Biophysical Journal Volume 75 October 1998 1964–1972

Primary Sequence and Solution Conformation of Ferrocytochrome *c*-552 from *Nitrosomonas europaea*

R. Timkovich,* D. Bergmann,# D. M. Arciero,# and A. B. Hooper#
*Department of Chemistry, University of Alabama, Tuscaloosa, Alabama 35487-0336, and #Department of Genetics and Cell Biology, University of Minnesota, St. Paul, Minnesota 55108

ABSTRACT Cytochrome *c*-552 from *Nitrosomonas europaea* is a 9.1-kDa monoheme protein that is a member of the bacterial cytochrome *c*-551 family. The gene encoding for *c*-552 has been cloned and sequenced and the primary sequence of the product deduced. Proton resonance assignments were made for all main-chain and most side-chain protons in the diamagnetic, reduced form by two-dimensional NMR techniques. Distance constraints (1056) were determined from nuclear Overhauser enhancements, and torsion angle constraints (88) were determined from scalar coupling estimates. Solution conformations for the protein were computed by the hybrid distance geometry-simulated annealing approach. For 20 computed structures, the root mean squared deviation from the average position of equivalent atoms was 0.84 Å (σ = 0.12) for backbone atoms over all residues. Analysis by residue revealed there were three regions clearly less well defined than the rest of the protein: the first two residues at the N-terminus, the last two at the C-terminus, and a loop region from residues 34 to 40. Omitting these regions from the comparison, the root mean squared deviation was 0.61 Å (σ = 0.13) for backbone atoms, 0.86 Å (σ = 0.12) for all associated heavy atoms, and 0.43 Å (σ = 0.17) for the heme group. The global folding of the protein is consistent with others in the *c*-551 family. A deletion at the N-terminus relative to other family members had no impact on the global folding, whereas an insertion at residue 65 did affect the way the polypeptide packs against the methionine-ligated side of the heme. The effects of specific substitutions will be discussed. The structure of *c*-552 serves to delineate essential features of the *c*-551 family.

INTRODUCTION

Cytochromes *c*-551 are redox hemoproteins that operate in prokaryotic electron transport chains in a manner functionally equivalent to mitochondrial cytochrome *c* (cyt *c*), namely, as highly specific, soluble electron carriers at a controlled redox potential. Composed of ~80 residues, they are smaller than mitochondrial cyt *c* (normally 103 residues) and therefore represent the most basic unit capable of the intended function. They are highly specific for their physiological redox partners (Dickerson and Timkovich, 1975), and rates of electron transfer to these partners are just as fast as for the mitochondrial counterparts (Yamanaka, 1967). In terms of self-electron exchange rates they are more efficient, in that the rate is three orders of magnitude faster (Timkovich et al., 1988). They operate at the same redox potential as mitochondrial cyt *c* (circa +250 mv), but some versions have the added capability of adjusting their redox potential depending upon pH (Leitch et al., 1984). Mitochondrial cyt *c* generally has a single electron donor (the bc_1 complex) and electron acceptor (cyt aa_3), but in some cell lines, *c*-551's can participate in multiple electron transport chains for aerobic and anaerobic respiration and peroxidase reactions.

Cyt *c*-551 from *Pseudomonas* has been most extensively studied. Crystal structures are known for the oxidized and reduced forms of *P. aeruginosa* (Matsuura et al., 1982), and solution structures based upon NMR are known for *P. aeruginosa* in both redox states (Detlefsen et al., 1991; Timkovich and Cai, 1993), and for reduced *P. stutzeri* (Cai et al., 1992) and *P. stutzeri* substrain *ZoBell* (formerly *P. perfectomarius*) (Cai and Timkovich, 1994). Although there is little sequence homology, the size of *Desulfovibrio* cyt *c*-553 has suggested a relation, and the global fold of *D. vulgaris* Hildenborough (Blackledge et al., 1995) is now known to be similar to that of the *c*-551 family, but with major differences that must contribute to the very low potential (40 mv) for this cytochrome. Cytochromes of the *c*-551 type are found in other prokaryotes such as *Azotobacter* (Campbell et al., 1973) and *Alcaligenes* (Timkovich and Cork, 1984), but have not been well characterized structurally.

Nitrosomonas europaea contains a small cytochrome originally designated as cyt *c*-552 (NE *c*-552) (Yamanaka and Shinra, 1974), in keeping with the conventional numbering of bacterial cytochromes on the basis of the wavelength maximum of the ferrous α-band. In high-resolution optical spectra the maximum is actually closer to 551, but the original name of *c*-552 has been kept to avoid confusion. NE *c*-552 performs multiple physiological tasks. In *N. europaea*, reduced *c*-552 is an electron donor to the terminal cytochrome oxidase (Yamazaki et al., 1988), to a Cu nitrite reductase (DiSpirito et al., 1985), and to a diheme cytochrome peroxidase (Arciero and Hooper, 1994). Oxidized *c*-552 accepts electrons from a diheme cytochrome *c*-554

Received for publication 1 April 1998 and in final form 1 July 1998.
Address reprint requests to Prof. Russell Timkovich, Department of Chemistry, P.O. Box 870336, University of Alabama, Tuscaloosa, AL 35487-0336. Tel.: 205-348-8439; Fax: 205-348-9104; E-mail: rtimkovi@ua1vm.ua.edu.

0006-3495/98/10/1964/09 $2.00

during turnover of hydroxylamine by hydroxylamine oxidase (HAO). Cyt *c*-552 does not accept electrons directly from HAO, but requires *c*-554 to mediate (Yamanaka and Shinra, 1974). Initial NMR characterization of NE *c*-552 indicated it belonged to the *c*-551 family, but also suggested some structural changes from other members (Timkovich et al., 1994). The complete amino acid sequence (this work) indicates, at most, 48% homology to other *Pseudomonas* *c*-551's. Insertions, deletions, and changes of previously considered invariant residues have occurred. An NMR-based solution structure for the reduced form of NE *c*-552 was determined to reveal the three-dimensional consequences of these changes within the *c*-551 family.

MATERIALS AND METHODS

Isolation and purification of NE *c*-552 have been described (Arciero et al., 1994). The DNA sequence of the gene coding for *c*-552 was determined by the following procedure. Genomic DNA of *N. europaea* was harvested as described by McTavish et al. (1993a). Purification of plasmid DNA by alkaline lysis, restriction digests, electrophoresis, fragment purification, and sequencing by the dideoxy-chain termination technique were performed by standard methods (Sambrook et al., 1989). DNA sequences were analyzed with the GCG Sequence Analysis Software Package (Genetics Computer Group, Madison, WI).

A degenerate oligonucleotide, 5′-AA(A/G)-AA(A/G)-AA(C/T)-AA(C/T)-TG(C/T)-ATI-GCI-TG(C/T)-CA(C/T)-CA-3′, was synthesized based on the N-terminal amino acid sequence of *c*-552 (Logan, 1991; Miller and Nicholas, 1986). This probe hybridized to a 3.1-kb *Eco*RI genomic fragment (McTavish et al., 1993b) and was used to screen a library of *Eco*RI fragments in pUC119, yielding a clone containing the entire open reading frame (*cyt*) encoding cyt *c*-552. The nucleotide sequence of cyt was deposited in the GenBank data base under accession number U86756.

Possible −35 (CTGTCA) and −10 (TAGAAT) σ70 consensus promotor sequences are present at bases 8–37. The start codon ATG is located at bases 74–76, nine bases downstream from a Shine-Delgarno sequence (AAAAGG). N-terminal sequence determination (Logan, 1991) agreed with a previously published N-terminal sequence (Miller and Nicholas, 1986) and indicated that an initial 22 residue signal sequence had been removed such that the mature protein began Asp-Ala-Asp-Leu, etc. The amino acid sequence deduced from the DNA gene sequence was further confirmed by mass spectrometry. Protein dissolved in 1:1 acetonitrile: water with 0.1% formic acid was injected at 10 μl/min through the electrospray interface of a Perkin-Elmer Sciex API/III spectrometer. The main cluster was observed for multiply charged species 6–10. The observed molecular mass of 9099 (σ = 1.4) agreed well with the calculated mass (9098.2) based upon the sequence. While this work was in progress, Fujiwara et al. (1995) independently published the *c*-552 sequence derived by Edman methods.

A series of samples were prepared in both 9:1 ^1H_2O:^2H_2O and 99.9% ^2H_2O at concentrations ranging from 2 to 5 mM and pH values from 4 to 8, although final chemical shifts will be reported for one set of conditions. The dilute samples were generally used to resolve assignment ambiguities. Spectra recorded at different pH values and temperatures allowed us to resolve overlapping sets of correlated resonances on the basis of their changes in chemical shifts under different experimental conditions. The more concentrated samples were used to improve signal to noise and quantify weak NOEs. HOHAHA spectra with mixing times from 9 to 75 ms were used to identify coupled spin subsystems. DQF-COSY spectra were used to confirm directly coupled spins and estimate coupling constants. NOESY spectra with mixing times from 75 to 150 ms were used to establish dipolar correlations, to group NOE intensities into constraint ranges, and to test for spin diffusion effects. Further details have been reported previously (Cai and Timkovich, 1994).

RESULTS

Resonance assignments

Because of the strong structural homologies that have been found for NE *c*-552 in relation to other *c*-551's, and to facilitate discussion of these, the sequence of residues in NE *c*-552 will be denoted by the number of the corresponding analogous residue in *P. aeruginosa*, the archetypical *c*-551. Thus NE *c*-552 begins with the N-terminal residue numbered as Asp^3. An insertion relative to *P. aeruginosa* causes two residues to be denoted as Val^{65A}-Asn^{65B}. The assignment process benefitted from the chemical shift homologies shown by NE *c*-552 to other *c*-551's, especially for landmark resonances such as the heme, ligands, thioethers, and invariant residues. However, the homologies were used only as an initial clue, and the assignments rest upon standard sequential connectivities. Sequential connectivities were broken only at Pro^{25}, Pro^{60}, and Pro^{62}. Pro^{60} and Pro^{62} flank the invariant heme ligand Met^{61}, which has highly characteristic chemical shifts because of the heme ring current. The three peptide segments 3–24, 26–59, and 63–82 have individualistic sequence features that make their placement very firm. Resonance assignments are reported in Table 1.

The secondary structure was dominated by four major α-helices, and there were no detectable segments of β-sheet. The α-helical regions showed the characteristic constraints of this type of regular secondary structure, including small NH-CαH coupling constants and a repetitive pattern of NOEs. Strong NOEs involving amide NH-NH(i, i + 1) were observed in the spans i = 5–15, i = 27–34, i = 39–50, and i = 68–80. NOEs involving amide NH-NH(i, i + 2) were observed in the spans i = 4–14, i = 27–32, i = 39–49, and i = 68–79. NOEs involving amide NH-CβH(i, i − 1) were observed for i = 4–17 (except for i = 14), i = 28–36, i = 40–51 (except for Gly residues in this span), and i = 68–81. NOEs involving amide NH-CαH(i, i − 3) were observed in the spans i = 6–15, i = 30–35, i = 43–50 (except i = 49), and i = 72–81. The associated segments ultimately refined into regular α-helical secondary structure.

Structure computations

The general strategy was based upon the hybrid distance geometry-dynamic simulated annealing (SA) approach of Nilges et al. (1988). Most details have already been provided in previous reports from this group (Cai and Timkovich, 1994), and this section will focus on modifications of past procedures. Distance constraints were categorized as follows. For NOEs involving an amide NH, strong was assigned as 1.8 to 2.9 Å, medium 1.8 to 3.5 Å, weak 1.8 to 5.0 Å, and very weak 1.8 to 6.0 Å. For NOEs involving an amide and a methyl, strong was assigned 1.8 to 3.4 Å, medium 1.8 to 4.0 Å, weak 1.8 to 5.5 Å, and very weak 1.8 to 6.5 Å. For NOEs involving a methyl with a nonamide, strong was assigned 1.8 to 3.2 Å, medium 1.8 to 3.8 Å, weak 1.8 to 5.5 Å, and very weak 1.8 to 6.5 Å. For NOEs

TABLE 1 Proton resonance assignments for *N. europaea* cyt *c*-552

Ordinal no.*	Residue#	Chemical shift (ppm)§ NH	Cα	Cβ	Cγ	Cδ	Cε	Other
1	Asp[3]		4.19	2.55				
				2.70				
2	Ala[4]	8.18	3.45	1.25				
3	Asp[5]	7.88	4.25	2.55				
4	Leu[6]	7.82	3.96	1.59	1.55	0.87		
				1.35		0.82		
5	Ala[7]	7.75	4.19	1.15				
6	Lys[8]	7.73	4.05	1.95	1.35			
				1.85	1.45			
7	Lys[9]	8.35	4.05	1.95	1.53			
				1.80				
8	Asn[10]	7.40	5.01	3.24				NH_2 7.69, 7.32
				2.98				
9	Asn[11]	8.04	4.76	3.62				NH_2 7.74, 6.90
				3.08				
10	Cys[12]	8.94	5.15	2.73				
				2.45				
11	Ile[13]	6.92	4.62	1.98	1.15	0.78		
					1.35			
					0.85 (CH_3)			
12	Ala[14]	7.37	4.12	1.71				
13	Cys[15]	6.46	4.37	1.97				
				1.05				
14	His[16]	7.13	3.75	0.85				NπH 9.23, C5 0.58, C2 0.45
15	Gln[17]	7.29	4.16	1.62	2.06			NH_2 7.19, 6.58
16	Val[18]	8.48	3.16	1.94	0.94			
					0.76			
17	Glu[19]	8.32	4.13	2.07	1.96			
				1.83				
18	Thr[20]	6.46	4.37	3.72	0.87			
19	Lys[21]	8.06	3.73	1.55	0.85	1.45	2.77	
					0.95			
20	Val[22]	7.22	3.87	1.19	0.68			
					0.40			
21	Val[23]	7.13	4.08	1.95	1.62			
					1.20			
22	Gly[24]	6.68	3.58					
			0.12					
23	Pro[25]		3.53	1.11	0.10	2.86		
				0.46	0.00	1.90		
24	Ala[26]	8.16	3.66	0.65				
25	Leu[27]	8.23	3.20	1.15	0.58	0.23		
				0.73		−0.93		
26	Lys[28]	8.42	3.77	1.62	1.24	1.50	2.77	
				1.50			2.85	
27	Asp[29]	6.76	4.49	2.63				
				2.52				
28	Ile[30]	7.12	3.97	1.56	0.75	−0.23		
					1.03			
					1.85 (CH_3)			
29	Ala[31]	7.67	4.12	1.45				
30	Ala[32]	7.41	4.02	1.52				
31	Lys[33]	7.72	3.92	1.90	0.92	1.22	2.50	
				1.70	0.42			
32	Tyr[34]	7.33	4.77	3.52				C2, 6 7.38: C3, 5 7.16
				2.55				

*Order of occurrence of the residues in the mature protein.

#Residue numbers are assigned based upon homology to *Pseudomonad* cyt's *c*-551. Compared to these, *c*-552 has a deletion of two residues at the amino terminus and an insertion of one residue at position 65.

§Chemical shifts were referenced against 3-(trimethylsilyl) tetradeutero sodium propionate as 0 ppm. Values are reported for pH 7 and 323 K. Underlined numbers are IUB-IUPAC designations.

TABLE 1 Continued

Ordinal no.*	Residue#	Chemical shift (ppm)§ NH	Cα	Cβ	Cγ	Cδ	Cε	Other
33	Ala[35]	7.47	4.05	1.49				
34	Asp[36]	8.23	4.69	2.65				
				2.70				
35	Lys[37]	7.63	4.42	2.00	1.40			
				1.55				
36	Asp[38]	8.68	4.38	2.72				
37	Asp[39]	8.71	4.67	2.68				
				2.79				
38	Ala[40]	7.48	3.86	1.73				
39	Ala[41]	8.86	3.99	1.52				
40	Thr[42]	7.53	3.82	4.13	1.20			
41	Tyr[43]	7.82	4.10	2.95				C2, 6 6.76; C3, 5 6.55
				3.19				
42	Leu[44]	8.96	3.65	2.20	2.15	1.12		
				1.35		1.08		
43	Ala[45]	8.32	3.85	1.34				
44	Gly[46]	7.46	3.61					
			3.76					
45	Lys[47]	7.18	3.52	1.56	0.81			
46	Ile[48]	7.91	1.57	1.77	1.90	1.36		
					1.05			
					0.70 (CH_3)			
47	Lys[49]	7.16	3.58	1.40				
				1.21				
48	Gly[50]	7.67	3.76					
			3.40					
49	Gly[51]	7.09	3.43					
			1.79					
50	Ser[52]	7.23	4.26	2.90				
				2.55				
51	Ser[53]	8.32	4.55	3.62				
52	Gly[54]	8.79	3.98					
			4.17					
53	Val[55]	10.23	3.83	2.55	0.63			
					1.03			
54	Trp[56]	10.69	4.66	3.83				C2, 7.78; Nε, 12.53;
				3.57				C4, 7.95; C5, 7.20;
								C6, 6.99; C7, 7.52
55	Gly[57]	7.72	4.10					
			4.45					
56	Gln[58]	8.90	4.55	2.05	2.40			NH_2 7.82, 7.53
					2.45			
57	Ile[59]	7.62	4.72	2.15	1.12	1.84		
					1.46			
					1.35 (CH_3)			
58	Pro[60]		4.81	2.05	1.50	3.90		
				2.10		4.10		
59	Met[61]	8.50	3.57	−1.02	−0.78	−3.15		
				−2.85	−3.73			
60	Pro[62]		4.09	2.14	1.87	2.95		
						3.85		
61	Pro[63]		2.90	1.45	1.74	3.24		
				1.60	1.82	3.32		
62	Asn[64]	6.48	4.71	2.39				NH_2 7.11, 3.35
				2.29				
63	Val[65A]	7.02	3.75	2.04	0.82			
					0.84			
64	Asn[65B]	8.58	4.70	2.75				NH_2 7.49, 6.80
					2.95			
65	Val[66]	7.12	4.14	1.90	0.96			
66	Ser[67]	9.16	4.35	4.09				
67	Asp[68]	8.73	4.16	2.57				
68	Ala[69]	8.32	4.08	1.38				

TABLE 1 Continued

Ordinal no.*	Residue#	Chemical shift (ppm)§ NH	Cα	Cβ	Cγ	Cδ	Cε	Other
69	Asp[70]	8.02	4.42	2.95				
				2.31				
70	Ala[71]	8.75	4.00	1.45				
71	Lys[72]	7.59	4.02	1.98	1.50			
				1.83				
72	Ala[73]	7.98	4.28	1.52				
73	Leu[74]	9.02	4.35	1.66	1.95	1.15		
				2.51		1.32		
74	Ala[75]	8.19	4.06	1.76				
75	Asp[76]	8.53	4.31	2.82				
				2.64				
76	Trp[77]	8.14	4.45	3.64				C2, 7.28; Nε 10.15;
				3.35				C4, 7.65; C5, 6.45;
								C6, 6.53; C7, 7.22
77	Ile[78]	8.92	2.70	2.06	0.55	1.37		
					2.30			
					0.74 (CH_3)			
78	Leu[79]	7.42	3.94	1.65	1.95	0.97		
				1.95		1.02		
79	Thr[80]	7.72	4.34	4.34	1.46			
80	Leu[81]	7.44	3.99	0.92	1.17	−0.17		
				1.53		0.26		
81	Lys[82]	7.02	3.93	1.63	1.59	1.28	2.95	
				1.61				

Heme assignments				
Substituent IUB-IUPAC designation, chemical shift (ppm)				
Meso	5 9.85	10 9.33	15 9.39	20 9.16
Methyl	2^1 3.52	7^1 3.70	12^1 3.19	18^1 3.29
Thioether	3^1 6.01	3^2 1.94	8^1 6.25	8^2 2.45
Propionate	13^1 4.66, 4.01	13^2 3.53, 2.58	17^1 4.45, 4.25	17^2 3.45, 2.78

not involving an amide and also not involving a methyl, strong was assigned 1.8 to 2.7 Å, medium 1.8 to 3.3 Å, weak 1.8 to 5.0 Å, and very weak 1.8 to 6.0. Torsion constraints and stereospecific assignments were as described previously (Cai and Timkovich, 1994).

Computations used the software package XPLOR, version 3.851, on a Cray C90. Molecular graphics and some simple geometry analysis were performed with the software package SYBYL (Tripos Associates), version 6.3. Side chains corresponding to the sequence of NE *c*-552 were added from the standard SYBYL library to the backbone coordinates of *P. ZoBell* (Brookhaven PDB accession 1CCH). The result was refined by conjugate gradient energy minimization to remove obvious steric conflicts created by the crude superimposition. The result was used strictly as a template to define local geometry for the next stage of distance-geometry calculations. A set of 88 torsion angle constraints and an initial set of 764 well-resolved and unambiguously assigned NOEs was used as experimental constraints to compute a family of 20 distance-geometry structures through the main steps of substructure embedding, full structure simulated annealing regularization, and simulated annealing refinement (Cai and Timkovich, 1994). The average of these was further refined by 1000 steps of restrained Powell energy minimization to produce a model of the NE *c*-552 structure. The model structure was then used to resolve ambiguous NOE assignments (see Cai et al., 1992, for specific examples of this process) and confirm very weak NOEs. This allowed the set of distance constraints to be expanded to the final set of 1056, including 400 intraresidue, 288 sequential (for residues i and j, $j - i = 1$), 158 medium-range ($1 < j - i < 5$), and 210 long-range ($j - I > 4$) constraints.

Using the expanded set of constraints, a new series of simulated annealing computations was performed with a protocol kindly donated by Dr. Mengli Cai, a postdoctoral fellow currently working in the laboratory of Dr. G. M. Clore, Laboratory of Chemical Physics, National Institutes of Health. The above model structure was assigned random initial velocities and annealed at 3000 K for 10 ps. Nonbonded energy terms were turned off for all atoms except α-carbons, which were given a pure repulsive term. Figuratively speaking, this allows side chains to pass through each other and cross the main chain while in general excluding two residues from occupying exactly the same space. The expanded set of distance constraints (2 kcal) and torsion constraints (10 kcal) were given low force constants. During a cooling period of 25 ps to 50 K, in temperature

steps of 50 K, maintaining coupling to the bath for 430 fs, the experimental constraints (2–30 kcal for distances; 10–200 kcal for torsions) and the nonbonded repulsive interactions for all atoms (1–1000 kcal) were continuously increased. After cooling, the structures were further refined by 500 steps of restrained Powell energy minimization.

Forty SA structures were computed. The 20 with lowest energy were selected for further statistical analysis. In this set there was a total of three distance violations greater than 0.5 Å and no torsion angle violations greater than 5°. The root mean squared displacement (rmsd) from the average structure was 0.84 Å ($\sigma = 0.12$) for backbone atoms over all residues (Fig. 1). Analysis by residue revealed there were three regions clearly less well defined than the rest of the protein: the first two residues at the N-terminus, the last two at the C-terminus, and a loop region from 34 to 40 that will be discussed in depth later. Omitting these regions from the comparison, the rmsd was 0.61 Å ($\sigma = 0.13$) for backbone atoms, 0.86 Å ($\sigma = 0.12$) for all associated heavy atoms, and 0.43 Å ($\sigma = 0.17$) for the heme group.

The average of the accepted SA structures was further refined by 1000 steps of restrained Powell energy minimization to remove distortions from idealized residue local geometry introduced by the simple averaging process. This structure, designated $\langle SA \rangle_r$, had no torsion angle violations greater than 5°, 43 distance violations greater than 0.1 Å, 13 distance violations greater than 0.25 Å, and no distance violations greater than 0.5 Å. The rmsds were as follows: for distance violations, 0.051 Å; for torsion violations, 0.070°; for bonds, 0.003 Å; for angles, 0.63°; for improper angles, 0.49°.

Up to this point no explicit hydrogen bonding interactions had been used as restraints, and the hydrogen bond energy term was not included. For reasons to be discussed in depth later, we explicitly identified the heme propionate oxygens as potential hydrogen bond acceptors and nearby amines and amides as potential donors, and introduced standard XPLOR hydrogen bonding energy terms with force constants between 3 and 30 kcal. Addition of this energy term did lead to the identification of potential hydrogen bonds involving the heme propionate. Additional putative hydrogen bonds in $\langle SA \rangle_r$ were identified involving residues in α-helices on the basis of donor, hydrogen, acceptor distances, and angles. These were added to the constraints list as tight distance constraints between the hydrogen and the acceptor (1.80–2.01 Å) and the donor and acceptor atoms (2.60–3.10 Å). The structure was then further refined with 1000 steps of restrained Powell energy minimization to generate a structure with idealized hydrogen bonds designated as $\langle SA \rangle_{hb}$. Coordinates for $\langle SA \rangle_r$ and $\langle SA \rangle_{hb}$ have been deposited with the Brookhaven Protein Data Bank as 1A8C and 1A56, respectively.

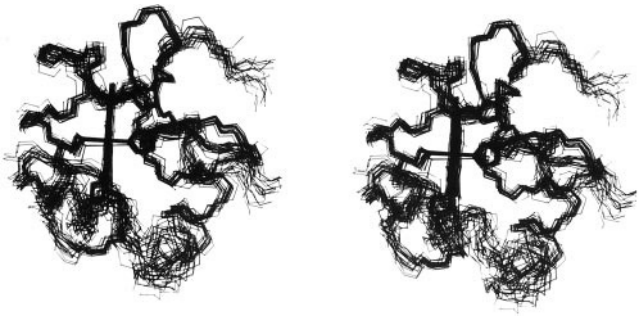

FIGURE 1 Stereo view of the best-fit superposition of the 20 simulated-annealing structures. The main-chain atoms are shown along with the heme, its ligands, and the covalent thioether bridges. The poorly defined regions are the N-terminus (*in the top right corner*), the C-terminus (*right center*), and the 34–40 loop (*bottom right*).

DISCUSSION

The global folding of NE *c*-552 (Fig. 1) is clearly homologous to previously studied members of the *c*-551 family. There is a high α-helical content. The N-terminus forms an α-helix that bends at the thioether bonds (Cys^{12} and Cys^{15}) to orient His^{16} for ligation to the heme. Then a loop from 17 to Pro^{25} forms the histidine side of the heme crevice. A second major segment of α-helix from residues 26 to 33 is then formed. A loop-type substructure from 34 to 40 was not well defined by the NMR constraints available for NE *c*-552. This segment is also poorly defined in the NMR-based structures of *P. stutzeri* (Cai et al., 1992) and *P. aeruginosa* (Detlefsen et al., 1991) *c*-551. In the case of *P. aeruginosa* there is additional data that indicate that the segment is disordered in solution. In the crystal determination, this region has abnormally high temperature factors (Matsuura et al., 1982), and NMR measurements of amide hydrogen exchange rates showed that the fastest rates were for this segment (Timkovich et al., 1992). Disorder of this segment is not a universal characteristic of the family, because in the solution determination for *P. ZoBell c*-551, the region was well defined by ample constraints (Cai and Timkovich, 1994). An interesting correlation is that the three ill-defined segments have an aromatic residue at position 34 (Tyr or Phe), whereas *P. ZoBell* has no aromatic here or anywhere else in the loop. It is a characteristic of NE *c*-552 and the other investigated *c*-551's that all Tyr and Phe aromatic rings are rapidly flipping on the NMR time scale. The side chain of residue 34 packs tightly against other hydrophobic residues and is only partially solvent exposed. It is intriguing to speculate (without proving causality) that motion of the aromatic ring is linked to segment flexibility (the tail wags the dog).

After the loop, another α-helix runs from 40 to 51. Then a loop structure folds back upon itself from 52 to 57. In this loop are residues highly conserved within the family: Gly^{51}, Ser^{52}, Gly^{54}, Val^{55}, Trp^{56}, and Gly^{57}. This loop allows the side chain of Trp^{56} to hydrogen bond to a heme propionate, which will be discussed at length later. In the original crystal structure of *P. aeruginosa c*-551 (Dickerson et al., 1976), it was noted that the span from 58 to 64, which served to orient the Met^{61} side chain for ligation to the

heme, contained a large number of proline residues, and adopted a conformation analogous to an extended polyproline helix. NE *c*-552 has two fewer proline residues in this span compared to the *Pseudomonads*, but the conformation remains the same. Asn^{64} is an invariant residue in the *c*-551 family. In all cases studied by NMR, it has a highly unusual chemical shift of ~3 ppm for one of the side-chain amide protons. This is because the side chain packs tightly against the heme with a close approach of 2.65 Å between 2HD2 and the heme pyrrole atom C1C (PDB atom nomenclature). The amide proton is above the heme plane and experiences a large upfield ring current shift contribution. Around 64 the polypeptide makes a ~90° degree turn, and the C-terminal residues from 67 to 82 form the final α-helix.

The backbone of NE *c*-552 was compared to the backbone of *P. ZoBell c*-551 by superposition of regions that do not involve disorder or insertions or deletions: residues 5–34, 42–63, and 67–80. The rmsd was 1.9 Å. The comparison might seem worse than it really is because it is the displacement between the two rather than the difference from an average, and side-chain differences do lead to subtle and different packing. If one compares smaller portions such as just the C-terminal helix backbone from residues 67–79, the rmsd is much smaller, 0.63 Å.

There are key features of the NE *c*-552 structure associated with special residues. The shortening of the N-terminus in NE *c*-552 compared to other *c*-551's has no obvious effects on the rest of the structure. The sequence of NE *c*-552 after the methionine ligand is M61-P62-P63-N64-V65A-N65B-V66, whereas in *Pseudomonads* the consensus is M61-P62-P63-N64-P65-V66. There is thus a net insertion that looks like a small-scale gene duplication of the NV motif. The DNA sequence is ACC(N64)-GTC(V65A)-ACC(N65B)-GTG(V66). The insertion occurs just before the turn to form the C-terminal helix. The extra length causes a bend that is viewed as packing closer to the heme and the Met ligand (see Fig. 2). The NMR spectrum of ferric NE *c*-552 indicated it had a different distribution of unpaired spin density around the heme (Timkovich et al., 1994), and its electron paramagnetic resonance spectrum indicated a highly axial low spin type signal different from *Pseudomonad c*-551's (Arciero et al., 1994). It is not possible to establish cause and effect at this time, but the different structure observed for NE *c*-552 may be related to these electronic perturbations. There are no obvious major structural differences on the His^{16} side.

The sulfur of free methionine is prochiral. When one of its lone pairs of electrons forms a permanent bond with iron, it is surrounded by four distinct groups ($C\epsilon H_3$, $C\gamma H_2$, Fe, and the remaining lone pair) and becomes chiral. In cytochromes the two isomers are distinguishable by characteristic NOEs to heme substituents (Timkovich et al., 1994, with references to earlier literature). The NMR data for NE *c*-552 showed its Met ligand had the same chirality as found in other *Pseudomonad c*-551's, which is opposite that of mitochondrial cyt *c*.

When cysteinyl sulfur adds across protoporphyrin IX vinyl groups to form the two covalent thioether bridges, two new chiral centers are created at the inner substituent carbons now surrounded by pyrrole, hydrogen, sulfur, and a methyl group (see Fig. 3). The chirality at these centers is readily determined by a characteristic NOE pattern that can be explained by the following considerations (see Fig. 4). The consensus heme c binding motif of -Cys-X-Y-Cys-His- is a short polypeptide span. Ligation of His to the heme iron is firmly established by the unusual and highly characteristic chemical shifts observed for the imidazole ring and Cβ protons, because of the strong heme ring current. The cysteinyl sulfurs are so close to the His ligand that they must approach the heme edge from the same side. For each chiral thioether, this fixes the relative positions of two of the four distinct groups (pyrrole and sulfur). In the S-configuration

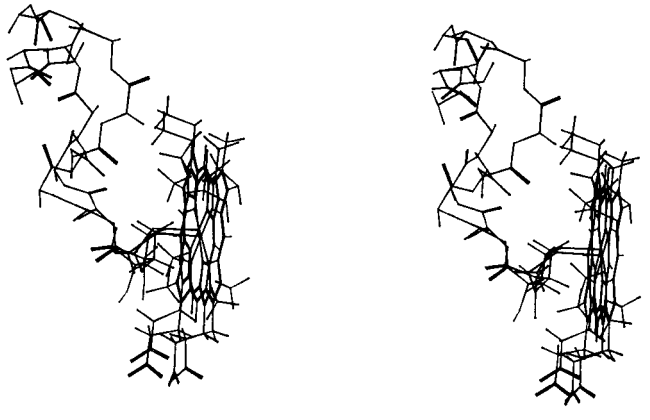

FIGURE 2 Stereo view of the folding of the peptide span from the methionine ligand Met^{61} to residue 67 and its relation to the heme. The structure from *P. ZoBell c*-551 (representative of all *Pseudomonads*) is superimposed on the deduced structure for NE *c*-552. The strand shifted to the right is the NE-c552 conformation.

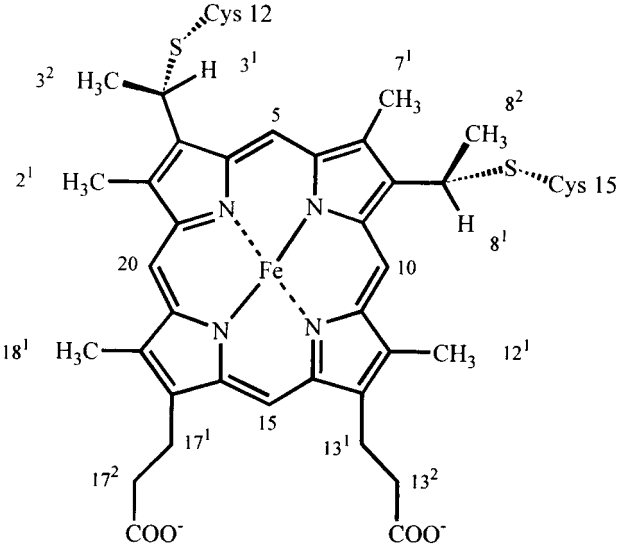

FIGURE 3 The covalent structure of the heme c group with substituents labeled according to IUB-IUPAC nomenclature for tetrapyrroles.

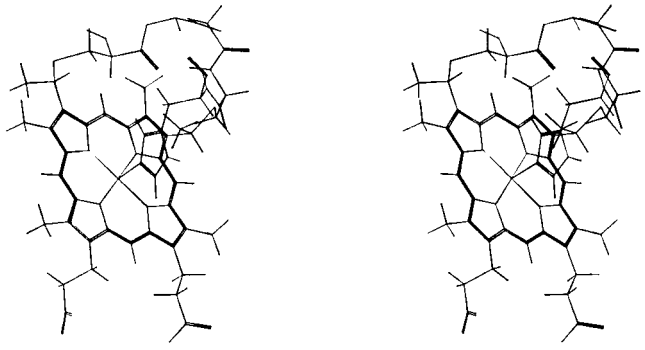

FIGURE 4 Stereo view of the thioether bridges in the computed structure of NE *c*-552.

(Ingold nomenclature) the methine proton (3^1 or 8^1) points directly at the adjacent meso proton (5 or 10), whereas the thioether methyl (3^2 or 8^2) points toward the ring methyl (2^1 or 7^1). In the R-configuration these must be reversed because, as explained, the pyrrole and sulfur are relatively fixed. In NE *c*-552 very intense NOEs were observed between 3^1 and meso 5 and 8^1 and meso 10. These are among the most intense NOEs observed in the protein, comparable to that observed for geminal protons, such as those in glycine residues. The NOEs 3^1 to 2^1 and 8^1 to 7^1 were weak, but the 3^2 methyl to 2^1 and the 8^2 methyl to 7^1 were strong. This established the S-configuration for NE *c*-552. In the refined structure, the methine-to-meso distances were 1.94 Å, methine-to-ring methyl carbon, 4.23 Å, and thioether methyl-to-ring methyl, 3.35 Å. This is the same chirality as observed in previous NMR-based and crystallography-based structures for *Pseudomonad c*-551's and is the same found in mitochondrial cyt *c*.

NE *c*-552 substitutes glutamine at position 58 for a conserved proline in *Pseudomonads*. No major main-chain conformational change results, but Gln^{58} demonstrates unusual spectroscopic characteristics. Its spin system is readily apparent in the fingerprint region of HOHAHA and NOESY spectra at pH 5 between 295 K and 323 K, but it disappears at pH 7 above a temperature of 300 K. The flanking residues Gly^{57} and Ile^{59} are readily observed at both pH values over the complete temperature range and have only a slight dependence for their chemical shift values on these variables. This suggests that there is no major conformational change in this region that depends upon pH or temperature. At pH 7 there is a very intense exchange cross-peak between the Gln^{58} amide and the water resonance. The conclusion is that the Gln^{58} amide is exchanging very rapidly with the solvent above pH 5. In our study of four members of the *c*-551 family, this is the only main-chain amide ever observed to be in such rapid exchange. Of course, the precise position 58 in the other *c*-551's cannot exchange, because there are proline residues here.

The crystal structure of *P. aeruginosa* first revealed a characteristic hydrogen bond between the carbonyl of Pro^{25} and His^{16} NπH. Without the inclusion of hydrogen bonding energy terms (vide infra), these atoms were found in $\langle SA \rangle_r$ to be in good potential hydrogen bonding arrangement in NE *c*-552. The donor-acceptor distance was 2.68 Å, and the hydrogen-acceptor distance was 1.81 Å. This hydrogen bond thus appears to be an invariant characteristic of the *c*-551 family and contributes to fixing of the orientation of the imidazole ring with respect to the heme plane (Timkovich, 1979).

The heme c in *c*-551 is unusual compared to hemoproteins like the globins in that the propionate substituents are mostly buried in the interior. The least solvent-accessible 17-propionate was found in both the crystal structure of *P. aeruginosa c*-551 and in other solution-based structures to be hydrogen bonded to the indole NH of the invariant Trp^{56} and the side chain of Arg/His^{47}. In the family of simulated annealing structures computed without hydrogen bond constraints or force constants, the final geometries indicated a strong hydrogen bond for Trp^{56}. The donor indole-nitrogen to acceptor propionate-oxygen distance averaged 2.7 Å (σ = 0.3 Å) over the 20 structures. Position 47 in other *c*-551's is either an arginine or a histidine residue, but in NE *c*-552 the side chain is lysine. The NMR-based computations were inconclusive about the location of the side chain of Lys^{47}. The donor nitrogen-to-acceptor propionate oxygen distance averaged 4.4 Å, but with a large standard deviation of 0.9 Å. In previous *c*-551 determinations, the full side chains of Arg/His^{47} were assigned, and NOE constraints involving these were obtained. In the present case a factor contributing to the failure to place the side chain definitively was the unfortunate degeneracy of the Lys^{47} β,γ-protons and our failure to complete the lysyl side chain assignments. Because of its size and charge, we feel that a lysine substitution is probably a near-neutral change and that the hydrogen bond does exist in the molecule. This cannot be proved with the current NMR constraints, but it can be shown that such a hydrogen bond is consistent with the rest of the data. The heme propionate oxygens and the Trp^{56} and Lys^{47} nitrogens were explicitly identified within the context of XPLOR as potential acceptors and donors, and a explicit hydrogen bonding energy term was included in further energy minimization trials starting with the coordinates of $\langle SA \rangle_r$. The force constant was varied from 3 to 30 kcal. At low values, local geometry did not change, but as the force constant increased, the side chain of Lys^{47} formed increasingly better hydrogen bonds without violating experimental constraints or raising the other energy terms. The final H-bond geometry was a distance of 2.92 Å for the oxygen-to-Lys^{47} nitrogen and an angle of 167° for O-H-N (Fig. 5).

An additional 32 hydrogen bond candidates were then identified for main-chain amides by two criteria. In $\langle SA \rangle_r$ they already possessed good hydrogen bond geometry (donor, H, acceptor distances, and angle), and they were among the slowest exchanging amides when samples were dissolved in deuterium oxide. They were all amide-to-carbonyl hydrogen bonds in α-helical segments, and the amides showed helical characteristic NOE patterns for $d_{NN}(i, i + 1)$, $d_{\beta N}(i, i + 1)$, and $d_{NN}(i, i + 2)$. They were included in

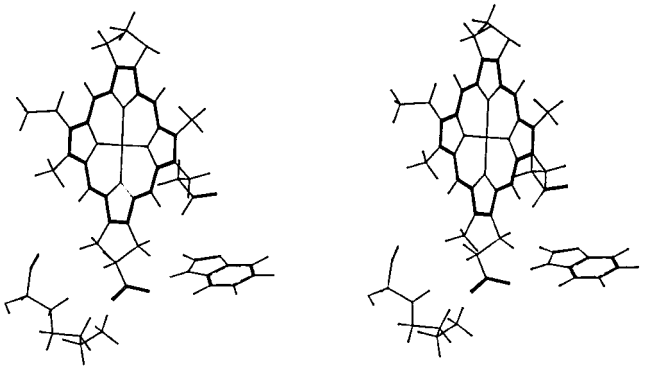

FIGURE 5 Stereo view of the proposed hydrogen bonding around the buried heme propionate substituent in NE *c*-552.

the experimental constraints as distance constraints between donor, hydrogen, and acceptor atoms, and the structure was refined by energy minimization to generate the final structure, termed $\langle SA \rangle_{hb}$. Coordinates with (1A56) and without (1A8C) the hydrogen bond constraints have been deposited with Brookhaven PDB.

In summary, despite insertions, deletions, and substitutions, NE *c*-552 clearly belongs to the *c*-551 family. Its changes help define the limits to which the family may be stretched, but show that the core of global folding is remarkably preserved.

REFERENCES

Arciero, D. M., and A. B. Hooper. 1994. A di-heme cytochrome *c* peroxidase from *Nitrosomonas europae* catalytically active in both the oxidized and half-reduced states. *J. Biol. Chem.* 269:11878–11886.

Arciero, D. M., Q. Peng, J. Peterson, and A. B. Hooper. 1994. Identification of axial ligands of cytochrome c552 from *Nitrosomonas europae*. *FEBS Lett.* 342:217–220.

Blackledge, M. J., S. Medvedeva, M. Poncin, F. Guerlesquin, M. Bruschi, and D. Marioin. 1995. Structure and dynamics of ferrocytochrome c553 from *Desulfovibrion vulgaris* studied by NMR spectroscopy and restrained molecular dynamics. *J. Mol. Biol.* 245:661–681.

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.

Cai, M., and R. Timkovich. 1994. The solution conformation of cytochrome *c*-551 from *P. stutzeri ZoBell* determined by NMR. *Biophys. J.* 67:1207–1215.

Campbell, W. H., W. H. Orme-Johnson, and R. H. Burris. 1973. A comparison of the physical and chemical properties of four cytochromes *c* from *Azotobacter vinelandii*. *Biochem. J.* 135:617–630.

Detlefsen, D. J., V. Thanabal, V. L. Pecoraro, and G. Wagner. 1991. Solution structure of Fe(II) cytochrome *c*551 from *Pseudomonas aeruginosa* as determined by two-dimensional ^{1}H NMR. *Biochemistry.* 30: 9040–9046.

Dickerson, R. E., and R. Timkovich. 1975. Cytochromes *c*. *In* The Enzymes. P. O. Boyer, editor. Academic Press, New York. 397–547.

Dickerson, R. E., R. Timkovich, and R. J. Almassy. 1976. The cytochrome fold and the evolution of bacterial energy metabolism. *J. Mol. Biol.* 100:473–491.

DiSpirito, A. A., L. R. Taaffee, J. D. Lipscomb, and A. B. Hooper. 1985. A "blue" copper oxidase from *Nitrosomonas europae*. *Biochim. Biophys. Acta.* 827:320–326.

Fujiwara, T., T. Yamanaka, and Y. Fukumori. 1995. The amino acid sequence of *Nitrosomonas europaea* cytochrome *c*-552. *Curr. Microbiol.* 31:1–4.

Leitch, F. A., G. R. Moore, and G. W. Pettigrew. 1984. Structural basis for the variation of pH-dependent redox potentials of *Pseudomonas* cytochromes *c*-551. *Biochemistry.* 23:1831–1838.

Logan, M. S. P. 1991. Hydroxylamine oxidoreductase from *Nitrosomonas europaea*: inhibition by cyanide and suicide substrates akyl and aryl hydrazines. Ph.D. thesis. Department of Genetics and Cell Biology, University of Minnesota, St. Paul, MN.

Matsuura, Y., T. Takano, and R. E. Dickerson. 1982. Structure of cytochrome *c*551 from *P. aeruginosa* refined at 1.6 Å resolution, and comparison of the two redox forms. *J. Mol. Biol.* 156:389–409.

McTavish, H. E., J. Fuchs, and A. B. Hooper. 1993a. Sequence of the gene for ammonia monooxygenase of *Nitrosomonas europae*. *J. Bacteriol.* 175:2436–2444.

McTavish, H. E., F. LaQuier, D. Arciero, M. Logan, J. Fuchs, and A. B. Hooper. 1993b. Multiple copies for genes for electron transport in the bacterium *Nitrosomonas europaea*. *J. Bacteriol.* 175:2445–2447.

Miller, D. J., and D. J. D. Nicholas. 1986. N-Terminal amino acid sequence of cytochrome *c*-552 from *Nitrosomonas europae*. *Biochem. Int.* 12: 176–172.

Nilges, M., G. M. Clore, and A. M. Gronenberg. 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.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Timkovich, R. 1979. Cytochrome *c*: the architecture of a protein-porphyrin complex. *In* The Porphyrins. D. Dolphin, editor. Academic Press, New York. 241–294.

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., M. Cai, D. M. Arciero, and A. B. Hooper. 1994. Characteristics of the paramagnetic ^{1}H NMR spectra of the ferricytochrome *c*-551 family. *Eur. J. Biochem.* 226:159–168.

Timkovich, R., M. Cai, and D. W. Dixon. 1988. Electron self-exchange in *Pseudomonas* cytochromes. *Biochem. Biophys. Res. Commun.* 150: 1044–1050.

Timkovich, R., and M. S. Cork. 1984. Proton NMR spectroscopy of cytochrome *c*-554 from *Alcaligenes faecalis*. *Biochemistry.* 23:851–860.

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

Yamanaka, T. 1967. Cytochrome *c* and evolution. *Nature.* 213:1183–1186.

Yamanaka, T., and M. Shinra. 1974. Cytochrome *c*-552 and cytochrome *c*-554 derived from *Nitrosomonas europaea*. *J. Biochem. (Tokyo).* 75: 1265–1273.

Yamazaki, T., Y. Fukumori, and T. Yamanaka. 1988. Catalytic properties of cytochrome *c* oxidase purified from *Nitrosomonas europae*. *J. Biochem.* 103:499–503.