the isotropic Debye–Waller factor profiles of the 1- and 145-atm structures are substantially the same (Figure 7). The average difference between the two structures is less than 1.0 Å^2. This is within the 15–20% expected error for the determination of a Debye–Waller factor at 2.0-Å resolution (Petsko & Ringe, 1984; Kuriyan, 1986). However, a small, new feature is observed in the Debye–Waller factor profile in the region of residues 61–69 in the the refined 145-atm structure. This region includes residues of the distal pocket and is opposite the heme group from the ligand binding sites.

DISCUSSION

The discussion will be divided into two parts. Part 1 will examine the ligands that are bound at high N_2 gas pressure. Part 2 will discuss the effects of high gas pressure on protein conformational rearrangement and protein dynamics.

Two new electron density features observed in preexisting packing defects or cavities are interpreted as bound ligands. Solely from contact distances between the ligand and surrounding atoms of the protein, the ligand bound in the proximal cavity is believed to be a nitrogen molecule and the ligand bound in cavity 3 is believed to be a water molecule. While

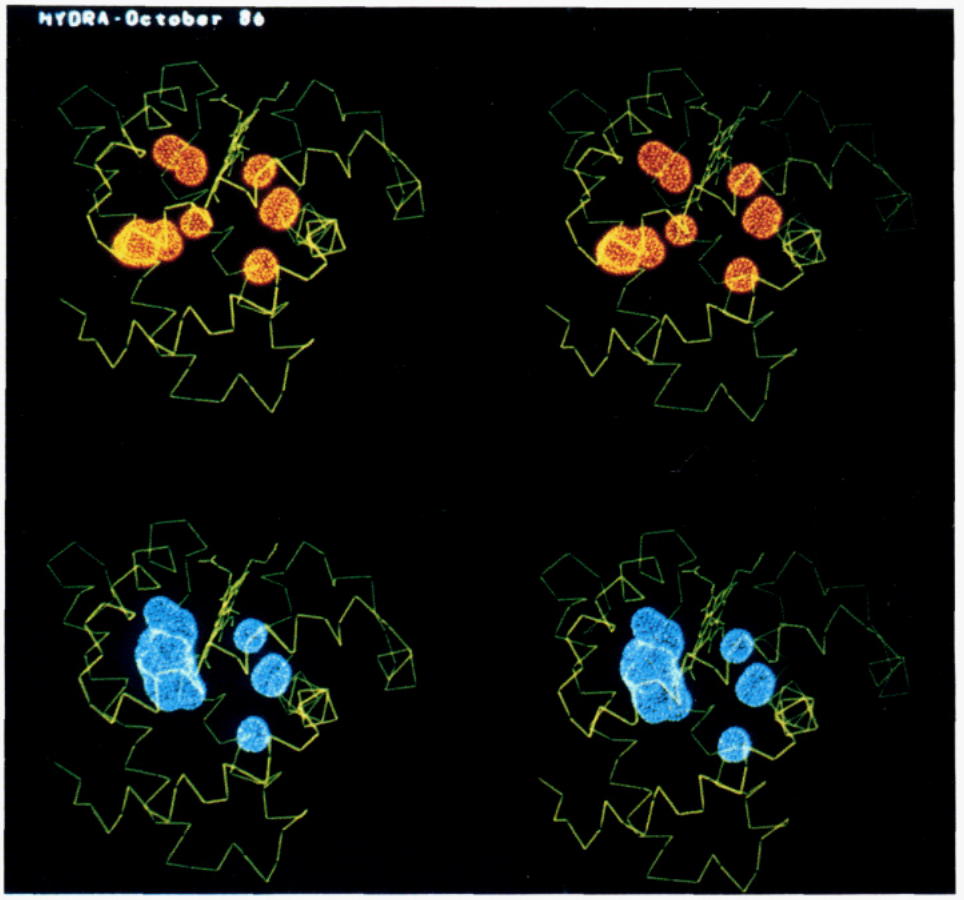

FIGURE 5: Cavity features computed with the Connolly algorithms using a 1.4-Å probe. (a, upper) Protein cavities with residues Leu-135, Phe-138, and Ile-142 in the single native conformation observed at 1 atm. (b, lower) Protein cavities with residues Leu-135, Phe-138, and Ile-142 in the alternate conformation observed at the elevated N_2 pressure. Ligands are removed for the computation of the cavity features. Pictures are generated with the program HYDRA written by Rod Hubbard (University of York, York, U.K.).

we stress that this X-ray experiment cannot distinguish between a N_2 and a water ligand, our interpretations are supported by the fact that the proximal cavity has the highest affinity for other inert gaseous ligands, e.g., Xe, and by molecular mechanics calculations that suggest that cavity 3 is the most likely internal site for water coordination (Tilton et al., 1986). The calculation indicates two hydrogen bonds between the bound water and the protein residues bordering the cavity—one with His-82 ND2 and a second with Leu-135 O. Similiarly, the X-ray structure indicates two hydrogen bonds—one with His-82 ND2 and a second with Gly-80 N. Leu-135 and Gly-80 are on opposite sides of cavity 3. The difference between the predicted structure and the experimental result is attributed to environmental effects of alternate conformations of Phe-138 and Leu-135 and to intrinsic bias in energy minimization toward the initial placement and orientation of the water molecule. While the evidence is circumstantial, we proceed under the assumption that we have correctly identified the two bound ligands.

If indeed a N_2 molecule is bound in the proximal cavity, it is almost certainly due to the elevated N_2 gas pressure. However, we do not, as yet, have an explanation for the presence of a water molecule in cavity 3. Two possibilities are that (1) direct pressure effects or (2) subtle perturbations in protein–water energetics by the high concentration of gas molecules have shifted the binding equilibrium to favor water binding. We do not believe that 145 atm of direct pressure is sufficient to produce such perturbations and speculate that interactions between the gas and solvent are responsible for changes in water activity and water binding affinity. Structural studies of myoglobin under high pressures of other gases and direct hydrostatic pressure might help to distinguish between the two possibilities.

While it is clear that ligand binding occurs in the vicinity of the alternate conformations of the protein, it is difficult to determine if ligand binding stablizes the alternate side-chain conformations, if fluctuations of protein side chains produce favorable ligand binding sites or if the two processes are completely independent. The difficulty arises, in part, from the unknown occupancies of the ligands and protein side-chain conformations and from the absence of information on the coordinated movements of the involved atoms. A striking observation, however, is that the alternate conformations do not conflict sterically with either the protein or the ligands. Instead it appears that the alternate conformations are able to be accommodated within the native conformation of the protein, reshuffling the available free volume and, in effect, repacking the protein interior. Even with the involvement of only three residues, dramatic changes in the atom-sized cavities are found. This fluctuation of interior cavity volume has been observed in recent molecular dynamics simulations of a myoglobin–xenon complex (Tilton et al., 1988) and supports the original hypothesis of Lumry and Rosenberg of "mobile defects" within proteins (Lumry & Rosenberg, 1975). If these alternate conformations can be accommodated within the native protein structure, why are they not observed in structures of the protein at 1 atm? The answer may involve general principles of protein packing and the interior free space within the protein.

While members of the globin family have a number of

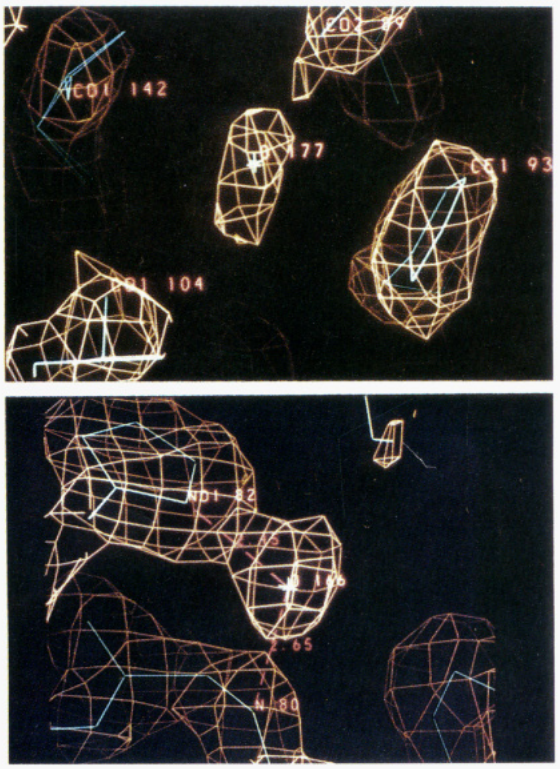

FIGURE 6: Electron density ($2F_o - F_c$ synthesis) and modeled positions for two bound ligands. The ligands are modeled as water molecules with full occupancy in the X-ray refinement. (a, upper) Ligand bound in cavity 1 (proximal) and believed to be a nitrogen molecule. (b, lower) Ligand bound in cavity 3 and believed to be a water molecule. Pictures were generated with the program FRODO (Jones, 1982). Close contacts are observed between the water molecule and the protein atoms N of Gly-80 (2.65 Å) and ND1 of His-82 (2.45 Å).

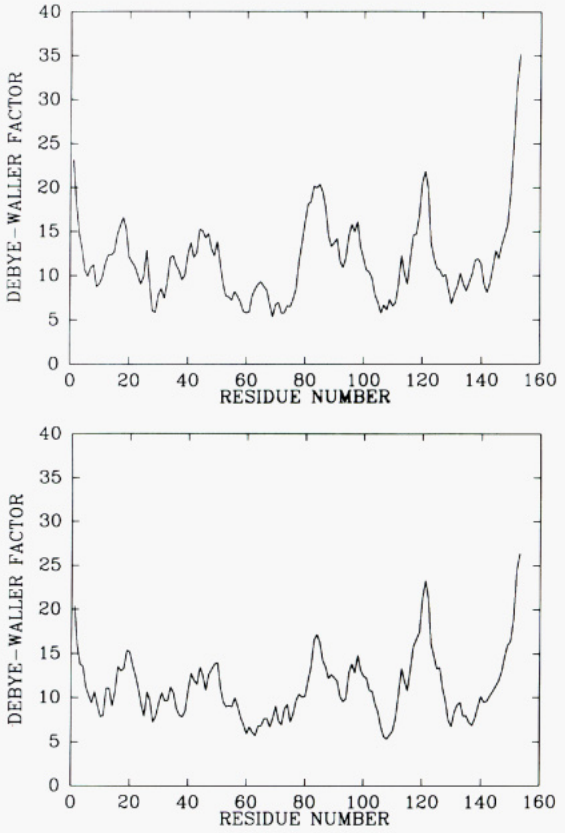

FIGURE 7: Average Debye–Waller factor (Å^2) versus residue number for metmyoglobin under 145 atm of N_2 pressure (top) and metmyoglobin under atmospheric pressure (bottom). Only main-chain atoms (C, CA, N) are included in the averaging. The profiles are qualitatively similar although a new small peak appears in the vicinity of residues 62–70 in the high gas pressure structure.

atom-sized cavities, proteins in general are remarkably well-packed molecules, lending validity to the axiom that "nature abhors a vacuum". The reason for such tight packing may involve the positive free energy associated with the creation of atom-sized cavities (Postma et al., 1981). Thus, proteins tend to minimize, if possible, the larger internal free volumes in order to increase stability. In virtually all instances, the observed alternate conformations of the high-pressure structure produce a significant increase in atom-sized free volume space. Because the overall protein volume remains unchanged, this increased cavity volume must arise from the reorganization of smaller packing-defect features (<10–15 Å^3) into larger atom-sized features. It is hypothesized that the process of coalescing these smaller packing defects into larger cavities would increase free energy and hence lead to a less favorable conformation. Therefore, native side-chain conformations, which minimize the large cavity volumes, would be the preferred state. It is only the additional perturbation of elevated gas pressure, possibly due to additional enthalpically favorable interactions between the protein and the two additional bound ligands, that stabilizes the alternate conformations and shifts the equilibrium.

The physical mechanisms that lead to side-chain fluctuations and internal rearrangements are still poorly understood. Spontaneous fluctuations and concerted motions of several protein residues as well as rearrangements of the protein–water structure are clearly involved. While not necessarily essential, the additional space provided by these atom-sized packing defects would promote atomic mobility and rearrangement. NMR data indeed suggest that Phe-138 in sperm whale myoglobin, a buried residue that borders cavities 2 and 3, is particularly mobile (Dalvit & Wright, 1987). Such protein fluctuations and transient nonnative states are crucial for ligand exchange and the phenomenon of hydrogen exchange (Woodward & Hilton, 1979; Englander & Kallenbach, 1984). The alternate side-chain conformations observed in this study, and the new cavity features that they define, provide structural evidence for necessary dynamic fluctuations within the protein interior.

X-ray crystallography can provide dynamic information in two ways. The first is implicit in the direct observation of alternate conformations and partial occupancies and the second is explicit in the high-frequency motions inherent in refined Debye–Waller factors. The effects of high gas pressure clearly indicate alternate conformations and internal rearrangements for interior protein residues. Moreover, the Debye–Waller profile indicates a small new peak in the region of residues 61–69 and a slight increase for residues 82–93 in the high-pressure structure as compared with the 1-atm structure. While errors of 10–15% of the Debye–Waller factor are expected and while there are difficulties in comparing Debye–Waller factors from structures solved with data from different crystals, the high-pressure structure appears to have greater mobility of the distal pocket and the EF corner. This is in contrast to the observation of lower Debye–Waller factors observed for myoglobin under a pressure of 7 atm of xenon gas where xenon atoms occupy four of the cavities and no alternate side-chain conformations are observed (Tilton et al., 1984). An explanation for the observed differences in the Debye–Waller factors is not available at this time.

Pressure represents a fundamental thermodynamic variable that can be used to study activation volumes of specific protein states and fluctuations (Wagner, 1980; Morishima et al., 1980;

Morild, 1981; Eden et al., 1982; Cooper, 1984; Kundrot & Richards, 1987). However, it is important to distinguish between the two types of pressure application. Hydrostatic and gas pressure may not produce the same perturbations. This is in large part due to the fact that gas molecules that enter the liquid and protein provide an additional perturbation. As an extreme case, gas molecules can, under moderate pressures (10–100's of atmospheres), promote clathrate formation and thus completely alter the properties of liquid water. This is believed to be important in the biological effects of anesthesia and narcosis that are produced with gases under moderate pressures (Halsey, 1980).

In addition, protein molecules with different folded topologies, packing densities, and amino acid compositions may respond differently to the application of pressure. This is suggested both by compressibility measurements (Gekko & Noguchi, 1979) and in a recent X-ray crystallographic study of HEW lysozyme under a hydrostatic pressure of 1000 atm (Kundrot & Richards, 1987). This crystallographic study revealed a small anisotropic compressibility of the protein but with no evidence of alternate side-chain conformations or protein repacking. This result contrasts with our finding on myoglobin at the more modest pressure of 145 atm of N_2, where no global compressibility is observed but substantial repacking of the protein interior is indicated. However, it is still premature to draw conclusions regarding the effects of pressure on protein molecules until a comprehensive study involving a wide range of hydrostatic and gas pressures is undertaken.

The physical limitations of future crystallographic studies of proteins under high pressure include the loss of crystallinity and pressure denaturation as well as experimental detection of the diffraction pattern. Different protein crystals will have different physical stabilities at high pressures. Adding higher concentrations of precipitant (Kundrot & Richards, 1986) or growing crystals at the elevated pressures may be useful techniques to extend the range of pressure. Data collection techniques using X-rays and a beryllium shroud are limited to pressures of 1500–2000 atm that approach the tensile strength of the beryllium metal. However, data collection with pressures to 4000 atm is possible with neutron crystallography using aluminum shrouds.

Conclusion

A 2.0-Å resolution structure of sperm whale metmyoglobin under a N_2 gas pressure of 145 atm reveals two bound ligands is preexisting cavities of the protein. One ligand, believed to be a single N_2 molecule, is bound in the proximal cavity (cavity 1). The second ligand forms two hydrogen bonds with His-82 and Gly-80 and is believed to be a water molecule. Three internal residues bordering the ligand binding sites (Leu-135, Phe-138, and Ile-142) are found to adopt alternate conformations at the elevated pressure. The alternate conformations do not produce steric overlap with the native protein or the bound ligands. These alternate conformations do, however, produce a rearrangement of the internal packing of the protein and result in a significant increase in the volume of atom-sized cavities.

Acknowledgments

We acknowledge helpful discussions with Rob Campbell, Fred Richards, Benno Schoenborn, John Kuriyan, Ken Dill, Steve Burley, and U. C. Singh and the programs PROLSQ, FRODO, and HYDRA. The high-pressure cell was designed in collaboration with Dave Costa and Dick Coco (Charles Stark Draper Laboratory) and fabricated in the beryllium facility at CSDL.

Registry No. Leu, 61-90-5; Phe, 63-91-2; Ile, 73-32-5; N_2, 7727-37-9; H_2O, 7732-18-5.

References

Brown, F. F., & Halsey, M. J. (1980) *Prog. Anesthesiol. 2*, 385–388.
Catchpool, J. F. (1968) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 27*, 884–887.
Conn, H. L., Jr. (1961) *J. Appl. Physiol. 16*, 1065–1070.
Connolly, M. L. (1981) Ph.D. Thesis, University of California, Berkeley.
Connolly, M. L. (1983) *J. Appl. Crystallogr. 16*, 548–558.
Connolly, M. L. (1985) *J. Am. Chem. Soc. 107*, 1118–1124.
Cooper, A. (1984) *Prog. Biophys. Mol. Biol. 44*, 181–214.
Dalvit, C., & Wright, P. E. (1987) *J. Mol. Biol. 194*, 313–327.
Eden, D., Mathew, J. B., Rosa, J. J., & Richards, F. M. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 815–819.
Englander, S. W., & Kallenbach, N. R. (1984) *Q. Rev. Biophys. 16*, 521–655.
Ewing, G., & Maestas, S. (1970) *J. Phys. Chem. 74*, 2341–2344.
Featherstone, R., Muehlbaecher, C., DeBon, F., & Forsaith, J. (1961) *Anesthesiology 22*, 977.
Frauenfelder, H., Petsko, G. A., & Tsernoglou, D. (1979) *Nature (London) 280*, 558–563.
Gekko, K., & Noguchi, (1979) *J. Phys. Chem. 83*, 2706–2713.
Halsey, M. J. (1980) *Gen. Anaesth. (4th Ed.)*, 45–65.
Hanson, J. C., & Schoenborn, B. P. (1981) *J. Mol. Biol. 153*, 117–146.
JCPDS (1951) *Powder Diffraction File*, Vol. 1, p 1291, JCPDS, Swarthmore, PA.
Jones, A. (1982) in *Computational Crystallography* (Sayre, D., Ed.) pp 3303–3317, Clarendon, Oxford.
Kendrew, J. C., & Parrish, R. G. (1956) *Proc. R. Soc. London, A 138*, 305–324.
Keyes, M., & Lumry, R. (1968) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 27*, 895–897.
Konnert, J. H., & Hendrickson, W. A. (1980) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A36*, 344–350.
Kundrot, C. E., & Richards, F. M. (1986) *J. Appl. Crystallogr. 19*, 208–213.
Kundrot, C. E., & Richards, F. M. (1987) *J. Mol. Biol. 193*, 157–170.
Kuriyan, J. (1986) Ph.D. Thesis, Massachusetts Institute of Technology.
Kuriyan, J., Wily, S., Karplus, M., & Petsko, G. A. (1986) *J. Mol. Biol. 192*, 133–154.
Lumry, R., & Rosenberg, A. (1975) *Colloq. Int. C.N.R.S. L'Eau Syst. Biol. 246*, 55–63.
Morild, E. (1981) *Adv. Protein Chem. 34*, 93–166.
Morishima, I., Ogawa, S., & Yamada, H. (1980) *Biochemistry 19*, 1569–1575.
Mortimer, R. G., & Bauer, N. (1960) *J. Phys. Chem. 64*, 387–390.
North, A. C., Phillips, D. C., & Mathews, F. S. (1968) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A24*, 351–359.
Petsko, G. A., & Ringe, D. (1984) *Annu. Rev. Biophys. Bioeng. 13*, 331–371.
Phillips, S. E. V. (1980) *J. Mol. Biol. 142*, 531–554.
Postma, J. P. M., Berendsen, H. J., & Haak, J. R. (1981) *Faraday Symp. Chem. Soc. 17*, 55–67.
Richards, F. M. (1977) *Annu. Rev. Biophys. Bioeng. 6*, 151–176.

Schoenborn, B. P. (1965) *Nature (London) 208*, 760–762.
Schoenborn, B. P. (1967) *Nature (London) 214*, 1120–1122.
Schoenborn, B. P., Watson, H. C., & Kendrew, J. C. (1965) *Nature (London) 207*, 28–30.
Smith, J. L., Hendrickson, W. A., Honzatko, R. B., & Sheriff, S. (1986) *Biochemistry 25*, 5018–5027.
Ten Eyck, L. F. (1977) *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. A33*, 486–492.
Tilton, R. F. (1988) *J. Appl. Crystallogr. 21*, 4–9.
Tilton, R. F., Jr., Kuntz, I. D., Jr., & Petsko, G. A. (1984) *Biochemistry 23*, 2849–2857.
Tilton, R. F., Jr., Singh, U. C., Weiner, S. J., Connolly, M. L., Kuntz, I. D., Jr., Kollman, P. A., Max, N., & Case, D. A. (1986) *J. Mol. Biol. 192*, 443–456.
Tilton, R. F., Singh, U. C., Kuntz, I. D., & Kollman, P. A. (1988) *J. Mol. Biol. 199*, 195–211.
Wagner, G. (1980) *FEBS Lett. 112*, 280–284.
Wishnia, A. (1969) *Biochemistry 8*, 5064–5070.
Woodward, C. K., & Hilton, B. D. (1979) *Annu. Rev. Biophys. Bioeng. 8*, 99–127.
Zipp, A., & Kauzmann, W. (1973) *Biochemistry 12*, 4217–4228.

Structural Comparison of Two Serine Proteinase–Protein Inhibitor Complexes: Eglin-C–Subtilisin Carlsberg and CI-2–Subtilisin Novo†

Catherine A. McPhalen‡ and Michael N. G. James*

Medical Research Council of Canada Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7

Received November 10, 1987; Revised Manuscript Received March 24, 1988

ABSTRACT: The crystal structures of the molecular complexes between two serine proteinases and two of their protein inhibitors have been determined: subtilisin Carlsberg with the recombinant form of eglin-c from the leech *Hirudo medicinalis* and subtilisin Novo with chymotrypsin inhibitor 2 from barley seeds. The structures have been fully refined by restrained-parameter least-squares methods to crystallographic R factors ($\sum||F_o| - |F_c||/\sum|F_o|$) of 0.136 at 1.8-Å resolution and 0.154 at 2.1-Å resolution, respectively. The 274 equivalent α-carbon atoms of the enzymes superpose with an rms deviation of 0.53 Å. Sequence changes between the enzymes result in localized structural adjustments. Functional groups in the active sites superpose with an rms deviation of 0.19 Å for 161 equivalent atoms; this close similarity in the conformation of active-site residues provides no obvious reason for known differences in catalytic activity between Carlsberg and Novo. Conformational changes in the active-site region indicate a small induced fit of enzyme and inhibitor. Some conformational differences are observed between equivalent active-site residues of subtilisin Carlsberg and α-chymotrypsin. Despite differences in tertiary architecture, most enzyme–substrate (inhibitor) interactions are maintained. Subtilisin Carlsberg has a rare *cis*-peptide bond preceding Thr211 (Gly211 in Novo). Both enzymes contain tightly bound Ca^{2+} ions. Site 1 is heptacoordinate with the oxygen atoms at the vertices of a pentagonal bipyramid. Site 2 in Carlsberg is probably occupied by a K^+ ion in Novo. Conserved water molecules appear to play important structural roles in the enzyme interior, in the inhibitor β-sheet, and at the enzyme–inhibitor interface. The 62 equivalent α-carbon atoms of the inhibitors superpose with an rms deviation of 1.68 Å. Sequence changes result in somewhat different packing of the α-helix, β-sheet, and reactive-site loop relative to each other. Hydrogen bonds and electrostatic interactions supporting the conformation of the reactive-site loop are conserved. The 24 main-chain plus C^β atoms of P_4 to P_1' overlap with an rms deviation of 0.19 Å. Features contributing to the inhibitory nature of eglin-c and CI-2 are discussed.

Subtilisin Carlsberg was the first bacterial serine proteinase to be discovered (Linderstrøm-Lang & Ottesen, 1947). It consists of a single polypeptide chain of 274 amino acids, M_r 27 292 (Smith et al., 1968) with extensive sequence identity with subtilisn Novo (Figure 1a). Recently, solution of the crystal structure of proteinase K (Pähler et al., 1984) demonstrated that this fungal enzyme also belongs to the subtilisin family. Subtilisins Novo and Carlsberg have broad substrate specificities with preferences for large aromatic or aliphatic amino acids in the P_1 position of the substrate.[1] The catalytic triad, histidine, aspartic acid, and serine, has a similar but not identical spatial arrangement in the subtilisins and the chymotrypsin family of serine proteinases (Kraut et al., 1972; Wright, 1972). These two families of enzymes form one of the best-known examples of convergent evolution, with completely different tertiary structures providing the framework for a similar disposition of active-site residues.

Eglin-c is a protein inhibitor of serine proteinases isolated from the leech *Hirudo medicinalis* (70 amino acids, M_r 8092;

† This research was generously supported by a group grant from the Medical Research Council of Canada. C.A.M. gratefully acknowledges the Alberta Heritage Foundation for Medical Research for financial support. The work described in this paper forms part of the doctoral thesis of C.A.M.

‡ Present address: Department of Structural Biology, Biozentrum, University of Basel, CH-4056 Basel, Switzerland.

[1] Nomenclature of Schechter and Berger (1967). Amino acid residues of substrates or inhibitors are numbered P_1, P_2, P_3, etc. toward the N-terminus and P_1', P_2', etc. toward the C-terminus from the reactive-site bond (the peptide bond in the substrate or inhibitor cleaved by the enzyme). The complementary subsites of the binding region on the enzyme are numbered S_1, S_2, etc. and S_1', S_2', etc.

0006-2960/88/0427-6582$01.50/0