Solution Conformation of Cytochrome c-551 from *Pseudomonas stutzeri* ZoBell Determined by NMR

Mengli Cai and Russell Timkovich
Department of Chemistry, University of Alabama, Tuscaloosa, Alabama 35487-0336 USA

ABSTRACT ^{1}H NMR spectroscopy and solution structure computations have been used to examine ferrocytochrome c-551 from *Pseudomonas stutzeri* ZoBell (ATCC 14405). Resonance assignments are proposed for all main-chain and most side-chain protons. Stereospecific assignments were also made for some of the β-methylene protons and valine methyl protons. Distance constraints were determined based upon nuclear Overhauser enhancements between pairs of protons. Dihedral angle constraints were determined from estimates of scalar coupling constants and intra-residue NOEs. Twenty structures were calculated by distance geometry and refined by energy minimization and simulated annealing on the basis of 1012 interproton distance and 74 torsion angle constraints. Both the main-chain and side-chain atoms are well defined except for two terminal residues, and some side-chain atoms located on the molecular surface. The average root mean squared deviation in the position for equivalent atoms between the 20 individual structures and the mean structure obtained by averaging their coordinates is 0.56 ± 0.10 A for the main-chain atoms, and 0.95 ± 0.09 A for all nonhydrogen atoms of residue 3 to 80 plus the heme group. The average structure was compared with an analogous protein, cytochrome c-551 from *Pseudomonas stutzeri*. The main-chain folding patterns are very consistent, but there are some differences, some of which can be attributed to the loss of normally conserved aromatic residues in the ZoBell c-551.

INTRODUCTION

Cyt c-551 is an 82 residue prokaryotic electron transport protein functionally equivalent to mitochondrial cyt c, yet with 20–25% fewer residues it affords an opportunity to determine the minimum structure required for successful function. Evolution has created an extensive family of c-551s among different species and even genera of prokaryotes. Within the family, there can be amino acid substitutions ranging from a few up to 45% of the primary structure. The respiratory chain is essential for cell survival, and all versions are successful in fulfilling the basic requirements of electron transport. Whether there are differences in the efficiency of function remains an unanswered question. The c-551 family presents an opportunity to investigate the impact of different residues on the final three-dimensional structure of the protein, by comparing the structure of individual members. NMR presents a method for structure determination of this size protein with the advantage of not requiring the growth of single crystals.

We have previously investigated the solution conformations of cyt c-551 from *Pseudomonas stutzeri* and *P. aeruginosa* (Cai et al., 1992; Chau et al., 1990; Timkovich and Cai, 1993). As part of a program to establish a comprehensive structural understanding of this family, the solution conformation of *Pseudomonas stutzeri* ZoBell (ATCC14405) has now been determined. This organism was isolated from marine sources and was formerly known as *Pseudomonas perfectomarius*, but reassessment of its phylogeny indicated a reclassification as a substrain of *P. stutzeri* (Dohler et al., 1987). The genetic relations notwithstanding, the ZoBell strain has a c-551 whose sequence has 13% different residues compared with c-551 of *P. stutzeri* (ATCC 17588) and 33% compared with c-551 from *P. aeruginosa* (ATCC 19428) (Jungst et al., 1991). To minimize confusion, we will refer to it as the ZoBell strain and the cytochrome as PZ c-551, to differentiate it from the *P. aeruginosa* protein (PA c-551) and the *P. stutzeri* protein (PS c-551).

Known members of the c-551 family generally have a small number of aromatic residues, but these residues are usually conserved. Excluding histidines, the usual set of aromatic residues includes Trp-56 and Trp-77, and then combinations of Phe or Tyr at positions 7, 27, and 34. All residues in this set are buried, although there is some possible edge contact with solvent. Our previous results established that in solution the three benzoid rings are rapidly flipping at room temperature or above (Chau et al., 1990). In terms of tertiary structure, the aromatics are found in two distinct hydrophobic clusters. The first is comprised of Trp-56 and Tyr/Phe-34 at the edge of the heme crevice in contact with the buried propionate 17 and the partially buried propionate 13. The second cluster is comprised of Trp-77 and Tyr/Phe-7 and -27 in contact with the heme crevice at the buried thioether bridge bonds to Cys-12 and Cys-14. PZ c-551 is atypical in its sequence with Leu at 27 and Asn at 34. One goal of the present study was to investigate the consequences of these substitutions in otherwise highly conserved locales.

Received for publication 21 March 1994 and in final form 26 May 1994.
Address reprint requests to Dr. Russell Timkovich, Department of Chemistry, University of Alabama, Tuscaloosa, Alabama 35487-0336. Tel.: 205-348-8439; Fax: 205-348-9104; E-mail rtimkovi@ua1vm.ua.edu.

Abbreviations used: NOE, nuclear Overhauser Enhancement; PA, *Pseudomonas aeruginosa*; PS, *Pseudomonas stutzeri* ATCC 17588; PZ, *Pseudomonas stutzeri* ZoBell, formerly *P. perfectomarius*, ATCC 14405; rmsd, root mean standard deviation; SA, simulated annealing.

0006-3495/94/09/1207/09 $2.00

MATERIALS AND METHODS

PZ c-551 was purified similar to the method described in Cai et al. (1992), which in turn had been adapted from an earlier procedure (Ambler and Wynn, 1973). Cells was grown in a 100-liter fermentor in the ACN medium (Korner et al., 1987) at pH 6.8 and 30°C for 20 h. The culture was grown anaerobically for the first 8 h, and then semi-aerobically with slowly bubbling of air through the culture medium. Cell walls were broken by sonication, followed by ammonium sulfate fractionation between 45–85% saturation. The crude protein solution was dialyzed against 50 mM ammonium acetate, pH 3.8, and loaded onto a carboxymethyl cellulose column. The very first fraction at pH 4.1 has a maximum absorbance at 554 nm in its reduced form, and could be a cytochrome detected earlier in the ZoBell strain (Zumft et al., 1988). Similar to the situation with PS c-551, two isozymes of PZ c-551 were isolated, which were eluted from the carboxymethyl cellulose at pH 4.2 (major) and pH 4.4 (minor), respectively. Detailed spectroscopic studies were performed only on the major isozyme. The loss of two aromatic residues in PZ c-551 was reflected in the uv-visible spectra of the cytochrome by decreased intensity in the 280-nm range. For the c-type cytochromes, a conventional criterion of purity has been the purity ratio, specified as the ratio of the absorbance of the reduced ferrocytochrome at 551 nm minus a baseline value of the absorbance of the ferrocytochrome at 570 nm to the absorbance of the oxidized ferricytochrome at 280 nm. For typical c-551s, this is in the range of 1.2–1.5, whereas for PZ c-551 it was 1.7.

Sample preparation and spectroscopic details for DQF-COSY, HOHAHA, and NOESY spectra have been described (Chau et al., 1990). Definitions of the secondary structure elements in terms of phi and psi angles have been taken from a standard source (Crawford et al., 1973).

Assignments

HOHAHA spectra in H_2O with spin lock times of 50 ms were obtained and used to categorize residue spin system types. Certain unique landmark residues, including the heme itself, ligand and thioether bridge residues, and the small number of aromatic residues, were readily assigned based upon distinctive NOEs to heme or aromatic protons. Sequential NOEs were then used to link and identify the remaining residues. Side-chain assignments were extended by examining HOHAHA spectra with variable spin-lock times and DQF-COSY spectra. Examples of this process have been given in detail (Chau et al., 1990). Although there is homology between PZ c-551 and previously studied cytochromes, the differences are sufficient to require a de novo analysis of the PZ spectra. However, once this is completed, the observed similarities and differences with respect to other c-551s adds high confidence to the assignments. An NOE connectivity diagram indicating the critical sequential NOEs is presented in Fig. 1, and assignments are given in Table 1.

NMR structure constraints

Distance constraints based on the NOESY crosspeak intensities were determined on spectra recorded with mixing time equal to 100 ms and were classified into groups of different distance ranges according the criteria described previously (Cai et al., 1992). A total of 1012 interproton distance constraints were determined, which included 238 intraresidue, 301 sequential (for residues i and j, $j - i = 1$), 172 short range ($j - i < 5$), 192 long range ($j - i > 5$), and 109 NOEs from the heme group to peptide protons. These distance constraints were obtained in two stages. In the first stage, only well resolved and unambiguously assigned NOEs were converted into distance constraints. A total of 723 distance constraints were identified, including 238 intraresidue, 396 interresidue, and 91 NOEs from the heme group to the peptide protons. These constraints along with 55 backbone dihedral angle constraints and 19 side-chain dihedral angle constraints, discussed in the following paragraph, were used to compute the preliminary structure for PZ c-551. Details of this are described in the structure computation subsection. With the preliminary structure available, additional NOEs could now be assigned with confidence in the second stage. NOEs

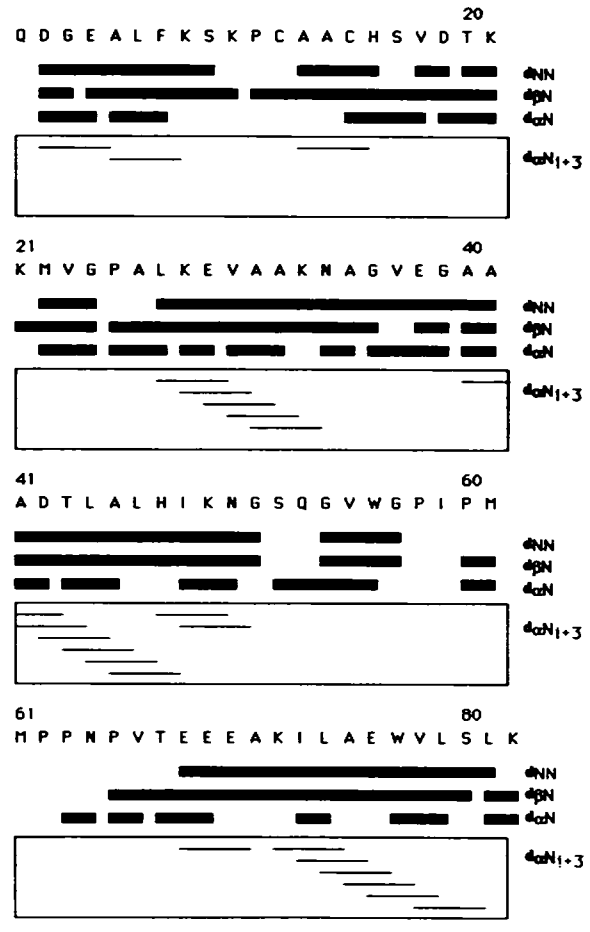

FIGURE 1 Summary of sequential connectivities for PZ c-551. A bar indicates a NOESY cross peak was observed between the two residues at the end of the bar, and the type of NOE is indicated by the designation on the right.

assigned from overlapping cross peaks, where intensity could be ambiguous because of the overlap, were assigned to the very weak class, and the upper boundary was set to 5 A.

Backbone dihedral angle constraints were estimated based upon scalar coupling as described previously (Driscoll et al., 1989). By estimating the width of the cross peak from a main-chain amide proton to a proton in the HOHAHA spectra, and by the peak-to-peak separation in high digital resolution DQF-COSY spectra, 15 coupling constants were identified that were larger than 8 Hz and 33 that were smaller than 5.5 Hz. Corresponding main-chain torsional angles were constrained to be in the range -80 to $-160°$ for the former and -40 to $-90°$ for the latter. All seven prolines were identified as being in the *trans* conformation and were constrained to an ω angle of 175–185°. No constraints were applied to glycine residues.

Stereospecific assignments were made for 13 β methylene protons sets on the basis of established criteria (Wagner et al., 1987), by examination of DQF-COSY, HOHAHA, and NOESY spectra. From these, eleven χ_1 were constrained to $-60 \pm 60°$, and two were constrained to $180 \pm 60°$. Four valine residues were identified with one belonging to the g+ rotamer (χ_1, $-60 \pm 60°$), and two to the t rotamer (χ_1, $180 \pm 60°$) conformation, according to the classification of Zuiderweg et al. (1985).

Structure computation

The steps in the structure computation for PZ c-551 were similar to those employed to determine the solution structure of PS c-551. An initial model was obtained by superimposing the sequence of PZ c-551 onto the global

TABLE 1 Proton resonance assignments for PZ c-551

Residue**	Chemical shift (ppm)†						
	NH	Cα	Cβ	Cγ	Cδ	Cε	Other
Q1							
D2	8.01	4.65	2.72				
G3	8.77	3.30					
E4	7.23	2.62	1.83	2.13			
A5	7.40	3.95	1.38				
L6	7.71	4.07	1.85	1.65	0.90		
			1.41				
F7	8.21	4.02	2.73				C2 & C6 6.99, C4 6.55
			2.63				C3 & C5 6.62
K8	7.52	3.83	1.79	1.70	2.08	2.95	
S9	7.73	4.53	3.97				
			4.07*				
K10	7.48	5.02	2.18	1.91	1.40	2.90	
			2.67*	2.00	1.55		
P11		4.82	2.65	2.03	3.86		
			2.82	2.25	4.10		
C12	8.05	4.73	2.62				
			2.25*				
A13	7.81	4.10	1.07				
A14	7.85	4.15	1.70				
C15	6.71	4.32	1.83				
			0.99*				
H16	6.92	4.03	1.14				C2 0.70, C5 0.63
							Nπ H 8.64
S17	8.11	4.42	3.68				
			3.78				
V18	8.48	3.15	1.83	0.65*			
				0.71			
D19	7.75	4.50	2.50				
			2.40				
T20	7.21	4.38	3.75	0.98			
K21	8.47	3.53	1.70	1.10	1.50	2.82	
			1.55*				
M22	7.24	4.33	1.65	1.75			
			1.35	2.00			
V23	6.98	3.91	1.90	1.60			
				1.17*			
G24	6.47	3.72					
		0.18					
P25		3.51	0.42	0.14	2.03		
			0.88	−0.07	2.85		
A26	8.23	3.80	0.67				
L27	7.87	3.33	0.88	0.50	−0.82		
			1.05*		−0.10		
K28	8.67	3.75	1.63	1.23	1.47	2.70	
						2.78	
E29	6.85	4.18	2.00	2.27			
V30	7.16	3.94	2.27	1.48			
				0.61*			
A31	7.92	3.85	1.58				
A32	7.74	4.08	1.53				
K33	7.90	4.21	2.10	2.02	1.51	2.85	
			2.27		1.63		
N34	7.37	4.79	2.83				NH_2 6.97, 8.40
			2.45*				
A35	7.48	4.20	1.48				
G36	8.69	4.00					
		3.90					
V37	7.50	4.00	2.25	0.95			
E38	8.79	4.15	2.01	2.30			
G39	8.95	4.00					
		4.15					
A40	7.64	4.12	1.60				
A41	8.82	3.83	1.40				
D42	7.50	4.32	2.65				
T43	7.66	3.62	4.10	0.87			
L44	8.76	3.90	1.46	2.13	1.19		
			2.07*		0.98		
A45	8.13	3.82	1.40				
L46	7.41	3.83	1.68	1.60	0.78		
H47	7.61	4.20	2.02				C2 8.70
			2.97*				C5 7.81

TABLE 1 *Continued*

Residue**	Chemical shift (ppm)†						
	NH	Cα	Cβ	Cγ	Cδ	Cε	Other
I48	8.30	1.38	1.55	0.75	1.32		
			0.65	1.42			
K49	7.28	3.72	1.61	1.28	1.45	3.05	
N50	8.09	4.49	2.66				NH_2 7.00, 7.10
G51	7.12	3.53					
		2.01					
S52	7.01	3.92	2.55				
Q53	7.54	4.20	1.85	1.99			NH_2 7.30, 6.61
				2.18			
G54	8.07	4.15					
		3.82					
V55	10.48	3.69	2.37	0.64			
				0.85*			
W56	10.57	4.55	3.58				C2 7.65, C4 7.95
			3.76				C5 7.23, C6 7.02
							C7 7.55, Nε 12.32
G57	7.97	4.55					
		4.20					
P58		4.75	2.26	2.00	3.93		
			2.26	2.17	3.65		
I59	7.42	4.66	2.20	1.50	1.12		
			1.35	1.81			
P60		4.83	1.45	1.87	3.87		
			1.95	2.10	4.12		
M61	8.50	3.65	−0.92	−3.75*			S-Me −2.92
			−2.75*	−0.55			
P62		4.15		1.72	3.61		
				1.90	2.88		
P63		3.33	2.00	1.55	3.20		
			1.48	1.72	3.16		
N64	7.01	5.09	1.91				NH_2 6.95, 3.11
			2.50*				
P65		4.83	1.90	2.45	3.80		
				2.10	3.85		
V66	8.02	5.06	2.45	0.55*			
				0.69			
T67	8.89	4.48	4.81	1.28			
E68	9.10	3.83	2.00	2.32			
E69	8.43	4.01	1.93	2.32			
			2.07				
E70	7.79	3.80	2.49	1.75			
A71	8.61	3.98	1.53				
K72	7.68	4.08	2.05	1.45	1.65	2.89	
			1.93				
I73	7.98	3.80	1.90	1.10	0.81		
			0.85	1.80			
L74	8.70	4.40	2.50*	1.93	1.19		
			1.81		1.32		
A75	8.72	4.02	1.73				
E76	8.38	4.00	2.19	2.57			
W77	8.44	4.37	3.78				C2 7.27, C4 7.61
			3.38				C5 6.15, C6 5.65
							C7 6.98, Nε 10.20
V78	9.04	2.69	2.48	0.85			
				1.42*			
L79	7.66	3.79	1.90	1.90	0.82		
			1.55		0.86		
S80	7.62	4.35	4.07				
			3.97				
L81	7.18	3.91	1.45*	1.20	0.09		
			0.85		−0.20		
K82	7.21	3.99	1.60	1.63	1.35	2.95	
			1.70				

Type of heme proton	IUB-IUPAC Designation	Chemical shift
Meso proton	5	9.75
	10	9.25
	15	9.23
	20	9.13

TABLE 1 *Continued*

Type of heme proton	IUB-IUPAC Designation	Chemical shift
Ring methyl	2^1	3.73
	7^1	3.70
	12^1	3.25
	18^1	3.40
Bridge methine	3^1	5.87
	8^1	6.17
Bridge methyl	3^2	1.86
	8^2	2.45
Propionate	13^{1a}	3.97
	13^{1b}	4.51
	13^{2a}	3.34
	13^{2b}	2.68
	17^{1a}	4.17
	17^{1b}	4.52
	17^{2a}	2.52
	17^{2b}	3.30

** Amino acids are listed with the one-letter code.

† Chemical shifts (ppm) are referenced to dioxin at 3.74 ppm, and the reported shifts were obtained from the samples recorded at pH 6.0 and 320 K. For some entries, only a single shift is given when there are two protons attached to the same carbon in the side chain. This means that we were unable to detect two discrete shifts under any condition. For side chains with successive sets of methylene protons (Glu, Gln, Arg, Lys, Met), the assignment to respective Cβ, Cγ, etc., was in some cases straightforward because of the resolution among the protons (e.g., M22 and M61). The more difficult cases where when one or more of the methylene protons sets were unresolved (e.g., E4) and successive sets were close in chemical shifts. The assignment to β, γ, and so forth was then made by examining COSY spectra and/or HOHAHA spectra taken with increasing spin-lock times, as described in detail for the assignment of lysine and proline side chains in Cai et al. (1992). The stereospecific assignments are denoted as follows. For the Cβ methylene protons, the asterisk indicates the Hβ2 proton; for Met 61 the asterisk at −3.75 indicates this has been assigned as Hγ2. For valine methyl groups, the asterisk indicates Cγ1. Methylene protons of residues 9, 10, 12, 15, 27, 34, 44, 47, 64, 74, and 81 are classified as t^2g^3 rotamers, and methylene protons of residues 21 and 61 are classified as g^2g^3 rotamers. Val 23, Val 30, Val 55, and Val 78 are classified as g^+ rotamers; Val 66 and Val 18 are classified as g^- rotamers.

folding of PS c-551. The graphics and molecular modeling software package SYBYL (Tripos Associates, St. Louis, MO) uses as many of the side-chain coordinates of PS c-551 as possible, whereas for atoms unique to PZ c-551, it adds them in standard relative geometry. This standard geometry does not account for inter-residue contacts and so, it does introduce steric conflicts. The initial model was refined by simulated annealing (Nilges et al., 1988) with an initial set of unambiguous distance and torsional angle constraints incorporated into the force field, as described for PS c-551 (Cai et al., 1992). With the preliminary model available, chemical shift ambiguities could be overcome and further constraints were then added.

On the basis of the full set of NMR constraints, solution structures were computed by distance geometry and simulated annealing using the software package X-PLOR (version 3.0). The protocol used in the structure computation process included three basic steps, i.e., substructure embedding and regularization, full structure SA-regularization, and SA refinement. In the first step, X-PLOR translates covalent geometry from X-PLOR parameter files into interatomic distance constraints, adds experimentally derived constraints, and creates a matrix of upper and lower bounds on the distances between all atoms of the protein; then a sub-set of atoms, including C, CαH, N, NH, Cβ, and Cγ, is embedded. The results were then minimized against an effective energy term, which represents all upper and lower bounds in the distance geometry (DG) matrix, whereas all the other energy terms were turned off. The correct enantiomer was selected on the basis of the lower rmsd to the preliminary model described earlier. A total of 24 structures were generated at this stage. These only contained a subset of atoms, and also required extensive regularization. The missing atoms were added to the molecule one residue a time by fitting a standard residue substructure onto the core atoms that had earlier been included (Nilges et al., 1988). These structures were then regularized by a multistage minimization protocol. During this refinement, nonbonded attractive potentials were never used. The nonbonded interactions were represented by a simple repulsion term, as described by Nilges et al. (1988), which replaced all the nonbonded terms of conventional molecular dynamics. This repulsion term equals zero when the distance between two atoms is larger or equal to the sum of standard van der Waal's radii as described in CHARMM. The regularization was started with only the experimental constraint terms and covalent bond energy, whereas all other potentials were turned off. After 100 steps of Powell energy minimization, the bond angle term was added to the force field and the structures were regularized with another 100 steps of Powell energy minimization. The improper torsion angle term was introduced slowly by increasing its weight from 0.005 to 1.000 during a series of restrained molecular dynamics refinement calculations. Finally, the structures were regularized with simulated annealing refinement with bond, angle, improper torsion, and the nonbonded repulsion term all included. Force constants, temperatures, and annealing times were as described previously (Cai et al., 1992) for regularization of embedded distance geometry structures.

All 24 structures generated had the same global folding. The rmsd to their mean structure was only 0.76 A for backbone atoms and 1.20 A for all heavy atoms. However, 2 of the 24 structures had higher NMR constraint energy and covalent structure energy compared with the others, and so were excluded. Lys-21 in two of the remaining 22 structures had a phi torsional angle of 70°, whereas in all others the value was 170°. These two were considered to represent a possible local energy minimum conformation and were also excluded.

A mean structure, designated ⟨SA⟩, was obtained by best-fitting the remaining 20 SA structures to each other and averaging their coordinates (Driscoll et al., 1989). Irregularities caused by the simple averaging process were corrected by 500 cycles of restrained energy minimization to yield the minimized mean structure, ⟨SA⟩r. Structure statistics for the best 20 SA structures and the minimized mean structure are given in Table 2. All 20 SA structures and the minimized mean structure display only small deviations from idealized covalent residue geometry and NMR constraints and have good nonbonded contacts as indicated by small values for the repulsive force and negative value for van der Waals energies. Coordinates for the minimized mean structure have been deposited with the Brookhaven Protein Data Bank (entry "1CCH").

RESULTS AND DISCUSSION

The overall fold of the protein is displayed in Fig. 2 by a ribbon diagram for ⟨SA⟩r. Superposition of the main-chain atoms for the 20 SA structures is shown in Fig. 3. The conformation of all main-chain atoms for the 20 SA structures

TABLE 2 Structure statistics for PZ c-551

	Average over the 20 SA*	Value for ⟨SA⟩r
Distance constraints, rmsd violations in A‡		
intraresidue	0.012 (±0.001)	0.009
sequential	0.012 (±0.004)	0.010
short range	0.011 (±0.002)	0.008
long range	0.010 (±0.002)	0.007
Angle constraints, rmsd violations in degrees		
	0.121 (±0.006)	0.118
Energies§ (kcal/mol)		
F_{NOE}	6.98 (±2.14)	3.81
F_{tor}	0.01 (±0.02)	0.00
F_{repel}	1.48 (±0.47)	0.85
F_{L-J}	−470.71 (±14.0)	−470.15
rmsd from idealized geometry		
bonds (A)	0.007 (±0.000)	0.007
angles (deg)	2.080 (±0.003)	2.076
improper (deg)	0.228 (±0.034)	0.186
rmsd displacements in A for	SA vs ⟨SA⟩	⟨SA⟩ vs. ⟨SA⟩r
whole protein		
main chain	0.69 (±0.12)	0.26
all heavy atoms	1.13 (±0.10)	0.53
residues 3–80 plus heme		
main chain	0.56 (±0.11)	0.18
all heavy atoms	0.95 (±0.11)	0.42

* SA represents the 20 SA individual structures refined by the simulated annealing method. ⟨SA⟩ is the mean structure obtained by simple averaging of the coordinates of 20 SA structures. ⟨SA⟩r is the minimized mean structure as described in the text.

‡ The *rmsd violation* was computed for a single SA structure; then the results for the 20 SA structures were averaged with the result listed in the column labeled "average"; then the rmsd of the individual 20 values from the 20 structure was computed and listed in parentheses. For ⟨SA⟩r there is only a single structure, so the actual *rmsd violation* is listed.

§ The force constants used for these calculations were 50 kcal/(molÅ^2) for NOE constraints, 500 kcal/(mol rad^2) for torsion angle constraints, and 4 kcal/(mol Å^2) for repulsion terms. F_{NOE}, F_{tor}, and F_{repel} are respectively, the violation energies associated with NOEs, torsion angles, and the van der Waals repulsion energy with hard-sphere van der Waals radii set to 0.8 times the standard value used in the CHARM empirical energy function. F_{L-J} is the van der Waals energy recalculated with the same coordinates but using the X-PLOR energy function.

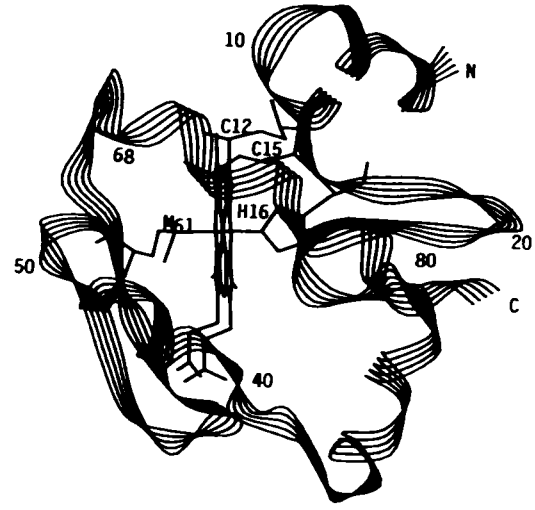

FIGURE 2 Ribbon diagram of the refined, minimized mean structure showing the global folding of PZ c-551.

FIGURE 3 Stereo view of the best-fit superposition of the 20 SA structures. The main-chain atoms are shown along with the heme, its ligands, and the covalent thioether bridges.

is well conserved except for the first two N-terminal residues and the last C-terminal residue. The distribution of rmsd values as a function of residue number is shown in Fig. 4. All nonglycine residues have sterically favorable torsional angles. Although main-chain conformations are very consistent, there are some backbone fluctuations based on torsion angle deviations caused by glycine backbone rotations. The phi angles of Gly-54 and Gly-57 fall into two obvious clusters. The average for Gly-54 is −94° in 10 out of 20 SA structures, and 88° in the other 10; for Gly-57 it is 102° in 5 SA structures and −145° in the other 15. For the two sets, the neighboring residues Gln-53 and Trp-56 correspondingly have average psi angles of −18° (in 10 SA structures) and −166° (in the other 10), and 39° in 5 SA structures and −64° in the other 15, respectively.

Despite the residue substitutions, the global folding of PZ c-551 is homologous to PA and PS c-551. The average rmsd of all main-chain atoms between the minimized average structure of PZ c-551 and PS c-551, and between the minimized average structure of PZ c-551 and the crystal structure of PA c-551 are 1.01 and 1. 19 A, respectively, from residue 3 to 80 plus the heme group. These values must be interpreted with consideration that they include some obvious differences imposed by amino acid substitutions between the proteins. The packing of side chains are different because of substitutions in some regions. This leads to a constant shift of some elements of secondary structure that contribute to increasing the rmsd, even though within this sub-element the folding homology is very high. For example, if only the C-terminal helix of PS and PZ c-551 from residues 69 through 79 are best-fitted to each other, then the rmsd would be 0.35 A.

Homology is not identity, and there are differences. Some are subtle and might reflect the inherent resolution of the structures imposed by the finite number of experimental constraints. Others are clear and can be rationalized in terms of amino acid substitutions, although it would have been impossible to predict de novo the actual observed change.

The N-terminal residue of both PS and PZ c-551 is glutamine. In the case of PS, the main isozyme isolated was a form in which this residue had cyclized to pyroglutamic acid.

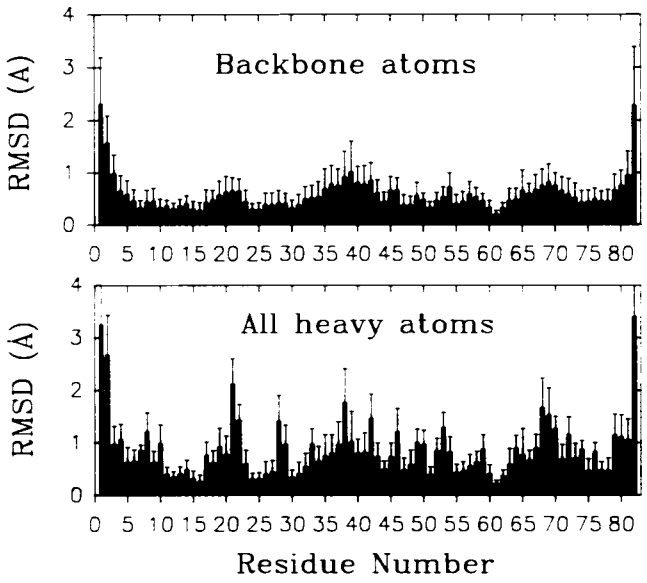

FIGURE 4 Atomic rmsd of the individual SA structures about the mean structure.

This was detected by the observation and assignment of a slowly exchanging pyrrolidone NH at residue 1. For unknown reasons, in the case of PZ the main fraction had normal Gln. There was a minor fraction that could have been the pyroglutamic form of PZ c-551, but there was an insufficient amount for further study. In the case of PS c-551, the isozyme with free Gln was accumulated over time and characterized. At the pH of the NMR sample (pH 5), we could not detect by NMR any conversion to the pyroglutamic form. The free Gln isozyme had minor chemical shift differences for the first few N-terminal residues compared with the pyro-Gln isozyme, but no detectable differences in NOE or torsion angle constraints, so the conclusion is that the state of the N-terminus has little impact on final global structure.

Previous solution state data for PS and PA c-551 indicated that a loop segment of the protein from residues 35 to 40 was disordered. There was a noticeable lack of NOEs and torsion angle constraints for this region, which contributed to multiple computed SA conformations. Quantification of rapid amide hydrogen exchange rates for residues in the loop supported the concept of conformational flexibility (Timkovich et al., 1992). It was a surprise, therefore, to find evidence that PZ c-551 is more ordered in this loop region as shown in the superposition of the SA structures in Fig. 3. One conformation dominated in 17 out of the 20 SA structures and in the minimized average structure. It consisted of type I turns in the sequences 32–35 and 37–40, and a type III turn in the sequence 34–37. For the sequence 37–40, a type II turn was identified in 2 of the 20 SA structures, and one SA structure had a turn not classifiable as any type. These alternative conformations at 37–40 did not violate any NMR constraint, nor did they have a significantly higher total energy. There is still some ambiguity, therefore, about these three residues but, overall, the NMR evidence indicated a higher degree of order for PZ than PA or PS c-551; the PZ data showed clear constraints in this region, whereas PA and PS did not. There were strong NOEs between amide protons of 34 to 33, 35 to 34, 39 to 38, and 40–39, and NOEs from the amide protons of 34 and 35 to the alpha protons of 33 and 34, respectively, and from the amide protons of 39 and 40 to the alpha protons of 38 and 39, respectively, consistent with type I turns in the sequence of 32–35 and 37–40. The small coupling constant, $^3J_{NH\alpha}$ (<5.5 Hz), at residues 33 and 38 and the large coupling constant (>8 Hz) at positions 34 and 39 also support the type I turns in these two segments.

There are three side-chain substitutions in this region among the set of PA, PS, and PZ c-551s. Because the side-chain change from PA Ala-38 to PS Asp-38 does not change the structure in this region (Cai et al., 1992), the additional substitution to PZ Glu-38 would also be expected to be neutral. Substitution PA, PS Gln-37 to PZ Val-37 can also be considered a neutral change, because both residues are uncharged, approximately the same size, and partially solvent exposed. The key change appears to be the substitution of PA Phe-34, PS Tyr-34 to PZ Asn-34. The benzoid ring of residue 34 had been shown to be rapidly flipping in PA and PS. We have no information on whether Asn-34 is flipping in PZ c-551 but, even if it is, its smaller volume leads to a different situation that could account for the more stable conformation observed in PZ. The 34 side chain is mostly buried and packs against the hydrophobic side-chains of the invariant residues Trp-56, Leu-44, Lys-33, and Val-30. The orientation of 34, in turn, directs the chain path for the 35–40 loop. Going to the smaller and possibly nonfluctuating side chain at 34 appears to allow this loop to attain one dominant conformation.

The loss of aromaticity at 34 led to a further consequence involving the orientation Lys 33. In PS and PA c-551, it was shown for this surface residue that the extended lysine aliphatic side chain was not randomly oriented, but had a preferred conformation in a hydrophobic groove created by the side chains of Phe/Tyr-34, Trp-56, and Val-30, with stabilization provided by hydrogen bonding of the terminal amine to the carbonyl of Val-55. The groove has been altered by the Asn-34 substitution in PZ c-551, and Lys-33 no longer possess the untypical scalar coupling features that were the distinguishing characteristic of the previous preferred conformation. The average distance from the terminal nitrogen of Lys-33 to the carboxyl oxygen of Val-55 is 6.8 A in the 20 SA structures, and they showed no preferred conformation as shown in Fig. 5,

FIGURE 5 Stereo view of the superposition of the 20 SA structures in the vicinity of Lys-33 and Asn-34. The view shows Trp-56 and Val-55 of the left side, Val-30 and Asn-34 on the right side, and Lys-33 at the bottom.

whereas it is only 3.39 A in the case of the final refined structure for PS c-551.

Loss of another aromatic residue, PS Phe-27 to PZ Leu-27, correlates with an additional structure change. In PA and PS c-551, the side chain of residue 81 is packed into a hydrophobic pocket comprised of the side chains of residues 27, 31, 77, 78, and 79. In PZ c-551, the substitutions PS Phe-27 (rapidly flipping) to PZ Leu-27 and PS Gln-81 to PZ Leu-81, change the size of the pocket. C-terminal Lys 82 in the SA structures of both PS and PZ was not constrained to a fixed conformation by either NMR data or energy considerations. PS Gln-81 was also disordered. However, data existed for a preferred orientation of Leu-81 in PZ, because a moderate NOE was observed from Cα of Val-78 to the amide of Leu-81, whereas no corresponding NOE was observed in PS. The packing of the two PZ Leu side chains might constitute a better fit serving to render the conformation in this region less flexible.

The heme is relatively buried in the global structure of c-551. Previous crystal and solution results for PA and PS c-551 showed that the polar propionic acid substitutes are stabilized by extensive hydrogen bonds. The inner propionate-17 has one carboxyl oxygen hydrogen-bonded to Trp-56 Nε1 and the other to Arg Nη, in PA c-551, whereas in PS c-551 substitution for His-47 leads to a homologous pattern except for the precise orientation of the imidazole versus guanidinium groups. PZ has His-47, and the NMR constraints of protons in this region were essentially identical to those observed for PS c-551. In the 20 SA structures, the average distances from one inner propionate oxygen to the imidazolium hydrogen and nitrogen of His-47 were 2.09 ± 0.26 and 2.93 ± 0.21 A, and for the other oxygen to the indole hydrogen and nitrogen of Trp-56, they were 1.86 ± 0.08 A, and 2.79 ± 0.07 A, respectively. After 1000 steps of energy minimization with electrostatic and hydrogen-bond terms included in the force field, these numbers reduced to 1.68 ± 0.02, 2.65 ± 0.02, 1.85 ± 0.05, and 2.82 ± 0.05 A, respectively. The hydrogen bond pattern from Trp-56 and His-47 to the heme group is shown in Fig. 6.

The pH behavior of protons involved with, or near, this hydrogen bond network was highly comparable for PZ and PS c-551. For example, titration results for PZ Trp-56 Nπ H as shown in Fig. 7 indicate transitions with pK_a of 3.3 and

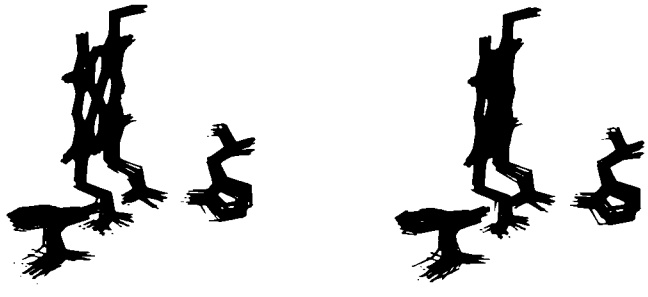

FIGURE 6 Stereo view of the hydrogen bonding pattern for the buried heme propionate side-chains. The final refined orientation of Trp-56 and His-47 are shown.

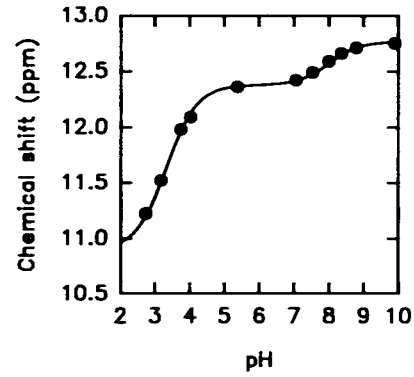

FIGURE 7 The dependence on pH of the chemical shift of PZ Trp-56 Nε1H. The solid line represents theoretical curves corresponding to transitions with pK_a values of 3.3 and 8.0.

8.0. In PS c-551, the values were 3.0 and 8.2 and were assigned to the ionization of both propionates 13 and 17 (pK_a = 3.0) and the imidazole of His-47 (pK_a = 8.2) (Cai and Timkovich, 1992). Thus, the hydrogen bond network is a highly conserved feature of Pseudomonad cytochromes.

PA c-551 has been considered the archtypical representative of the c-551 family, because it was one of the first members to be purified and have primary and tertiary structure determined. The presence of five aromatic residues was thought to be invariant, based upon the sequences available at first. The structure of PZ c-551 now clearly shows that they are not absolutely necessary for attainment of the global c-551 fold, and nonaromatic but relatively hydrophobic residues can be accommodated by rearrangement of only local surroundings.

Financial Support was provided in part by Grant GM43292 from the National Institutes of Health.

REFERENCES

Ambler, R. P., and M. Wynn. 1973. The amino acid sequences of cytochrome c-551 from three species of *Pseudomonas*. *Biochem. J.* 131: 485–498.

Cai, M., E. G. Bradford, and R. Timkovich. 1992. Investigation of the solution conformation of cytochrome c-551 from *Pseudomonas stutzeri*. *Biochemistry*. 31:8603–8612.

Chau, M. H., M. Cai, and R. Timkovich. 1990. NMR comparison of prokaryotic, and eukaryotic cytochromes c. *Biochemistry*. 29:5076–5087.

Crawford, J. L., W. N. Lipscomb, and C. G. Schellman. 1973. The reverse turn as a polypeptide conformation in globular proteins. *Proc. Natl. Acad. Sci. USA*. 70:538–542.

Dohler, K., V. A. R. Huss, and W. G. Zumft. 1987. Transfer of *Pseudomonas perfectomarius* Baumann, Bowditch, Baumann, and Beaman 1983. to *Pseudomonas stutzeri* (Lehmann and Neumann 1896) Sijderius 1946. *Int. J. Syst. Bacteriol.* 37:1–3.

Driscoll, P. C., A. M. Gronenberg, L. Beress, and G. M. Clore. 1989. Determination of the three-dimensional solution structure of the antihypertensive, and antiviral protein BDS-I from the sea anemone *Anemonia sulcata*. *Biochemistry*. 28:2188–2198.

Jungst, A., S. Wakabayashi, H. Matsubara, and W. G. Zumft. 1991. The *nir*STBM region coding for cytochrome cd_1-dependent nitrite respiration of *Pseudomonas stutzeri* consists of a cluster of mono-, di-, and tetraheme proteins. *FEBS Lett.* 279:205–209.

Korner, H., K. Frunzke, K. Dohler, and W. G. Zumft. 1987. Immunochemical patterns of distribution of nitrous oxide reductase, and nitrite reductase (cytochrome cd_1) among denitrifying pseudomonads. *Arch. Microbiol.* 148:20–24.

Nilges, M., G. M. Clore, and A. M. Gronenborn. 1988. Determination of three-dimensional structures of proteins from interproton distance data by hybride distance geometry-dynamical simulated annealing calculations. *FEBS Lett.* 229:317–324.

Timkovich, R., and M. Cai. 1993. Investigation of the structure of oxidized *Pseudomonas aeruginosa* cyt c-551 by NMR: comparison of observed, and paramagnetic shifts, and calculated pseudocontact shifts. *Biochemistry.* 32:11516–11523.

Timkovich, R., L. A. Walker, and M. Cai. 1992. Hydrogen exchange in *Pseudomonas* cytochrome c-551. *Biochim. Biophys. Acta.* 1121:8–15.

Wagner, G., W. Braun, T. F. Havel, T. Schaumman, N. Go, and K. Wuthrich. 1987. Protein structures in solution by NMR, and distance geometry. *J. Mol. Biol.* 196:611:639.

Zuiderweg, E. R. P., R. Boelens, and R. Kaptein. 1985. Stereospecific assignments of 1H-NMR methyl lines, and conformation of valyl residues in the lac repressor headpiece. *Biopolymers.* 24:601–611.

Zumft, W. G., K. Dohler, H. Korner, S. Lochelt, V. A., and K. Frunzke. 1988. Defects in cytochrome cd_1-dependent nitrite respiration of transposon Tn5-induced mutants from *Pseudomonas stutzeri*. *Arch. Microbiol.* 149:492–498.