Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Hyperstability and crystal structure of cytochrome c_{555} from hyperthermophilic *Aquifex aeolicus*

Marii Obuchi,[a]‡ Kazuki Kawahara,[b]‡ Daisuke Motooka,[c] Shota Nakamura,[b] Masaru Yamanaka,[a] Taku Takeda,[a] Susumu Uchiyama,[b] Yuji Kobayashi,[c] Tadayasu Ohkubo[b] and Yoshihiro Sambongi[a]*

[a]Hiroshima University, Japan, [b]Osaka University, Japan, and [c]Osaka University of Pharmaceutical Science, Japan

‡ These authors contributed equally to this work.

Correspondence e-mail: sambongi@hiroshima-u.ac.jp

Received 27 February 2009
Accepted 8 May 2009

PDB Reference: AA c_{555}, 2zxy, r2zxysf.

In order to elucidate the relationship between the stability and the structure of the monohaem cytochrome c_{555} (AA c_{555}) from the hyperthermophilic bacterium *Aquifex aeolicus*, chemical denaturation and crystal structure determination were carried out. AA c_{555} exhibited higher stability than the thermophilic *Hydrogenobacter thermophilus* cytochrome c_{552} (HT c_{552}), which is one of the most stable cytochromes *c*. The three-dimensional crystal structure of AA c_{555}, which was determined using the multiple anomalous dispersion technique at 1.15 Å resolution, included a unique 14-residue extra helix, while the side-chain interactions of several amino-acid residues responsible for the stability of HT c_{552} were conserved in AA c_{555}. The side chain of the Met61 residue in the extra helix was aligned towards the haem, forming a coordination bond between the Met S and haem Fe atoms. In other cytochromes *c* the corresponding regions always form Ω loops which also include the haem-liganding Met residue and are known to be involved in the initial step in cytochrome *c* denaturation. The formation of the extra helix in AA c_{555} results in the highest helix content, 59.8%, among the monohaem cytochromes *c*. The extra helix should mainly contribute to the hyperstability of AA c_{555} and is presumed to be a novel strategy of cytochromes *c* for adaptation to a hyperthermophilic environment.

1. Introduction

Organisms that have adapted to high temperatures are called hyperthermophiles (growth temperatures over 353 K) and thermophiles (growth temperatures 318–348 K). The hyperthermophilic *Aquifex aeolicus*, which grows optimally at 358 K, belongs to the phylum Aquificae (Deckert *et al.*, 1998). Thermophilic genera such as *Hydrogenobacter* and *Hydrogenobaculum*, which grow optimally at around 345 K, are also included in the phylum Aquificae (Reysenbach, 2001). These hyperthermophilic and thermophilic Aquificae are chemolithotrophic hydrogen oxidizers with respiratory chains into which electrons are derived from H_2 molecules (Reysenbach, 2001). In the respiratory electron-transport chains of *A. aeolicus* and *Hydrogenobacter thermophilus*, ubiquitous monohaem soluble cytochromes *c* function as electron carriers on the outer surface of the cytoplasmic membrane. The 87-residue monohaem cytochrome c_{555} (AA c_{555}) of *A. aeolicus* and the 80-residue monohaem cytochrome c_{552} (HT c_{552}) of *H. thermophilus* are presumed to be functionally equivalent because they exhibit similar redox-potential values (Aubert *et al.*, 2001; Terui *et al.*, 2003).

In the last decade, we have investigated the structures and stabilities of ubiquitous cytochromes *c* such as HT c_{552} and its homologous counterparts from the moderately thermophilic *Hydrogenophilus thermoluteolus*, which grows optimally at 325 K, and the mesophilic *Pseudomonas aeruginosa*, which grows at 310 K (Sambongi *et al.*, 2002). HT c_{552} has been found to be more stable than *H. thermoluteolus* cytochrome c_{552} (PH c_{552}) and *P. aeruginosa* cytochrome c_{551} (PA c_{551}) towards thermal and chemical denaturation, although these three proteins exhibit greater than ∼50% sequence identity and similar main-chain structures, with root-mean-square deviation (r.m.s.d.) values of ∼1 Å (Hasegawa *et al.*, 1999, 2000; Uchiyama *et al.*, 2002; Nakamura *et al.*, 2006; Oikawa *et al.*, 2005; Hakamada *et al.*, 2008). The order of stability of these three proteins corresponds to that of the optimal growth temperatures of the respective bacteria. From the results of mutagenesis studies on these cytochromes *c*, we concluded that the higher stability of HT c_{552} mainly arises from a few specific side-chain interactions (Oikawa *et al.*, 2005).

Since AA c_{555} originates from a hyperthermophilic bacterium, we expected it to exhibit higher stability than the thermophile-originated HT c_{552}, which is one of the most stable cytochromes *c*. In a previous thermal denaturation experiment on AA c_{555} up to 373 K, the complete protein-denaturation profile could not be confirmed because of its extraordinary thermal stability (Aubert *et al.*, 2001). Sequence comparison revealed that AA c_{555} exhibits more than ∼30% sequence identity or ∼50% similarity to HT c_{552}, PH c_{552} and PA c_{551}, indicating that they possess similar overall molecular architectures with different stabilities. Therefore, AA c_{555} may be an excellent model for obtaining clues to protein hyperstability at the atomic level.

The hyperstability of AA c_{555} has been proposed based on sequence comparison and homology modelling by Aubert *et al.* (2001). They pointed out that AA c_{555} has a stretch of amino acids (residues 66–70) on the C-terminal side of Met61, an axial ligand to the haem (Fig. 1). This stretch is not found in the sequences of other cytochromes *c* including HT c_{552}. Aubert *et al.* (2001) predicted that this stretch forms a loop structure and is correlated to the hyperstability of AA c_{555}. However, complete denaturation experiments and three-dimensional structure determination have not yet been performed for AA c_{555}.

Here, we show the hyperstability of AA c_{555} through complete denaturation experiments and determine the three-dimensional crystal structure of AA c_{555} by the multi-wavelength anomalous dispersion (MAD) phasing method with refinement to a resolution of 1.15 Å. The AA c_{555} structure is compared with those of other monohaem cytochromes *c*, especially that of HT c_{552}. Based on the results, the stabilization mechanism of AA c_{555} is discussed.

2. Experimental procedures

2.1. Protein preparation

Overexpression and correct post-translational modification of AA c_{555} in *Escherichia coli* were achieved by the methods used previously in the case of HT c_{552}, where the signal sequence was changed to that of PA c_{551} and the cytochrome *c* maturation (*ccm*) genes were co-expressed (Oikawa *et al.*, 2005). AA c_{555} was overexpressed with co-transformed pEC86 carrying *ccm* genes in *E. coli* RI90 (*dsbA*$^-$ strain) cells and purified from the periplasmic protein fraction as described previously (Oikawa *et al.*, 2005).

Initially, the protein fraction including AA c_{555} was passed through a DEAE column using 10 m*M* Tris–HCl pH 8.0 and further purification was carried out by Hi-Trap SP column chromatography (GE Healthcare, USA) with a sodium chloride concentration gradient (0–500 m*M*) in 25 m*M* 2-morpholinoethanesulfonic acid (MES) buffer pH 6.0. Finally, gel-filtration chromatography was performed on a Superdex 75 column (GE Healthcare, USA) equilibrated and eluted with 25 m*M* MES buffer pH 6.0. The purity of the eluted AA c_{555} was confirmed by SDS–PAGE. The heterologously expressed AA c_{555} exhibited the same spectroscopic properties as those of the authentic protein (Aubert *et al.*, 2001).

2.2. Guanidine thiocyanate-induced denaturation

The stability of air-oxidized AA c_{555} and HT c_{552} towards chemical denaturation was assessed by means of guanidine thiocyanate-induced denaturation experiments. The guanidine thiocyanate (GdnSCN) concentration was determined using the refractive index measured on a refractometer (Reichert Inc., USA). As GdnSCN interferes with the far-UV circular-dichroism (CD) spectrum, which is usually used for moni-

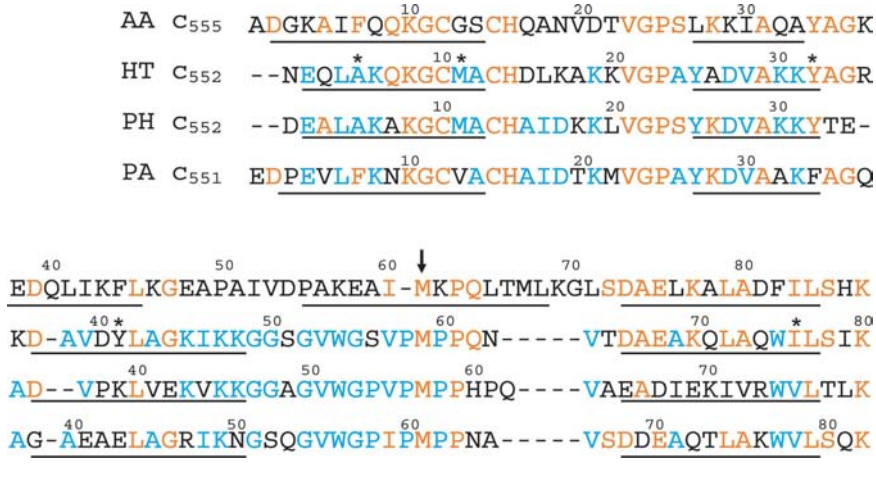

Figure 1
Sequence alignment of AA c_{555}, HT c_{552}, PH c_{552} and PA c_{551}. Residue numbers are indicated above each sequence. Helical structures assigned on the basis of three-dimensional structural analyses are underlined (those for AA c_{555} were determined in this study). Conserved residues in AA c_{555} and the other cytochromes *c* are highlighted in orange, while residues that are conserved apart from in AA c_{555} are highlighted in blue. The conserved axially haem-liganded Met residues are indicated by an arrow. The five amino-acid residues indicated by asterisks in the HT c_{552} sequence are responsible for its high stability.

toring protein denaturation, the absorption spectrum around a specific peak of air-oxidized cytochrome *c* (Soret band) was measured. The absorption spectra (200–800 nm) of protein samples were obtained using a 1 cm path-length cuvette in a Jasco V-530 spectrophotometer (Japan) at 298 K. In order to equilibrate the proteins with GdnSCN, 10 µ*M* protein solutions were incubated in 20 m*M* sodium phosphate buffer pH 7.0 containing various concentrations of GdnSCN at 298 K for 2 h before the stability measurements. An incubation time of 2 h was sufficient for reaching equilibration; 1–6 h incubation provided essentially the same results. The reversibility of protein denaturation was verified by simple dilution of a tested sample and measurement of the spectra.

Nonlinear least-squares fitting of the data obtained after stability measurements was performed according to the method described previously (Oikawa *et al.*, 2005). Data points were corrected for the slope of the baselines for the native and denatured forms and normalized to calculate the fraction of protein denatured. The fraction denatured was plotted as a function of the GdnSCN concentration to estimate the midpoint of GdnSCN denaturation (C_m).

Table 1
Data-collection, phasing and refinement statistics.

Values in parentheses are for the highest resolution shell.

		MAD		
	Native	Fe peak	Fe edge	Fe remote
Data-collection statistics				
Temperature (K)	100			
Wavelength (Å)	0.70000	1.73786	1.74011	1.69260
Resolution range (Å)	50.0–1.15 (1.19–1.15)	50.0–2.86 (2.96–2.86)	50.0–2.87 (2.97–2.87)	50.0–2.79 (2.89–2.79)
R_{merge}†	0.047	0.040	0.042	0.055
Unique reflections	25402	1725	1705	1859
Redundancy	3.71	3.18	3.51	3.56
Completeness (%)	97 (89.4)	97.2 (92.2)	99.2 (96.4)	99.7 (97.3)
$\langle I\rangle/\langle\sigma(I)\rangle$ in highest resolution shell	10.01	18.92	27.94	32.70
Space group	*C*2			
Unit-cell parameters				
a (Å)	48.529	48.562	48.570	48.586
b (Å)	32.993	32.976	32.980	32.988
c (Å)	49.668	49.770	49.798	49.821
Mosaicity	0.180	0.903	0.905	0.924
Summary of phasing statistics				
Figure of merit‡	NA	0.459		
No. of sites	1 Fe	1 Fe		
Crystallographic analysis and refinement statistics				
Refinement resolution range (Å)	15.92–1.15			
R factor§	0.127			
R_{free}¶	0.164			
Overall temperature factor (Å^2)	10.410			
R.m.s.d. from ideal				
Bonds (Å)	0.005			
Angles (°)	0.706			
Ramachandran plot				
Most favoured regions (%)	89.2			
Additional allowed regions (%)	8.1			
Generously allowed regions (%)	2.7			
Disallowed regions (%)	0.0			

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the average intensity of reflection *hkl* and symmetry-related reflections. ‡ Figure of merit (FOM) is the mean of the cosine of error in the phase angle: FOM = $|F(hkl)_{best}|/|F(hkl)|$, where $F(hkl)_{best} = \sum_{hkl} P(\alpha)F_{hkl}(\alpha)/\sum_{hkl} P(\alpha)$. § *R* factor = $\sum_{hkl} ||F_o| - |F_c||/\sum_{hkl} |F_o|$, where F_o and F_c are the observed and calculated structure-factor amplitudes, respectively. ¶ $R_{free} = R_{factor}$ calculated using 5% of the reflection data chosen randomly and set aside from the start of refinement.

2.3. Crystallization and data collection

A solution of purified AA c_{555} was concentrated to 8.75 mg ml^{-1} by ultrafiltration (Centriplus YM-3, Millipore, USA) and then filter-sterilized (0.1 µm Ultrafree-MC, Millipore, USA). Initial crystals were obtained by the hanging-drop vapour-diffusion method using Crystal Screen kits I and II (Hampton Research, USA) at 277 K. Small plate-like crystals of AA c_{555} grew to typical dimensions of 0.1 × 0.1 × 0.2 mm within a week with 30%(*w*/*v*) polyethylene glycol 4000, 0.2 *M* ammonium acetate and 0.1 *M* trisodium citrate dihydrate pH 5.6. To improve the quality of the crystals, optimization of these conditions was carried out using the sitting-drop vapour-diffusion method and the microseeding technique. The best crystals for crystallographic analysis were obtained at 277 K from drops consisting of 2 µl protein solution (8.75 mg ml^{-1}) and 2 µl reservoir solution containing 39%(*w*/*v*) polyethylene glycol 4000, 0.8 *M* ammonium acetate and 0.1 *M* trisodium citrate dihydrate pH 5.6. The crystal morphology changed to a hexagonal shape with typical dimensions of approximately 0.2 × 0.3 × 0.2 mm.

X-ray diffraction data were collected at 100 K from AA c_{555} crystals mounted on a cryoloop (0.3–0.4 mm, Hampton Research, USA) on SPring-8 BL38B1 (Hyogo, Japan) and were processed using *DENZO* and *SCALEPACK* from the *HKL*-2000 package (Otwinowski & Minor, 1997). Further data processing was performed with programs from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). After measurement of an X-ray fluorescence spectrum around the Fe *K* edge, a total of three wavelengths were selected for MAD data collection (Table 1). The AA c_{555} crystals diffracted to a resolution of 1.15 Å for the native data and 2.87 Å for the MAD data at the Fe edge. The diffraction spots from the unit cell were indexed in space group *C*2, with unit-cell parameters *a* = 48.529, *b* = 32.993, *c* = 49.668 Å. Assuming the presence of one 9.8 kDa molecule in the asymmetric unit, the Matthews coefficient was 1.84 Å^3 Da^{-1} and the solvent content was 36.7% (Matthews, 1968). The data-collection and refinement statistics are presented in Table 1.

2.4. Structure determination and refinement

Structure determination was performed by the MAD technique using the anomalous dispersion of the intrinsic

haem Fe atom (Hendrickson, 1991). Phase calculations were performed using three programs, *SHELXC*, *SHELXD* and *SHELXE*, from the *SHELX* suite (Sheldrick, 2008). After data preparation with *SHELXC*, the Fe-atom position in the asymmetric unit was clearly found using *SHELXD*. Subsequent phase calculations and electron-density improvement were carried out using *SHELXE* (Table 1). The resulting electron density allowed automatic model building using the program *ARP*/*wARP* (Morris *et al.*, 2002). 69 of 87 residues were automatically built and most of the side chains were traced with *ARP*/*wARP*. The missing parts were manually rebuilt with the molecular-graphics program *Coot* (Emsley & Cowtan, 2004) and refined with the program *REFMAC*5 against the native high-resolution data set (Murshudov *et al.*, 1999). During the final stages of refinement and model building, 145 water molecules were added to the model by manual inspection of $2F_o - F_c$ electron-density maps. The final model of AA c_{555} consisted of 711 non-H protein atoms, 43 haem atoms and 145 water molecules. The R factor and R_{free} of the final model were 0.127 and 0.164, respectively (Table 1). The correctness of the structure was assured with *PROCHECK* (Laskowski *et al.*, 1993). A total of 89.2% of the residues occupied most favoured regions and 10.8% of the residues occupied additional allowed regions on a Ramachandran plot, with no residues in the disallowed regions. Secondary-structure characterization was carried out using the program *DSSP* (Kabsch & Sander, 1983). Figures representing the protein structure were made using *PyMOL* (DeLano, 2002).

Table 2
Structural comparisons among AA c_{555} and other monohaem cytochromes *c*.

PDB code	Description	Content of α- and 3_{10}-helices (%)	Z score†
2zxy	*Aquifex aeolicus* cytochrome c_{555} (AA c_{555})	59.8	–
1dvh	*Desulfovibrio vulgaris* cytochrome c_{553} (DV c_{553})	54.4	6.1
1cno	*Pseudomonas nautica* cytochrome c_{552}	53.3	6.9
1a56	*Nitrosomonas europaea* cytochrome c_{552}	53.1	6.3
1h1o	*Acidithiobacillus ferrooxidans* cytochrome c_4	52.8	3.9
1c75	*Bacillus pasteurii* cytochrome c_{553}	52.1	6.7
1gks	*Ectothiorhodospira halophila* cytochrome c_{551}	51.3	5.4
1cch	*Pseudomonas stutzeri* cytochrome c_{551} (PS c_{551})	51.2	9.0
2d0s	*Hydrogenophilus thermoluteolus* cytochrome c_{552} (PH c_{552})	50.6	10.2
1ynr	*Hydrogenobacter thermophilus* cytochrome c_{552} (HT c_{552})	50.5	9.8
351c	*Pseudomonas aeruginosa* cytochrome c_{551} (PA c_{551})	50.0	9.8
1i8o	*Rhodopseudomonas palustris* cytochrome c_2	47.4	5.8
1cc5	*Azotobacter vinelandii* cytochrome c_5	47.0	5.5
1c52	*Thermus thermophilus* cytochrome c_{552}	45.8	5.4
1jdl	*Rhodospirillum centenum* cytochrome c_2	45.8	5.7
1f1c	*Arthrospira maxima* cytochrome c_{549}	45.6	5.0
1yea	Yeast iso-2-cytochrome *c*	45.5	5.8
1qn2	*Methylobacterium extorquens* cytochrome c_2	44.8	6.9
1kx2	*Shewanella putrefaciens* cytochrome *c*	44.4	5.6
2aiu	Mouse testicular cytochrome *c*	43.3	6.4
3c2c	*Rhodospirillum rubrum* cytochrome c_2	42.9	5.9
1ccr	Rice cytochrome *c*	42.3	5.9
1j3s	Human cytochrome *c*	42.3	5.2
1ycc	Yeast iso-1-cytochrome *c*	40.7	6.1
2b4z	Bovine heart cytochrome *c*	40.4	6.4
1c7m	*Paracoccus denitrificans* cytochrome c_{552}	39.0	5.8
1hro	*Rhodopila globiformis* cytochrome c_2	38.9	6.5
1cot	*Paracoccus denitrificans* cytochrome c_2	38.8	4.9
1co6	*Rhodopseudomonas viridis* cytochrome c_2	38.3	6.0
155c	*Paracoccus denitrificans* cytochrome c_{550}	26.1	4.2

† As defined by Holm & Sander (1998).

2.5. Secondary-structure assessment and three-dimensional structural alignment

A total of 50 PDB entities which were classified as bacterial and mitochondrial monohaem cytochromes *c* were obtained from the SCOP database. To avoid unnecessary data duplication, 29 coordinate sets were then further extracted from the data set according to the following criteria: (i) they contained His/Met coordinated bonds, (ii) they formed a nonredundant set of amino-acid sequences and (iii) they were cytochromes *c* of the respiratory electron-transport chain. The *DSSP* program was used for assignment of secondary structures to the selected polypeptides and all pairwise structural comparisons were carried out with the *DaliLite* program (*DALI*; Holm & Park, 2000). The resulting total helix contents and Z scores are listed in Table 2.

3. Results and discussion

3.1. GdnSCN-induced denaturation

Our preliminary experiments showed that air-oxidized AA c_{555} was so stable that its chemical denaturation could not be monitored by measuring the 222 nm ellipticity of CD spectra even at pH 1.0 with adjustment with HCl or at pH 5.0 with 6 *M* guanidine hydrochloride (GdnHCl). Therefore, in order to achieve complete denaturation of AA c_{555}, GdnSCN, a stronger chemical denaturant than GdnHCl, was used (Mason *et al.*, 2003).

The absorption spectra (360–440 nm) of AA c_{555} in the presence of various concentrations of GdnSCN at pH 7.0 and 298 K are shown in Fig. 2. This shows that the Soret absorption peak became blue-shifted with an increase in GdnSCN concentration. The existence of an isosbestic point at 409 nm indicates that the GdnSCN concentration dependency of the spectral change can be described as a simple two-state process.

Difference spectra were obtained by subtraction of the AA c_{555} spectra without GdnSCN from those in the presence of various concentrations of GdnSCN. A positive peak at 400 nm and a negative peak at 416 nm were observed (Fig. 2). The GdnSCN-induced denaturation of the AA c_{555} protein was sharply reflected in the GdnSCN concentration dependence of the 400 nm peak. Therefore, the fraction of denatured protein as a function of the GdnSCN concentration was calculated from the raw data obtained for the 400 nm peak (Fig. 2). For

comparison, the results for the air-oxidized HT c_{552} protein (Takeda *et al.*, 2009) under the same experimental conditions are also presented. The S-shaped denaturation curves can be adequately fitted to a two-state model of protein denaturation. The GdnSCN concentration values at the midpoint of denaturation (C_m) of AA c_{555} and HT c_{552} were 3.09 ± 0.02 and 1.88 ± 0.01 *M*, respectively.

The above stability comparison indicated that AA c_{555} exhibits higher stability than HT c_{552} during GdnSCN-induced denaturation. As HT c_{552} has been recognized as one of the most stable monohaem cytochromes *c* (Sanbongi *et al.*, 1989), AA c_{555} can be regarded as a 'hyperstable' cytochrome *c*. Identical GdnSCN-induced denaturation experiments showed that the C_m values for air-oxidized PH c_{552} and PA c_{551} were 1.12 and 0.63 *M*, respectively (Takeda *et al.*, 2009). Therefore, the order of stability of these four cytochromes *c* corresponds to that of the optimal growth temperatures of the respective bacteria (*A. aeolicus*, 358 K; *H. thermophilus*, 345 K; *H. thermoluteolus*, 325 K; *P. aeruginosa*, 310 K).

3.2. Structure determination of AA c_{555} and haem geometry

To elucidate the structural basis of the AA c_{555} hyperstability in detail, we determined its three-dimensional structure by X-ray crystallography. Despite a number of trials to determine the initial phase by the molecular-replacement method using a variety of search models, we failed to obtain clear solutions that produced a fully traceable electron-density map. Therefore, the initial AA c_{555} structure was determined by the MAD technique using the *K*-edge anomalous dispersion effects of the Fe atom present in the haem and structure refinement was carried out to a resolution of 1.15 Å. The electron-density map calculated from the MAD phasing (using reflections to 2.87 Å resolution) and solvent flattening was well defined, so we could position the haem and construct the entire atomic model between residues 1 and 86 (the C-terminal residue 87 was omitted from the final model owing to disordered electron density).

An overall view of the crystal structure of AA c_{555} is presented in Fig. 3. The secondary structure was assigned with the program *DSSP* and the model consists of five α-helices, *i.e.* αA (residues 3–9), αB (residues 27–33), αC (residues 38–45), αD (residues 58–61) and αE (residues 73–84), and three 3_{10} helices, *i.e.* 3_{10}A (residues 12–14), 3_{10}B (residues 55–57) and 3_{10}C (residues 62–68).

A single haem is buried in a hydrophobic pocket in the AA c_{555} structure (Fig. 3). A representative part of the haem with the surrounding side chains is depicted in Fig. 3. The haem is covalently bound to Cys12 and Cys15 through thioether linkages and the Fe atom also lies in the plane of the haem, axially liganded to His16 and Met61. The haem is further held in place by hydrogen bonds between haem 17-propionate O2A and Tyr34 OH, 17-propionate O2A and Ile52 N, 17-propionate O1A and Ile52 N, 17-propionate O1A and Val53 N and 13-propionate O1D and Asp54 N. Unusually, the two haem propionate groups are hydrogen bonded to each other. Owing to such a unique hydrogen bond between the two propionates in the AA c_{555} haem, the 13-propionate group is bent unlike those of other monohaem cytochromes *c*, which have straight propionate groups.

3.3. Structural comparison of AA c_{555}, HT c_{552}, PH c_{552} and PA c_{551}

The newly determined crystal structure of AA c_{555} was compared with those of HT c_{552} (PDB code 1ynr), PH c_{552} (PDB code 2d0s) and PA c_{551} (PDB code 351c). The overall

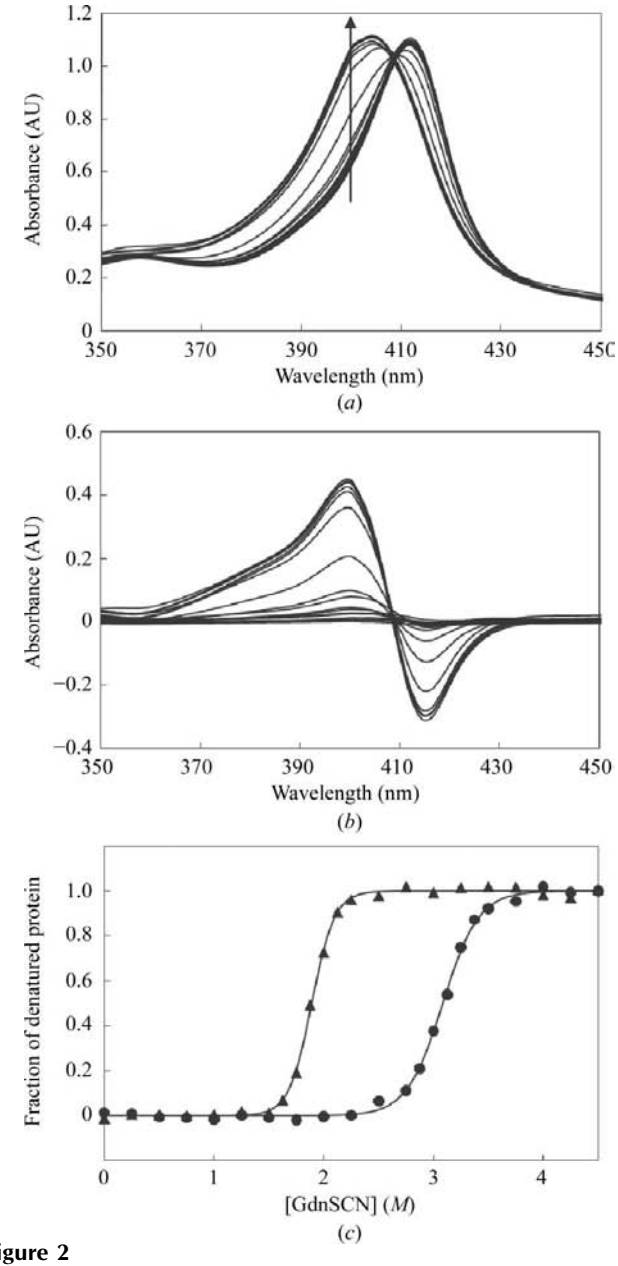

Figure 2
GdnSCN-induced denaturation of air-oxidized AA c_{555} and HT c_{552}. (*a*) Visible absorption spectra of AA c_{555} in the presence of various concentrations of GdnSCN. The arrow indicates the increase in GdnSCN concentration from 0 *M*. (*b*) Difference spectra of AA c_{555} in the presence of various concentrations of GdnSCN after subtraction of the data without GdnSCN. (*c*) Fractions of denatured AA c_{555} (closed circles) and HT c_{552} (closed triangles) plotted as a function of the GdnSCN concentration. The curve fitted for each experiment is also shown.

structures of AA c_{555} and the other three cytochromes *c* (more than ∼30% and ∼50% sequence identity and similarity, respectively) could be superimposed (Fig. 4) with average backbone r.m.s.d. values of 1.82 Å for AA c_{555}–HT c_{552}, 1.67 Å for AA c_{555}–PH c_{552} and 1.74 Å for AA c_{555}–PA c_{552}. These values are considerably larger than those for the main-chain folds of the HT c_{552}, PH c_{552} and PA c_{551} structures, which have r.m.s.d. values of ∼1 Å (Nakamura *et al.*, 2006).

A major structural difference between AA c_{555} and the other three cytochromes *c* is in the region connecting the αC helix and αE helix (*i.e.* residues 47–72 for AA c_{555}, residues 49–65 for HT c_{552}, residues 47–64 for PH c_{552} and residues 51–67 for PA c_{551}). In the case of the AA c_{555} structure, there are three helices (3_{10}B, αD and 3_{10}C) forming a consecutively long helix flanked by short loops in this region (Fig. 4). The axial haem-liganding Met61 residue is located in the middle of this helix. In contrast, HT c_{552}, PH c_{552} and PA c_{551} do not have any helix structure and have a long loop structure in the corresponding region (Fig. 4) which contains an equivalent Met residue. We call this 3_{10}B–αD–3_{10}C helix in AA c_{555} the 'extra helix'. From measurement of the far-UV CD ellipticity at 222 nm in 20 m*M* sodium acetate buffer pH 5.0, the helical content of AA c_{555} was estimated to be ∼60%, which was significantly higher than those of the other three cytochromes *c* (MY, unpublished observation). These indicate that the formation of the extra helix is an intrinsic feature of AA c_{555} and is not an artifact of the crystallization conditions.

Apart from the extra helix and its corresponding regions in these four cytochromes *c*, the other parts superimpose quite well (black-line models in Fig. 4). The average r.m.s.d. values for the C^{α} atoms of the other parts are 1.59 Å for AA c_{555}–HT c_{552}, 1.45 Å for AA c_{555}–PH c_{552} and 1.42 Å for AA c_{555}–PA c_{552}.

With these considerations, we first focus on the side-chain interactions within the parts other than the extra helix and adjacent loops in AA c_{555} to discern the structural basis for its

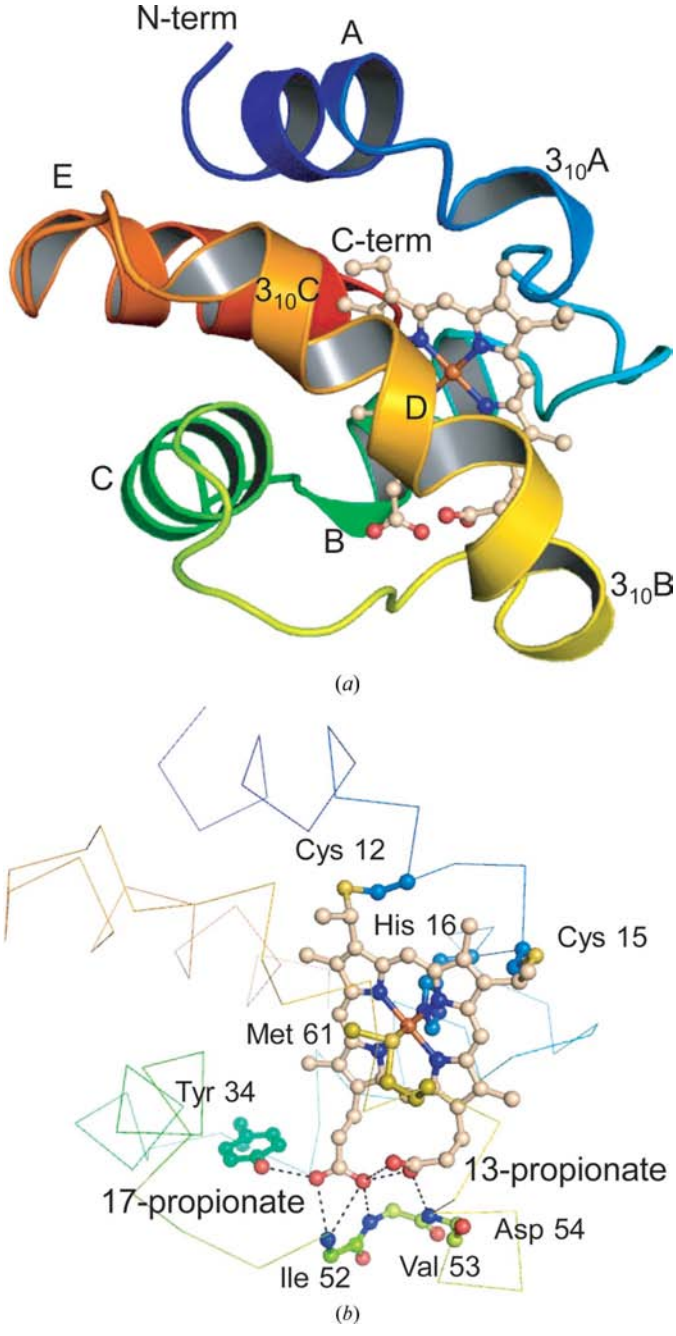

Figure 3
Three-dimensional structure of AA c_{555} and its haem geometry. (*a*) Overall structure of AA c_{555}. The main chain and haem are presented as ribbon and ball-and-stick models, respectively. The N- and C-termini as well as helix regions are indicated. (*b*) Haem geometry. The main chain and haem are presented as line and ball-and-stick models, respectively. Relevant amino-acid residues that are chemically in contact with the haem are indicated, together with the haem propionates. Relevant hydrogen bonds are shown by dashed lines.

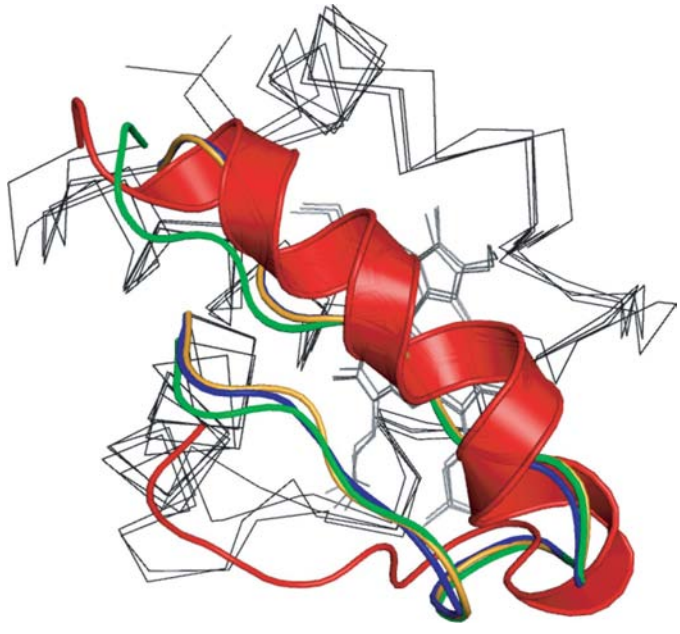

Figure 4
Three-dimensional structural comparison of AA c_{555}, HT c_{552}, PH c_{552} and PA c_{551}. The extra helix and adjacent loops in AA c_{555} are presented as a ribbon model in red. The corresponding regions in HT c_{552}, PH c_{552} and PA c_{551} are coloured orange, green and blue, respectively. The other parts of polypeptides and haems are presented as a line model (black).

hyperstability in comparison with HT c_{552}. Previous structure and mutagenesis studies on HT c_{552}, PH c_{552} and PA c_{551} revealed that only five amino-acid residues of HT c_{552} (Ala5, Met11, Tyr32, Tyr41 and Ile76; indicated by asterisks in Fig. 1) distributed in three spatially separate regions designated regions I, II and III (Fig. 5) are responsible for the higher stability compared with PA c_{551} (Hasegawa *et al.*, 1999, 2000; Uchiyama *et al.*, 2002; Nakamura *et al.*, 2006; Oikawa *et al.*, 2005; Hakamada *et al.*, 2008). The residues corresponding to these in AA c_{555} are Phe7, Gly13, Tyr34, Phe44 and Ile83 (Fig. 1), respectively, which occupy the same positions in its three-dimensional structure as in that of HT c_{552} (Fig. 5).

Region I of HT c_{552} is stabilized through tight van der Waals interactions of the side chains of Ala5 and Met11 (Fig. 5). The corresponding residues in PA c_{551} are Phe7 and Val13; the Phe side chain occupies a rather large space inside the molecule so that Val protrudes into the external solvent, which has been concluded to be a destabilizing factor in PA c_{551} compared with HT c_{552} (Hasegawa *et al.*, 2000). As in the case of PA c_{551}, Phe7 occupies the corresponding position in AA c_{555}. However, Gly, with the smallest side chain, occupies position 13 in AA c_{555} and thus Phe7 can be located in the middle of the void space created by Leu27, Phe82 and itself (Fig. 5), which leads to tight hydrophobic packing in region I of AA c_{555}.

Region II of HT c_{552} is stabilized by hydrogen bonds between Tyr32 (OH) and haem 17-propionate (O1A) and between Tyr41 (OH) and haem 17-propionate (O2A) (Fig. 5). These hydrogen bonds do not exist in PA c_{551}, the corresponding residues being Phe34 and Glu43 (Hasegawa *et al.*, 2000). In AA c_{555}, the corresponding Tyr34 residue forms a hydrogen bond to the haem 17-propionate, while Phe44 without an OH group cannot form the same bond. However, Phe44 extends further in the AA c_{555} molecule than the corresponding Tyr41 residue does in HT c_{552}. Thus, Phe44 in AA c_{555} contributes to the tighter packing in region II.

The two corresponding residues Ile83 in AA c_{555} and Ile76 in HT c_{552} are in hydrophobic contact with the respective haem in region III (Fig. 5). In PA c_{551} there is a gap between the corresponding Val78 residue and the haem (Nakamura *et al.*, 2006). This gap is filled by the methyl groups of Ile83 in AA c_{555} and Ile76 in HT c_{552}. This difference in packing density should naturally enhance hydrophobic interactions and thus the Ile residue in AA c_{555} contributes to the stability in a similar manner to the case of HT c_{552}.

Together, these results suggest that the contributions of five amino-acid residues in regions I, II and III responsible for the stability of HT c_{552} are basically the same in the case of AA c_{555}.

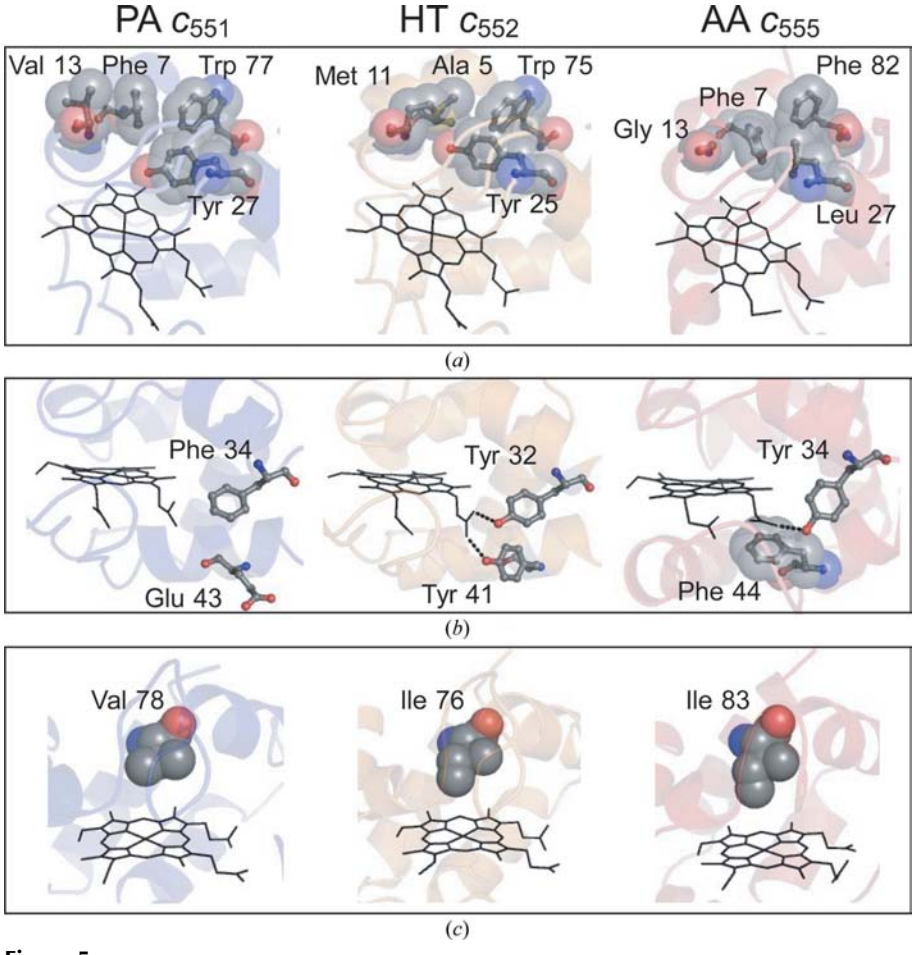

Figure 5
Five amino-acid residues responsible for the stability of HT c_{552}. (*a*) Region I. (*b*) Region II. Relevant hydrogen bonds are shown by dashed lines. (*c*) Region III. In all panels, the side chains of relevant amino-acid residues are indicated and backbones and haems are presented as ribbon and line models, respectively.

3.4. Contribution of the extra helix and adjacent loops to the hyperstability

The presence of the extra helix and adjacent loops in AA c_{555} should lead to its hyperstability as AA c_{555} exhibits almost the same strategy for stabilization as that of HT c_{552} in the other parts of the structure, as discussed above. On sequence comparison, Aubert *et al.* predicted that one region (residues 66–70) forms a loop structure which is responsible for the hyperstability (Aubert *et al.*, 2001). However, our present results clearly demonstrate that this region (residues 66–70) is part of the extra helix. This extra helix is an important factor in the hyperstability as ten backbone–backbone hydrogen bonds are formed within the extra helix (Fig. 6). In addition to these ten bonds, three backbone–backbone hydrogen bonds are formed in loops adjacent to the extra helix (Fig. 6). In contrast to these 13 backbone–backbone hydrogen bonds in the extra helix and adjacent loop in AA c_{555}, the corresponding long loop structure in HT c_{552} (Gly49–Thr65)

contains only four backbone–backbone hydrogen bonds (Fig. 6) (Travaglini-Allocatelli *et al.*, 2005).

The loop (Lys46–Asp54) preceding the extra helix in AA c_{555} (Fig. 6) contains the unique Pro50 residue (Fig. 1), which generally constrains the loop conformation. Although Gly47 exists in the loop in AA c_{555}, the corresponding HT c_{552} loop contains a larger number of Gly residues, at positions 49, 50, 52 and 55 (Figs. 1 and 6), which cause conformational flexibility compared with AA c_{555}. Consistently, the loop (Lys46–Asp54) in AA c_{555} limits the conformational flexibility as reflected by the crystallographic *B* factors (Fig. 6) and the corresponding loop in HT c_{552} is considerably more dynamic in the native state. In most other monohaem cytochromes *c*, including HT c_{552}, the corresponding regions usually form Ω-loop structures, which facilitate coordination-bond formation between the axially liganded Met S and haem Fe atoms. Hydrogen-exchange experiments indicated that the Ω loop is the most susceptible part in cytochrome *c* denaturation (Bai *et al.*, 1995). Therefore, by minimizing conformational flexibility through the hydrogen bonds and the existence of the Pro residue, the extra helix and adjacent loops in AA c_{555} contribute to protection against protein denaturation unlike in other cytochromes *c*.

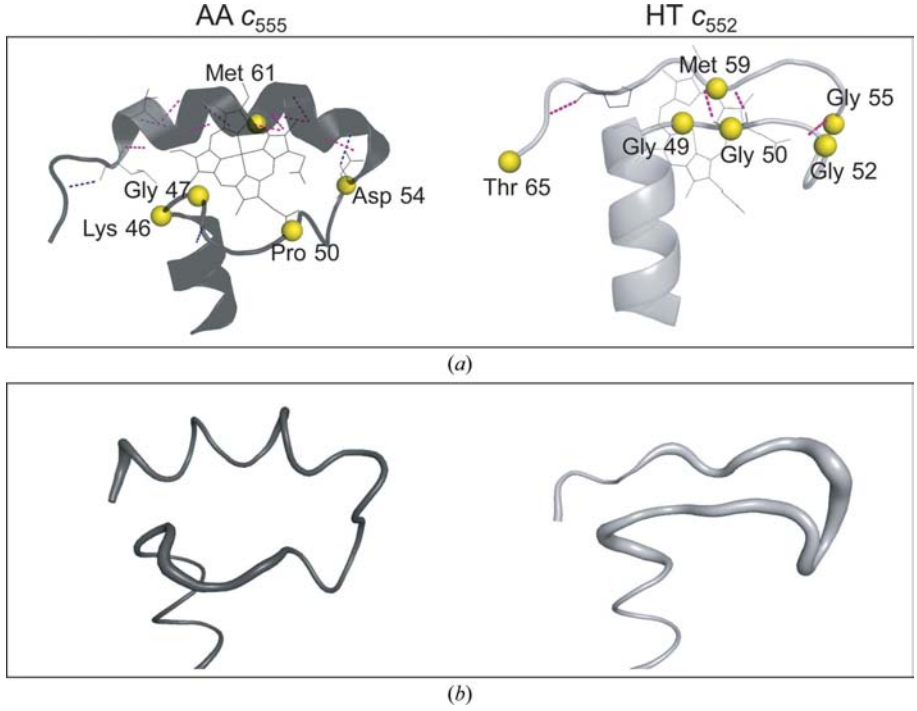

Figure 6
Extra helix and adjacent loops in AA c_{555} and the corresponding region in HT c_{552}. (*a*) The main chains around the AA c_{555} extra helix and corresponding regions in HT c_{552} are presented as ribbon models. Haems are presented as stick models. The C^{α} atoms of relevant amino-acid residues are indicated (yellow). In AA c_{555}, hydrogen bonds are shown by pink and dark blue dashed lines within the extra helix and adjacent loops, respectively. In HT c_{552}, hydrogen bonds within the corresponding region are represented by pink dashed lines. (*b*) Conformational flexibility as observed from the crystallographic *B* factors in the regions shown in (*a*) is represented by the thickness of the tube model.

3.5. Comparison with other cytochromes c

The newly determined AA c_{555} crystal structure was further compared with those of other bacterial and mitochondrial monohaem cytochromes *c* to reveal any structural similarities and differences. More than 50 structures have been examined in terms of their *Z* scores and helix contents using the programs *DALI* and *DSSP* (see §2 for details).

The backbone structure of AA c_{555} is most similar to that of PH c_{552} (PDB code 2d0s), as judged from the *Z* score (10.2), among the monohaem cytochromes *c* examined (Table 2). AA c_{555} also exhibits high structural similarity to HT c_{552} (PDB code 1ynr, $Z = 9.8$) and PA c_{551} (PDB code 2d0s, $Z = 9.8$), followed by *Pseudomonas stutzeri* cytochrome c_{551} (PS c_{551}; PDB code 1cch, $Z = 9.0$). PH c_{552}, HT c_{552}, PA c_{551} and PS c_{551} belong to the class ID monohaem cytochromes *c* and exhibit more than 50% sequence identity to each other. Although the AA c_{555} sequence clearly differs from those of typical class ID cytochromes *c* (Fig. 1), its backbone structure most closely resembles ID cytochromes *c* compared with other monohaem cytochromes *c*.

For comparative three-dimensional structure analysis of AA c_{555} and other cytochromes *c*, we selected the following five bacterial and mitochondrial cytochromes *c* (PDB codes in parentheses): *Thermus thermophilus* cytochrome c_{552} (1c52), *Bacillus pasteurii* cytochrome c_{553} (1c75), *Desulfovibrio vulgaris* cytochrome c_{553} (1dvh), *Saccharomyces cerevisiae* iso-1-cytochrome *c* (1ycc) and *Azotobacter vinelandii* cytochrome c_5 (1cc5). These cytochromes *c* together with HT c_{552} discussed above represent the various monohaem cytochrome *c* structures known to date.

The three-dimensional structure comparison (Fig. 7) showed that AA c_{555} and the other cytochromes *c* share N- and C-terminal α-helices (denoted as αA and αE in Fig. 3) and confirmed that AA c_{555} uniquely contains the extra helix. The other cytochromes *c* have a Ω-loop structure in the corresponding region, similar to class ID cytochromes *c*, including HT c_{552}. Although the axial haem-liganding Met61 residue is located in the middle of the extra helix, AA c_{555} has a redox potential of ∼220 mV (Aubert *et al.*, 2001), which is similar to those of other class ID cytochromes *c* such as HT c_{552} (237 mV, Takeda *et al.*, 2009). This indicates that the presence of the extra helix does not strongly affect the haem environment in the protein.

D. vulgaris cytochrome c_{553} (DV c_{553}; Blackledge *et al.*, 1995) is the only other example that has a helix structure (residues 55–61) in the corresponding region. However, its

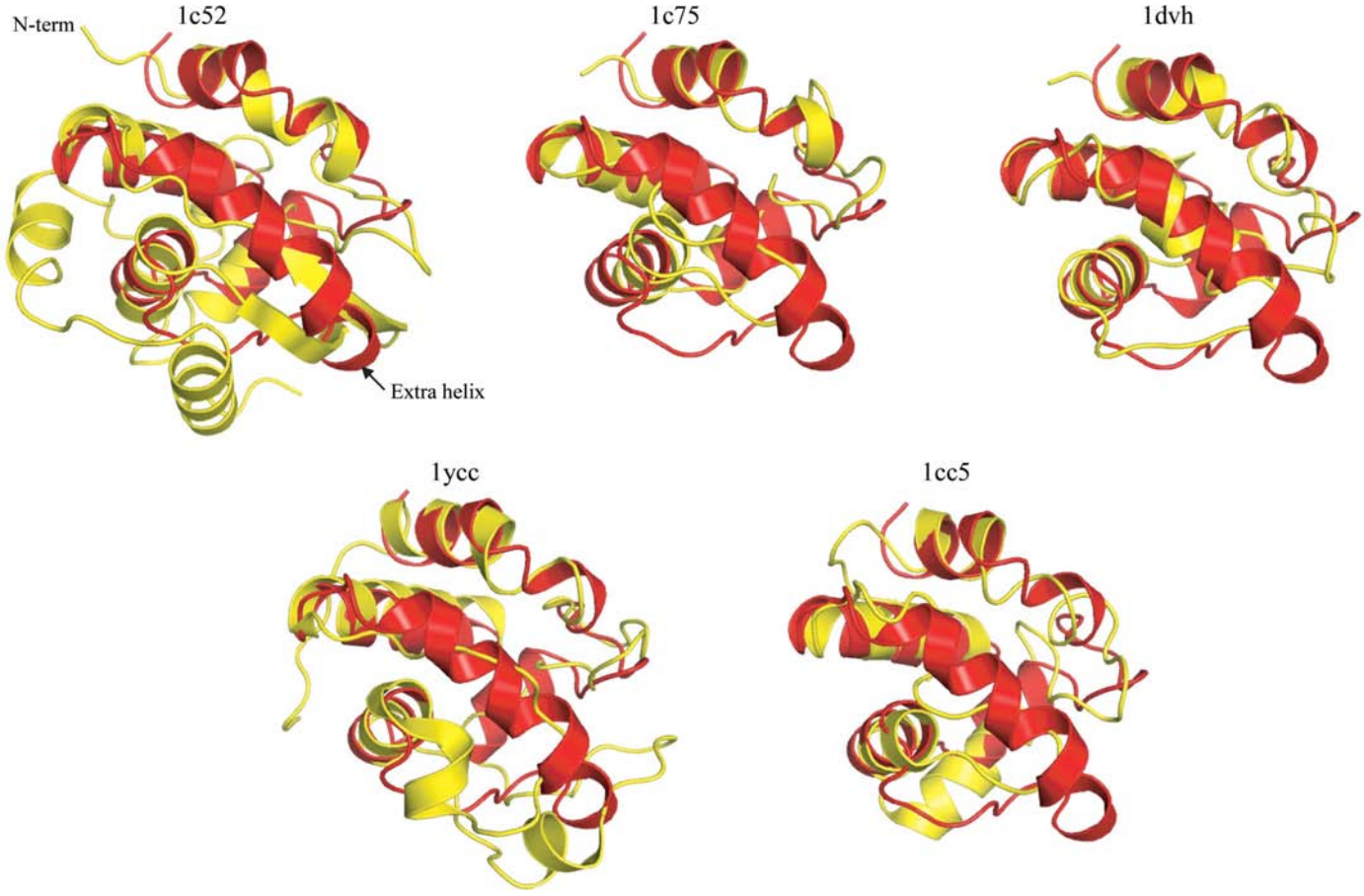

Figure 7
Superpositioning of the main chains. The structures were aligned on the basis of r.m.s. deviations for the superposition of nine haem atoms using *PyMOL*. The main chain of AA c_{555} is shown in red. The other aligned cytochromes *c* are shown in yellow. The N-terminus and extra helix are indicated in an alignment with 1c52.

length is only half that of the extra helix of AA c_{555} and the total number of helix structures is four in DV c_{553} instead of five as in AA c_{555}. The results of structure comparison between AA c_{555} and DV c_{553} show that AA c_{555} has the highest helix content (59.8%) of the monohaem cytochromes *c*, which is followed by that of DV c_{553} (Table 2). This should be a consequence of the presence of the extra helix in AA c_{555}.

4. Conclusions

The structure of AA c_{555} is the first reported crystal structure of a monohaem cytochrome *c* from a hyperthermophilic bacterium. Through the present structural analysis, it has been shown that AA c_{555} has a novel cytochrome *c* fold with a unique extra helix. We conclude that the presence of the extra helix should contribute to the further higher stability (hyperstability) of AA c_{555}, which represents the upper limit of cytochrome *c* stability. The stable HT c_{552} structure arises through improvement of side-chain interactions within class ID cytochromes *c*. However, for further enhancement of the protein stability of AA c_{555}, the main-chain conformation as well as the side-chain interactions identified in this study is required.

The authors thank Miss Yuko Sasaki for her technical assistance. We also thank Dr Seiki Baba of the Japan Synchrotron Radiation Research Institute for his help during the data collection. The X-ray data collection was carried out with the approval of the organizing committee of SPring-8 (Proposal 2007B1635.).

References

Aubert, C., Guerlesquin, F., Bianco, P., Leroy, G., Tron, P., Stetter, K. O. & Bruschi, M. (2001). *Biochemistry*, **40**, 13690–13698.
Bai, Y., Sosnick, T. R., Mayne, L. & Englander, S. W. (1995). *Science*, **269**, 192–197.
Blackledge, M. J., Medvedeva, S., Poncin, M., Guerlesquin, F., Bruschi, M. & Marion, D. (1995). *J. Mol. Biol.* **245**, 661–681.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., Huber, R., Feldman, R. A., Short, J. M., Olsen, G. J. & Swanson, R. V. (1998). *Nature (London)*, **392**, 353–358.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* http://pymol.sourceforge.net.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Hakamada, S., Sonoyama, T., Ichiki, S., Nakamura, S., Uchiyama, S., Kobayashi, Y. & Sambongi, Y. (2008). *Biosci. Biotechnol. Biochem.* **72**, 2103–2109.
Hasegawa, J., Shimahara, H., Mizutani, M., Uchiyama, S., Arai, H., Ishii, M., Kobayashi, Y., Ferguson, S. J., Sambongi, Y. & Igarashi, Y. (1999). *J. Biol. Chem.* **274**, 37533–37537.
Hasegawa, J., Uchiyama, S., Tanimoto, Y., Mizutani, M., Kobayashi, Y., Sambongi, Y. & Igarashi, Y. (2000). *J. Biol. Chem.* **275**, 37824–37828.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
Holm, L. & Park, J. (2000). *Bioinformatics*, **16**, 566–567.
Holm, L. & Sander, C. (1998). *Proteins*, **33**, 88–96.
Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Mason, P. E., Neilson, G. W., Dempsey, C. E., Barnes, A. C. & Cruickshank, J. M. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 4557–4561.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Morris, R. J., Perrakis, A. & Lamzin, V. S. (2002). *Acta Cryst.* D**58**, 968–975.
Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.
Nakamura, S., Ichiki, S., Takashima, H., Uchiyama, S., Hasegawa, J., Kobayashi, Y., Sambongi, Y. & Ohkubo, T. (2006). *Biochemistry*, **45**, 6115–6123.
Oikawa, K., Nakamura, S., Sonoyama, T., Ohshima, A., Kobayashi, Y., Takayama, S. J., Yamamoto, Y., Uchiyama, S., Hasegawa, J. & Sambongi, Y. (2005). *J. Biol. Chem.* **280**, 5527–5532.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Reysenbach, A. L. (2001). *Bergey's Manual of Systematic Bacteriology*, 2nd ed., Vol. 1, edited by D. R. Boone, R. W. Castenholz & G. M. Garrity, pp. 359–367. New York: Springer-Verlag.
Sanbongi, Y., Igarashi, Y. & Kodama, T. (1989). *Biochemistry*, **28**, 9574–9578.
Sambongi, Y., Uchiyama, S., Kobayashi, Y., Igarashi, Y. & Hasegawa, J. (2002). *Eur. J. Biochem.* **269**, 3355–3361.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Takeda, T., Sonoyama, T., Takayama, S. J., Mita, H., Yamamoto, Y. & Sambongi, Y. (2009). *Biosci. Biotechnol. Biochem.* **73**, 366–371.
Terui, N., Tachiiri, N., Matsuo, H., Hasegawa, J., Uchiyama, S., Kobayashi, Y., Igarashi, Y., Sambongi, Y. & Yamamoto, Y. (2003). *J. Am. Chem. Soc.* **125**, 1360–13651.
Travaglini-Allocatelli, C., Gianni, S., Dubey, V. K., Borgia, A., Di Matteo, A., Bonivento, D., Cutruzzola, F., Bren, K. L. & Brunori, M. (2005). *J. Biol. Chem.* **280**, 25729–25734.
Uchiyama, S., Hasegawa, J., Tanimoto, Y., Moriguchi, H., Mizutani, M., Igarashi, Y., Sambongi, Y. & Kobayashi, Y. (2002). *Protein Eng.* **15**, 455–461.