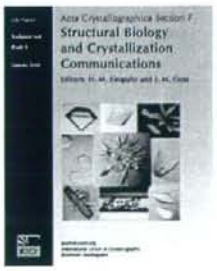
Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091
Editors: H. M. Einspahr and J. M. Guss

Cloning, expression and purification of cytochrome c_6 from the brown alga *Hizikia fusiformis* and complete X-ray diffraction analysis of the structure

Hideharu Akazaki, Fumihiro Kawai, Hirotaka Chida, Yuichirou Matsumoto, Mao Hirayama, Ken Hoshikawa, Satoru Unzai, Wataru Hakamata, Toshiyuki Nishio, Sam-Yong Park and Tadatake Oku

Acta Cryst. (2008). F**64**, 674–680



Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Cloning, expression and purification of cytochrome c_6 from the brown alga *Hizikia fusiformis* and complete X-ray diffraction analysis of the structure

Hideharu Akazaki,[a] Fumihiro Kawai,[b] Hirotaka Chida,[a] Yuichirou Matsumoto,[a] Mao Hirayama,[a] Ken Hoshikawa,[a] Satoru Unzai,[b] Wataru Hakamata,[a] Toshiyuki Nishio,[a] Sam-Yong Park[a] and Tadatake Oku[a]*

[a]Bio-organic Chemistry Laboratory, Graduate School of Bioresource Sciences, Nihon University, Kameino 1866, Fujisawa-shi, Kanagawa 252-8510, Japan, and [b]Protein Design Laboratory, Graduate School of Integrated Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan

Correspondence e-mail: oku@brs.nihon-u.ac.jp

Received 26 April 2008
Accepted 11 June 2008

PDB Reference: cytochrome c_6, 2zbo, r2zbosf.

The primary sequence of cytochrome c_6 from the brown alga *Hizikia fusiformis* has been determined by cDNA cloning and the crystal structure has been solved at 1.6 Å resolution. The crystal belonged to the tetragonal space group $P4_12_12$, with unit-cell parameters $a = b = 84.58$, $c = 232.91$ Å and six molecules per asymmetric unit. The genome code, amino-acid sequence and crystal structure of *H. fusiformis* cytochrome c_6 were most similar to those of red algal cytochrome c_6. These results support the hypothesis that brown algae acquired their chloroplasts *via* secondary endosymbiosis involving a red algal endosymbiont and a eukaryote host.

1. Introduction

Soluble *c*-type monohaem cytochromes are ubiquitously distributed haem proteins which act as electron carriers in mitochondria, bacteria, algal chloroplasts and cyanobacteria. Cytochrome c_6 is a soluble low-spin haem protein that functions in oxygenic photosynthesis as an electron carrier between cytochrome *f*, which is part of the membrane-embedded cytochrome b_6f complex, and the P700 reaction centre of photosystem I (Kerfeld & Krogmann, 1998). This cytochrome c_6 is classified as a class I *c*-type cytochrome, in which the haem iron has histidine–methionine axial coordination. Plastocyanin is a blue copper protein with the same function as cytochrome c_6. Cytochrome c_6 and plastocyanin have completely different amino-acid sequences and secondary and tertiary structures, but they contain similar acidic and hydrophobic patches on their surface for recognition of their interaction partners (Frazão *et al.*, 1995; Ullmann *et al.*, 1997).

Although chloroplasts are thought to have evolutionally arisen from cyanobacteria (Aitken, 1976), there are differences in the expression and genome coding of cytochrome c_6 in green and red algae. In the green alga *Chlamydomonas reinhardtii*, the gene for cytochrome c_6 exists in the genomic DNA and its coding region is interrupted by two introns (Hill *et al.*, 1991). On the other hand, in the red alga *Porphyra purpurea* the *petJ* gene encoding cytochrome c_6 exists in the chloroplast genome (Reith & Munholland, 1993).

In eukaryotic brown algae, which contain no plastocyanin, photosynthetic electron transport between cytochrome *f* and photosystem I is only performed by cytochrome c_6. It is generally considered that brown algae acquired their chloroplasts *via* secondary endosymbiosis involving a primitive red algal endosymbiont and a nonphotosynthetic eukaryote host (Cavalier-Smith, 2000; McFadden, 1999). Although the physicochemical properties and amino-acid sequences of cytochromes c_6 from the brown algae *Petalonia fascia* and *Alaria esculenta* have been determined (Sugimura *et al.*, 1981; Laycock, 1975), the genome code and tertiary structure of brown algal cytochrome c_6 remain to be studied. In this study, we determined the genome code of the brown algal cytochrome c_6 gene from the brown alga *Hizikia fusiformis*, determined the crystal structure of the protein and compared it with those of cyanobacterial and red and green algal cytochromes c_6.

2. Materials and methods

2.1. Sequence determination

The brown alga *H. fusiformis* was collected in the coastal area off Hayama, Japan. Total RNA was isolated from the brown alga using the RNeasy Plant Mini Kit (Qiagen). Poly(A)$^+$ mRNA and poly(A)$^-$ mRNA were separated from the total RNA using Oligotex-dT30 (Takara). Additional of adenine at the 3′-terminus of poly(A)$^-$ mRNA was carried out for 40 min at 310 K in a reaction mixture containing 2 µg poly(A)$^-$ mRNA, 50 m*M* Tris–HCl pH 7.9, 50 m*M* $MgCl_2$, 10 m*M* $MnCl_2$, 500 m*M* NaCl, 2.5 m*M* DTT, 0.5% BSA, 1 m*M* ATP, 121 U ribonuclease inhibitor and 1.5 U poly(A) polymerase (Takara). First-strand cDNA was synthesized using a 1st Strand cDNA Synthesis Kit with AMV Reverse Transcriptase (Life Science Inc.) and the oligonucleotide primer 5′-CGGGATCC$(T)_{25}$-3′, designated primer P1 (reverse). To obtain the clone encoding the 3′-region of cytochrome c_6 from *H. fusiformis*, we designed the degenerate oligonucleotide primer P2 (forward), 5′-AAYTGYGCIGCIGCITGYCAYGCI-3′, based on the highly conserved residues around the haem *c* motif (Asn-Cys-Ala-Ala-Cys-His-Ala) of cytochrome c_6 from the cyanobacteria *Synechocystis* PCC6803 and *Anabaena* 7119, the green alga *C. reinhardtii*, the euglena *Euglena gracillis* and the cyanelle *Cyanophora paradoxa*. PCR products were subcloned into a pGEM T-Easy vector (Promega). DNA sequencing was performed by the dideoxy chain-termination method using a Thermo Sequence fluorescent-labelled primer cycle sequencing kit with 7-deaza-dGTP (Amasham) and an automated DSQ 2000L DNA sequencer (Shimadzu, Japan). The first-strand cDNA from *H. fusiformis* were dC-tailed at their 3′-ends using the 5′ RACE system for Rapid Amplification of cDNA Ends Reagent Assembly v.2.0 (Life Technologies Inc.). The 5′-region of the cytochrome c_6 gene from *H. fusiformis* was amplified by the polymerase chain reaction (PCR) using a forward primer complementary to the dC tail [P3, 5′-GGCCACGCGTCGACTAGTAC$(G)_{16}$-3′] and a gene-specific primer designed based on the 3′-region sequences of the cytochrome c_6 cDNA (Hf1, 5′-TCAGGCATAATAATAACATTATTACCGCC-3′). The PCR product was subcloned and sequenced by the same methods as used in 3′ RACE. Genome DNA from *H. fusiformis* was extracted using Isoplant II (Nippon Gene). To obtain the genome sequence of *H. fusiformis* cytochrome c_6, we designed gene-specific primers for amplification of the full-length cytochrome c_6 gene on the basis of the cDNA sequence of cytochrome c_6 (Hf2, 5′-ATGGGGGGGTGGAAAATTTATTATT-3′, forward; Hf3, 5′-TCAACGTTCCAGGTCCAATAATATCATAA-3′, reverse). The PCR product was subcloned and sequenced according to 3′ RACE.

2.2. Construction of expression vector

Construction and overproduction of the cytochrome c_6 gene (*petJ*) in *Escherichia coli* was performed according to the method described by Satoh *et al.* (2002) with slight modifications. The mature cytochrome c_6 sequence was amplified using the forward primer ExP1 (5′-CATGCCATGGGCTGATATTAATCATGGAG-3′) corresponding to codons for the amino-acid residues of the cytochrome c_6 N-terminal region and the reverse primer ExP2 (5′-GCGGATCCTTAGTTCCAACTCTTTTCAG-3′) corresponding to codons for the amino-acid residues of the C-terminal region. The amplified mature cytochrome c_6 sequence was ligated to the *pelB* signal sequence adapter (Genset Co. Ltd). The resulting *pelB*–cytochrome c_6 hybrid gene was cloned into *Nde*I–*Bam*HI sites of pET22b(+) (Novagen Co. Ltd) to construct the plasmid pET22bHfc6.

The cytochrome *c* maturation genes *ccmA–H* were amplified using the polymerase chain reaction from *E. coli* MC1061 genomic DNA using the forward primer P3, 5′-CCAGAATTCGGTTGCCGCGAAGATGCAT-3′, corresponding to upstream of the *ccmA* gene from the *E. coli* K12 MG1655 genome sequence (AE000309), and the reverse primer P4, 5′-TTCCTGCAGCAACGCGGGGCACAATAAA-3′, corresponding to downstream of the *ccmH* gene. The resulting *ccmA–H* gene was cloned into the *Eco*RI–*Pst*I sites of pSTV28 (Takara Shuzo Co.) to create the plasmid pSTV28*ccmA–H*.

2.3. Protein expression and purification

For the overproduction of *H. fusiformis* cytochrome c_6, both pET22bHfc6 and pSTV28*ccmA–H* were co-introduced into *E. coli* BL21 (DE3). Transformed *E. coli* cells were grown in 1 l Luria–Bertani (LB) medium supplemented with 100 mg l^{-1} ampicillin and 20 mg l^{-1} chloramphenicol at 303 K for 36 h. Cells were harvested by centrifugation at 6000*g* (277 K) for 5 min. The pellet was resuspended in 80 ml PBS buffer and disrupted using a high-pressure homogenizer (Mini Lab 8.30H, Rannie). The suspension was fractionated with ammonium sulfate (40–80% saturation). The precipitate was dissolved in a small amount of 20 m*M* sodium acetate buffer pH 5.5 and dialyzed against the same buffer. The sample was applied onto a DE52 cellulose column (Whatman, 2.0 × 40.0 cm) equilibrated with 20 m*M* sodium acetate buffer pH 5.5. After the column had been washed with the same buffer, the proteins were eluted using a linear gradient of sodium acetate pH 5.5 (20–200 m*M*). Fractions containing cytochrome c_6 were pooled and dialyzed against 20 m*M* sodium acetate buffer pH 5.5 and the dialyzed sample was applied onto a Poros HQ/20 column (Applied Biosystems) previously equilibrated with the same buffer. After the column had been washed with the same buffer, the proteins were eluted with an NaCl gradient (0–500 m*M*) in the same buffer. The sample thus obtained was used as purified recombinant *H. fusiformis* cytochrome c_6. The degree of purity was confirmed by tricine SDS–PAGE (Schägger & von Jagow, 1987) and UV–visible spectroscopy. UV–visible spectra of *H. fusiformis* cytochrome c_6 were measured with a Hitachi U3310 spectrophotometer using quartz cuvettes of 1.0 cm path length. The concentration of the cytochrome c_6 was determined spectrophotometrically from the pyridine ferrohaemochrome spectrum (550 nm, 29.1 mM^{-1} cm^{-1}). Potassium ferricyanide and sodium dithionite were used as the oxidant and the reductant, respectively.

2.4. Crystallization and refinement

The purified protein was dissolved in 10 m*M* sodium phosphate buffer pH 7.0 to prepare a concentrated protein solution of 20 mg ml^{-1}. Initial crystals were obtained using the Wizard I random sparse-matrix crystallization screen (Emerald BioSystem). *H. fusiformis* cytochrome c_6 was crystallized by vapour diffusion using the hanging-drop method at 293 K. Each drop consisted of 2 µl protein solution and 2 µl reservoir solution. An initial crystal of *H. fusiformis* cytochrome c_6 grew within a week using condition No. 33 [2.0 *M* $(NH_4)_2SO_4$, 0.1 *M* CAPS pH 10.5 and 0.2 *M* Li_2SO_4]. To improve the quality of the crystal, further screening for crystallization was performed and crystals were obtained reproducibly using 0.1 *M* CAPS pH 10.5, 0.2 *M* Li_2SO_4, 2.2 *M* $(NH_4)_2SO_4$ and 3% glycerol. X-ray diffraction data were collected on BL-5A, Photon Factory, Tsukuba, Japan. The data set was processed with *HKL*-2000 and scaled with *SCALEPACK* (Otwinowski & Minor, 1997). The structure of *H. fusiformis* cytochrome c_6 was determined by molecular replacement using the program *MOLREP* (Collaborative Computational Project, Number 4, 1994). The search model used was

Table 1
Crystal parameters and data-collection and structure refinement.

Values in parentheses are for the outer shell (1.66–1.60 Å).

Data-collection statistics	
Temperature (K)	100
Resolution range (Å)	50.0–1.6
Space group	$P4_12_12$
Unit-cell parameters (Å)	$a = b = 84.578$, $c = 232.911$
Reflections (measured/unique)	699906/107513 (9592)
Completeness (%)	95.6 (86.8)
R_{merge}† (%)	4.8 (23.5)
Redundancy	6.6 (3.5)
Mean $I/\sigma(I)$	20.5
Mosaicity	0.33
Refinement statistics	
Resolution range (Å)	20.0–1.6
σ Cutoff/reflections used	0.0/107239
R factor/R_{free}‡ (%)	18.4/20.9
R.m.s.d. bond lengths (Å)/bond angles (°)	0.011/1.165
B factors (Å^2)	
Average	22.4
Protein	21.1
Haem	14.8
Water	34.0
Sulfate	40.5
Ramachandran plot	
Residues in most favourable region (%)	82.7
Residues in additional allowed region (%)	16.0
Residues in disallowed region (%)	1.3

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation and $\langle I(hkl)\rangle$ is the mean value for the unique reflection; summations are over all reflections. ‡ R factor $= \sum_h |F_o(h) - F_c(h)| / \sum_h F_o(h)$, where F_o and F_c are the observed and calculated structure-factor amplitudes, respectively. The free R factor was calculated using 5% of the data, which were excluded from the refinement.

Porphyra yezoensis cytochrome c_6 (Yamada *et al.*, 2000). The structure of *H. fusiformis* cytochrome c_6 was refined with *REFMAC* from the *CCP*4 program suite. Water molecules were added using a water-pick script in *CNS* and refinement was continued using *REFMAC*5 (Collaborative Computational Project, Number 4, 1994). The final model obtained had an R factor of 18.4% and a free R factor of 20.9%. Manual model building was carried out using *Coot* (Emsley & Cowtan, 2004). Solvent molecules were placed at positions where spherical electron-density peaks were found above 1.5σ in the $|2F_o - F_c|$ map and above 3.0σ in the $|F_o - F_c|$ map and where stereochemically reasonable hydrogen bonds were allowed. A summary of the data-collection and refinement statistics is given in Table 1.

3. Results and discussion

3.1. Sequence of *H. fusiformis* cytochrome c_6

To elucidate the genome code of a cytochrome c_6 gene from a brown alga, we determined the protein cDNA from the brown alga *H. fusiformis* as shown in Fig. 1 (Genbank accession No. AB105058). *H. fusiformis* cytochrome c_6 genes were amplified using cDNA, which was performed by the reverse transcription of poly(A)$^-$ mRNA. The polyadenylation signal sequences (AAUAAA) necessary for the addition of polyadenylic acid were not included in the 3′-region of the cytochrome c_6 gene from *H. fusiformis*, but the 3′-regions of the cDNA of the cytochrome c_6 that contained the sequence that can form a stem-loop structure that stabilizes mRNA were transcribed from the chloroplast genome (Drager *et al.*, 1996; Yang & David, 1997). The gene that was transcribed from the chloroplast genome does not add polyadenylic acids (Sagher *et al.*, 1976). Generally, the addition of polyadenylic acids that participate in mRNA stability occurs after transcription inside the nucleus (Darnell *et al.*, 1971; O'Hara *et al.*, 1995). The Shine–Dalgano (SD) sequence, a 16S-ribosomal RNA-binding site that is rich in purine 3–9 bases upstream of the initiation codon of prokaryotic cell mRNA (Bonham-Smith &

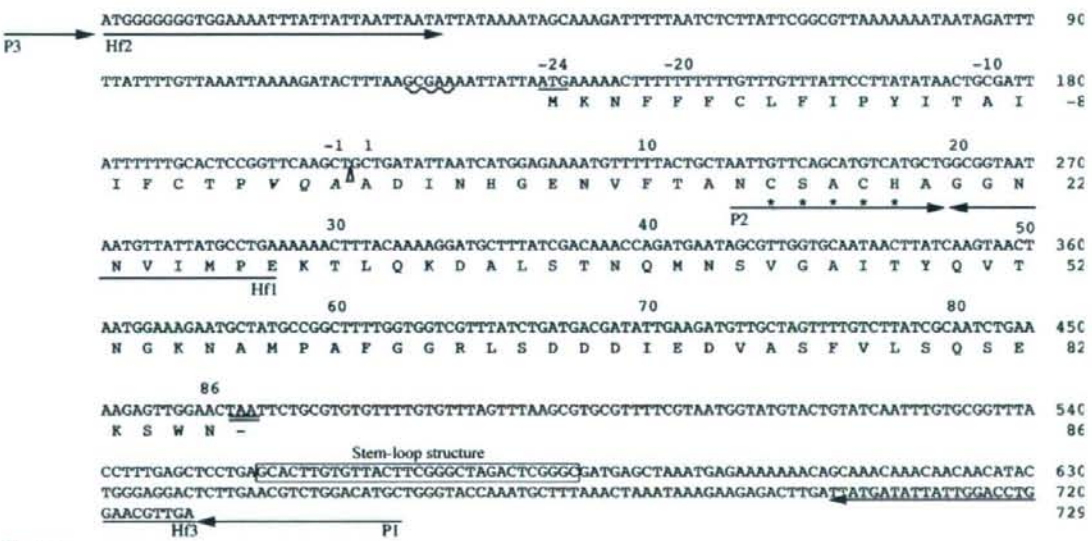


Figure 1
Nucleotide sequence and deduced amino-acid sequence of the cDNA encoding cytochrome c_6 from the brown alga *H. fusiformis*. The amino-acid residues numbered −1 to −24 and 1–86 constituted the putative transit sequence and the mature peptide sequence, respectively. Underlined and double-underlined nucleotides indicate the initiation codon and termination codon, respectively. Nucleotides underlined with a wavy line indicate the Shine–Dalgarno-like sequence. The open arrowhead indicates the putative transit-peptide cleavage site. The arrows indicates the primers used. The residues marked with asterisks and the amino-acid residues in italics (−1 to −3) correspond to the haem *c* motif (Cys-*X*-*X*-Cys-His) and cleavage motif (Val-*X*-Ala), respectively. The box indicates the stem-loop structure-formation sequence.

Bourque, 1989), was present ten bases upstream of the initiation codon in the *H. fusiformis* cytochrome c_6 gene (Fig. 1). There were also no SD sequences in the cytochrome c_6 genes of the green alga *C. reinhardtii* (Merchant & Bogorad, 1987), the euglena *Euglena gracilis* (Vacula *et al.*, 1999) and the cyanelle *Cyanophora paradoxa* (Steiner *et al.*, 2000) that are encoded in the nuclear genome. We obtained a genomic DNA clone of approximately 730 bp that was amplified using primers constructed based on cDNA sequences (Fig. 1). The gene that encodes cytochrome c_6 was not inserted with an intron. The green alga *C. reinhardtii* gene encoding cytochrome c_6 has been reported to have its coding region interrupted by two introns (Hill *et al.*, 1991). Genes encoded in nuclear genomes are usually inserted with introns (The Arabidopsis Initiative, 2000), but genes encoded in chloroplast genomes do not have these insertions (Shinozaki *et al.*, 1986). These results showed that the cytochrome c_6 gene from the brown alga *H. fusiformis* was encoded in the chloroplast genome. At present, only red and brown algae have been reported to have a cytochrome c_6 gene encoded in the chloroplast genome.

3.2. Protein expression and purification

E. coli BL21 (DE3) harbouring both pET22bHfc6 and pSTV28*ccmA–H* was used as a source of recombinant *H. fusiformis* cytochrome c_6. Recombinant *H. fusiformis* cytochrome c_6 was purified by ammonium sulfate pecipitation and two-step anion-exchange chromatography. The degree of homogeneity was confirmed by tricine SDS–PAGE and UV–visible spectroscopy. After initial purification by anion-exchange chromatography, tricine SDS–PAGE analysis displayed a predominant cytochrome c_6 band and a minor band and the fractions with an $A_{275}/A_{552.5}$ ratio lower than 2.0 were pooled and concentrated. After a second chromatography purification step, tricine SDS–PAGE analysis showed only the cytochrome c_6 band and the purification ratio ($A_{275}/A_{552.5}$) of *H. fusiformis* cytochrome c_6 was 0.90, which was similar to that of other cytochromes c_6.

The protein consists of 86 amino acids and one *c*-type haem and its molecular weight was calculated to be 9762.4 Da. From SDS–PAGE analysis a value of 8.0 kDa was obtained, which is somewhat lower than that deduced from the sequence (Fig. 2*a*). Similar discrepancies have been observed in other small negatively charged proteins, such as the green alga *Monoraphidium braunii* cytochrome c_6 and plastocyanin (Campos *et al.*, 1993).

The UV–visible spectra of reduced and oxidized recombinant *H. fusiformis* cytochrome c_6 are shown in Fig. 2(*b*). In the reduced form, the α, β, γ (Soret) and δ absorption maxima peaks appear at 552.5, 522.0, 416.0 and 318.0 nm, respectively. For the oxidized form of the cytochrome c_6, the $\alpha + \beta$ and γ (Soret) absorption maxima peaks were 528.0 and 408.5 nm, respectively; a shoulder peak at 695.0 nm, indicating His–Fe–Met coordination, was observed (Fig. 2*b*, inset).

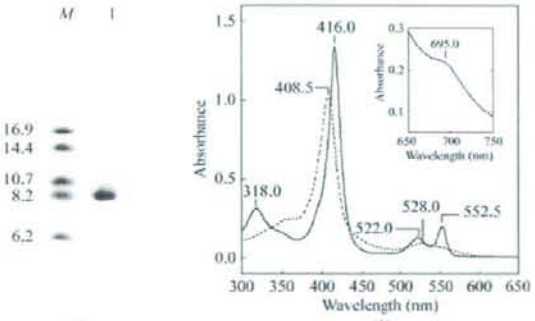


Figure 2
SDS–PAGE analysis and UV–visible spectra of purified *H. fusiformis* cytochrome c_6. (*a*) Proteins were analysed on 16.5% tricine SDS–PAGE and stained with Coomassie Blue. Lane *M*, molecular-weight markers (kDa); lane 1, purified *H. fusiformis* cytochrome c_6. (*b*) UV–visible spectra of dithionite-reduced (solid line) and ferricyanide-oxidized (dotted line) forms of *H. fusiformis* cytochrome c_6 (10 μM) were measured in 10 m*M* sodium phosphate pH 7.0 at 298 K. The inset shows the 695 nm band of the oxidant form at 500 μM.

3.3. Crystallization of *H. fusiformis* cytochrome c_6

A crystallization droplet was prepared by mixing 2 µl protein solution (20 mg ml^{-1} protein) in 10 m*M* sodium phosphate buffer pH 7.0 and 2 µl reservoir solution consisting of 0.1 *M* CAPS pH 10.5, 0.2 *M* Li_2SO_4, 2.2 *M* $(NH_4)_2SO_4$ and 3% glycerol and was equilibrated against 500 µl of the same reservoir solution at 293 K. Diffraction-quality crystals appeared within a week (Fig. 3). This reservoir solution used for *H. fusiformis* cytochrome c_6 is somewhat similar to that used for cytochrome c_6 from the cyanobacterium *Arthrospira maxima* [reservoir solution containing 0.1 *M* Tris pH 7.8, 0.2 *M* Li_2SO_4, 2.2 *M* $(NH_4)_2SO_4$ and 1% glycerol (Sawaya *et al.*, 2001)], but few similarities were found between the reservoir solutions used for *H. fusiformis* cytochrome c_6 and those used for other algal and cyanobacterial cytochromes c_6 (Kerfeld *et al.*, 1995; Frazão *et al.*, 1995; Schnackenberg *et al.*, 1999; Yamada *et al.*, 2000; Dikiy *et al.*, 2002; Worrall *et al.*, 2007).

3.4. Overall structure of *H. fusiformis* cytochrome c_6

The crystal structure of *H. fusiformis* cytochrome c_6 has been determined at 1.6 Å resolution. This is the first cytochrome c_6 crystal structure for a brown secondary symbiotic alga. The crystal belonged to space group $P4_12_12$, with unit-cell parameters $a = b = 84.58$, $c = 232.9$ Å and six molecules (*A*, *B*, *C*, *D*, *E* and *F*) per asymmetric unit (Fig. 4*a*). These six molecules could be superimposed with main-chain root-mean-square deviation (r.m.s.d.) values of 0.1–0.4 Å, as determined using the *DALI* program (Holm & Park, 2000). The hexamer contains four sulfate ions which may be derived from the ammonium sulfate and lithium sulfate included in the crystallization solution. The cytochrome c_6 hexamer was formed of a dimer of trimers (*ABC* and *DEF* trimers; Fig. 4*b*). An intermolecular hydrogen bond was formed in the *ABC* trimer between each pair of molecules

Figure 3
Crystal of *H. fusiformis* cytochrome c_6 grown in 0.1 *M* CAPS pH 10.5, 0.2 *M* Li_2SO_4, 2.2 *M* $(NH_4)_2SO_4$ and 3% glycerol.

in the trimer through Ala60 N and Arg64 O and one sulfate ion was centred between the Arg64 side chains of the three molecules (Fig. 4c). This arrangement of a sulfate ion enclosed by a basic amino-acid residue has been also found in the crystal structure of *Hydrogenobacter thermophilus* cytochrome c_{552} (Travaglini-Allocatelli *et al.*, 2005). Considering that the crystals of *H. fusiformis* cytochrome c_6 were obtained in the presence of sulfate ions, the sulfate ions were convenient for crystallization and might contribute to crystal-packing stabilization by neutralization of charge repulsion in this region. Therefore, we deduce that the crystallographic hexamer is a nonphysiological crystal-packing artifact. An intermolecular hydrogen bond was formed between each pair of molecules of the second trimer through the C-terminal Asn86 $N^{\delta2}$ and Asn86 OX (Fig. 4c). Hydrogen bonds between the two trimers were formed between Asn22 $O^{\delta1}$ of the *ABC* trimer and Arg64 $N^{\varepsilon2}$ of the *DEF* trimer.

An oligomeric arrangement of molecules has been found in the crystal structures of other cytochromes c_6. A trimeric arrangement of molecules has been found in the structures of cytochrome c_6 from *C. reinhardtii* form I (Kerfeld *et al.*, 1995) and *M. braunii* (Frazão *et al.*, 1995). The proteins from *Scenedesmus obliquus* (Schnackenberg *et al.*, 1999), *A. maxima* (Sawaya *et al.*, 2001) and *Phormidium laminosum* (Worrall *et al.*, 2007) have been crystallized as dimers. These oligomers of cytochromes c_6 were formed by the packing of different molecules and were not superimposed. It has been reported that the observed differences in oligomerization between various cytochromes c_6 may be determined by subtle differences in their surface electrostatic potential properties (Dikiy *et al.*, 2002). In

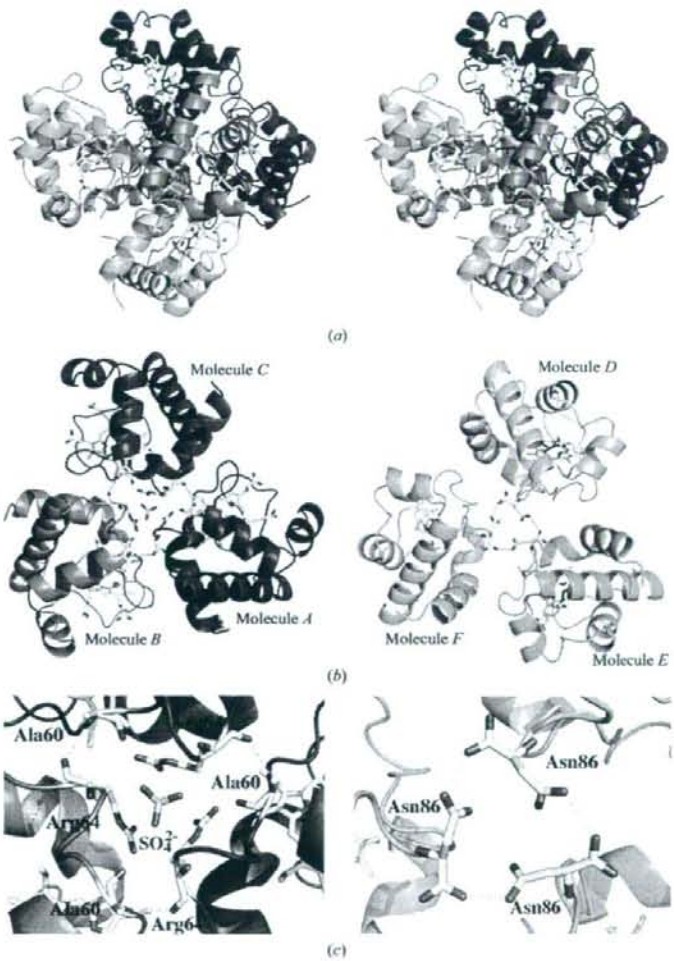


Figure 4
H. fusiformis cytochrome c_6 hexamer. Six protein molecules are displayed, with each molecule in a different colour (red, molecule *A*; marine, molecule *B*; magenta, molecule *C*; lemon, molecule *D*; green, molecule *E*; cyan, molecule *F*). The amino-acid residues and haem group are represented by a stick model with atom-specific colours: white, carbon; blue, nitrogen; red, oxygen; yellow, sulfur; iron, orange. This figure was drawn with *PyMOL*.

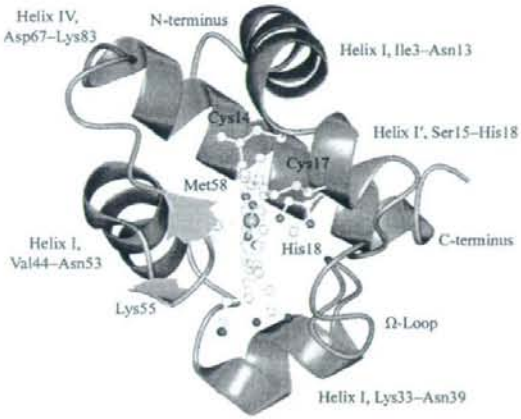


Figure 5
Overall structure of cytochrome c_6 from the brown alga *H. fusiformis*. The α-helix (marine) and β-sheet (green) are indicated as a cartoon model. Cys14, Cys17, His18, Met58 and haem are represented using a ball-and-stick model in the same colour scheme used in Fig. 1. This figure was drawn with *CCP*4 (Collaborative Computational Project, Number 4, 1994).

contrast, the cytochromes c_6 from *P. yezoensis* (Yamada *et al.*, 2000) and *Cladophora glomerate* (Dikiy *et al.*, 2002) are monomeric in the crystal.

The structure of *H. fusiformis* cytochrome c_6 belongs to the class I *c*-type cytochromes, which are composed of four α-helices and tight turns (Fig. 5). The protein consists of a single polypeptide chain folded around the haem prosthetic group. The secondary structures have been classified according to the criteria of Kabsch & Sander (1983). Four α-helices, Asp2–Asn13 (I), Ser15–His18 (I'), Lys33–Ala38 (II), Ile44–Asn53 (III) and Asp67–Lys83 (IV), are found as elements of a regular secondary structure, with helices I and IV overlapping at about 90° (Fig. 3). A two-stranded antiparallel β-sheet was formed with two interchain hydrogen bonds between Lys55 and Met58, which form a type II' β-turn with Asn56 and Ala57. A short β-sheet has commonly been observed in the structures of cyanobacterial, green and red algal cytochromes c_6.

3.5. Structural comparison between *H. fusiformis* and other cytochromes c_6

The crystal structures of four cytochromes c_6 from the eukaryotic green algae *C. reinharditii* (Kerfeld *et al.*, 1995), *M. braunii* (Frazão *et al.*, 1995), *S. obliquus* (Schnackenberg *et al.*, 1999) and *C. glomerata* (Dikiy *et al.*, 2002), of one from the eukaryotic red alga *P. yezoensis* (Yamada *et al.*, 2000) and of two from the prokaryotic cyanobacteria *A. maxima* (Sawaya *et al.*, 2001) and *P. laminosum* (Worrall *et al.*, 2007) have been determined. They are composed of 85–90 amino acids and their main secondary-structural elements are α-helices wrapping around the haem prosthetic group. An amino-acid sequence comparison of *H. fusiformis* cytochrome c_6 with those from *C. reinharditii*, *M. braunii*, *C. glomerata*, *S. obliquus*, *P. yezoensis*, *A. maxima* and *P. laminosum* revealed similarities of 46.67, 47.78, 45.05, 47.19, 72.09, 53.33 and 59.77%, respectively (Fig. 6) and the amino-acid sequence of *H. fusiformis* cytochrome c_6 is most similar to that of *P. yezoensis* cytochrome c_6. The main-chain r.m.s.d.s between *H. fusiformis*, *C. reinharditii*, *P. yezoensis* and *A. maxima* cytochromes c_6 are 0.5–1.1 Å, as determined using the *DALI* program (Holm & Park, 2000). A C^{α} trace of *H. fusiformis* cytochrome c_6 shows a high overall similarity between the green algal and cyanobacterial cytochromes c_6, as well as subtle differences (Fig. 7). The largest deviation in the C^{α} trace between the brown alga *H. fusiformis* cytochrome c_6 and green algal and cyanobacterial cytochromes c_6 was found in the second interconnecting loop (Gln40–Ser43; Fig. 7). The green algal and cyanobacterial cytochromes c_6 have a small insertion of 2–4 amino acids in this region compared with *H. fusiformis* cytochrome c_6. The loop region in *H. fusiformis* cytochrome c_6 resembles that in cytochrome c_6 from the red alga *P. yezoensis*, which also lacks two amino acids in this region compared with green algal and cyanobacterial cytochromes c_6. Considering that the loop region of cytochromes c_6 has a poorly conserved amino-acid sequence compared with other regions, this region may have no common biological functional role. In the structure of other cytochromes, functional roles have not been reported for this region.

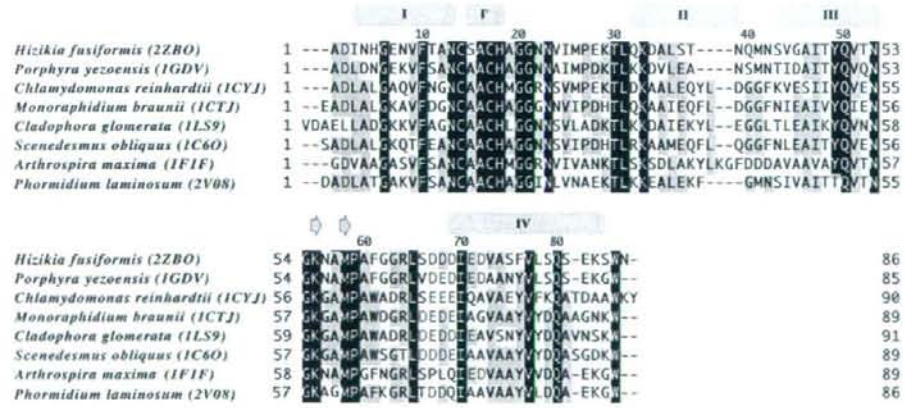


Figure 6
Aligned amino-acid sequences of cytochromes c_6 with reported crystal structures. The conserved and semi-conserved amino-acid residues among the six algal species and two cyanobacterial species are indicated by black and grey boxes, respectively. The haem ligands and the residues forming the acidic patch of the exposed surface are shown in yellow and red, respectively. The secondary structure of cytochrome c_6 from *H. fusiformis* is indicated: orange cylinders, α-helices; blue arrows, β-sheets.

Figure 7
Superimposition of the C^{α} traces of oxidized cytochromes c_6 from the brown alga *H. fusiformis* (orange; PDB code 2zbo), the red alga *P. yezoensis* (red; PDB code 1gdv), the green alga *C. reinhardtii* (green; PDB code 1cyj) and the cyanobacterium *A. maxima* (marine; PDB code 1f1f). The alignment was prepared using the *DALI* program (Holm & Park, 2000).

In this study, we showed that the cytochrome c_6 gene from the brown alga *H. fusiformis* was encoded in the chloroplast genome. To date, the cytochrome c_6 gene has only been found to be encoded in the chloroplast genome in red and brown algae. The amino-acid sequence and tertiary structure of *H. fusiformis* cytochrome c_6 were very similar to those of a red algal cytochrome c_6 rather than those of green algal cytochromes c_6. The present results support the hypothesis that brown algae gained their chloroplasts *via* secondary endosymbiosis involving a primitive red algal endosymbiont and a nonphotosynthetic eukaryote host.

We thank Messrs Daisuke Tamura, Naoya Terunuma and Masaki Hosokawa and Mses Ayako Ohsuzu, Ayumi Hisamitsu and Kasumi Suzuki of Nihon University for the expression and purification of *H. fusiformis* cytochrome c_6. This work was supported in part by a Nihon University Multidisciplinary Research Grant for 2008.

References

Aitken, A. (1976). *Nature (London)*, **263**, 793–796.
Bonham-Smith, P. C. & Bourque, D. P. (1989). *Nucleic Acids Res.* **17**, 2057–2080.
Campos, A. P., Aguiar, A. P., Hervás, M., Regalla, M., Navarro, J. A., Ortega, J. M., Xavier, A. V., De La Rosa, M. A. & Teixeira, M. (1993). *Eur. J. Biochem.* **216**, 329–341.
Cavalier-Smith, T. (2000). *Trends Plant Sci.* **5**, 174–182.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Darnell, J. E., Philipson, L., Wall, R. & Adesnic, M. (1971). *Science*, **174**, 507–510.
Dikiy, A., Carpentier, W., Vandenberghe, I., Borsari, M., Safarov, N., Dikaya, E., Van Beeumen, J. & Ciurli, S. (2002). *Biochemistry*, **41**, 14689–14699.
Drager, R. G., Zeidler, M., Simpson, C. L. & Stern, D. B. (1996). *RNA*, **2**, 652–663.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Frazão, C., Soares, C. M., Carrondo, M. A., Pohl, E., Dauter, Z., Wilson, K. S., Hervás, M., Navarro, J. A., De La Rosa, M. A. & Sheldrick, G. M. (1995). *Structure*, **3**, 1159–1169.
Hill, K. L., Li, H. H., Singer, J. S. & Merchant, S. (1991). *J. Biol. Chem.* **266**, 15060–15067.
Holm, L. & Park, J. (2000). *Bioinformatics*, **16**, 566–567.
Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.
Kerfeld, C. A., Anwar, H. P., Interrante, R., Merchant, S. & Yeates, T. O. (1995). *J. Mol. Biol.* **250**, 627–647.
Kerfeld, C. A. & Krogmann, D. W. (1998). *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**, 397–425.
Laycock, M. V. (1975). *Biochem. J.* **149**, 271–279.
McFadden, G. I. (1999). *J. Eukaryot. Microbiol.* **46**, 339–346.
Merchant, S. & Bogorad, L. (1987). *J. Biol. Chem.* **262**, 9062–9067.
O'Hara, E. B., Chekanova, J. A., Ingle, C. A., Kushner, Z. R., Peters, E. & Kushner, S. R. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 1807–1811.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Reith, M. & Munholland, J. (1993). *Plant Cell*, **5**, 465–475.
Satoh, T., Itoga, A., Isogai, Y., Kurihara, M., Yamada, S., Natori, M., Suzuki, N., Suruga, K., Kawachi, R., Arahira, M., Nishio, T., Fukazawa, C. & Oku, T. (2002). *FEBS Lett.* **531**, 543–547.
Sagher, D., Grosfeld, H. & Edelman, M. (1976). *Proc. Natl Acad. Sci. USA*, **73**, 722–726.
Sawaya, M. R., Krogmann, D. W., Serag, A., Ho, K. K., Yeates, T. O. & Kerfeld, C. A. (2001). *Biochemistry*, **40**, 9215–9225.
Schägger, H. & von Jagow, G. (1987). *Anal. Biochem.* **166**, 368–379.
Schnackenberg, J., Than, M. E., Mann, K., Wiegand, G., Huber, R. & Reuter, W. (1999). *J. Mol. Biol.* **290**, 1019–1030.
Shinozaki, K. *et al.* (1986). *EMBO J.* **5**, 2043–2049.
Steiner, J. M., Serrano, A., Allmaier, G., Jakowitsch, J. & Löffelhardt, W. (2000). *Eur. J. Biochem.* **267**, 4232–4241.
Sugimura, Y., Hase, T., Matsubara, H. & Shimokoriyama, M. (1981). *J. Biochem. (Tokyo)*, **90**, 1213–1219.
The Arabidopsis Initiative (2000). *Nature (London)*, **408**, 796–815.
Travaglini-Allocatelli, C., Gianni, S., Dubey, V. K., Borgia, A., Di Matteo, A., Bonivento, D., Cutruzzolà, F., Bren, K. L. & Brunori, M. (2005). *J. Biol. Chem.* **280**, 25729–25734.
Ullmann, G. M., Hauswald, M., Jensen, A., Kostic, N. M. & Knapp, E. W. (1997). *Biochemistry*, **36**, 16187–16196.
Vacula, R., Steiner, J. M., Krajcovic, J., Ebringer, L. & Löffelhardt, W. (1999). *DNA Res.* **6**, 45–49.
Worrall, J. A., Schlarb-Ridley, B. G., Reda, T., Marcaida, M. J., Moorlen, R. J., Wastl, J., Hirst, J., Bendall, D. S., Luisi, B. F. & Howe, C. J. (2007). *J. Am. Chem. Soc.* **129**, 9468–9475.
Yamada, S., Park, S.-Y., Shimizu, H., Koshizuka, Y., Kadokura, K., Satoh, T., Suruga, K., Ogawa, M., Isogai, Y., Nishio, T., Shiro, Y. & Oku, T. (2000). *Acta Cryst.* D**56**, 1577–1582.
Yang, J. & David, B. S. (1997). *J. Biol. Chem.* **272**, 12874–12880.

Biosci. Biotechnol. Biochem., **72** (10), 2791–2794, 2008

Note

Physicochemical Properties of Diheme Cytochrome c_4 of Unknown Function from *Vibrio parahaemolyticus* Strain RIMD2210633

Hideharu AKAZAKI,[1] Yoshio FUTAMI,[1] Naoya SHIBAYAMA,[2] Ikuko SHIRASAKI,[1] Harumi NAKADE,[1] Hirotaka CHIDA,[1] Wataru HAKAMATA,[1] Sam-Yong PARK,[3] Toshiyuki NISHIO,[1] and Tadatake OKU[1,†]

[1]*Bio-Organic Chemistry Laboratory, Graduate School of Bioresource Sciences, Nihon University, 1866 Kameino, Fujisawa-shi, Kanagawa 252-8510, Japan*
[2]*Department of Physiology, Division of Biophysics, Jichi Medical School, 3311-1 Yakushiji, Shimotsuke-shi, Tochigi 329-0498, Japan*
[3]*Protein Design Laboratory, Guraduate School of Integrated Science, Yokohama City University, 1-7-29 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan*

Received June 5, 2008; Accepted July 7, 2008; Online Publication, October 7, 2008
[doi:10.1271/bbb.80380]

To characterize a diheme cytochrome c_4 of unknown functional of the *Vibrio* genus for the first time, the *Vibrio parahaemolyticus* cytochrome c_4 was overexpressed in *Escherichia coli* periplasm using the endogenous signal sequence. The physicochemical properties of the purified recombinant protein, *viz.*, molecular mass, UV/Vis, and CD spectra, and the redox potentials of the N- and C-terminal domain hemes were determined.

Key words: diheme cytochrome c_4; physicochemical property; recombinant protein; *Vibrio parahaemolyticus*

Bacterial respiratory chains are composed of a variety of electron transport constituents, such as flavoproteins, ironsulfur proteins, quinones, and cytochromes (cyts), and the constituents of the system differ with the kind of microorganisms.[1] In the case of mitochondrial respiratory chains, only cyt *c* functions as an electron carrier between the cyt bc_1 complex and terminal aa_3-type cyt *c* oxidase. On the other hand, electron transfer between the cyt bc_1 complex and the various terminal oxidases in the bacterial respiratory chain are predicted to be performed not only by cyt *c* but also by other monoheme and diheme *c*-type cyts, but the protein interaction between these electron carrier proteins has scarcely been elucidated. Cyt c_4 is a diheme *c*-type cyt, composed of two globular domains (the N-terminal domain and the C-terminal domain), each with heme-iron Met-His axial coordination,[2,3] and is distributed in some genera of γ proteobacteria. To date, cyts c_4 from the acidophilic *Thiobacillus ferrooxidans*,[4] the soil bacterium *Azotobacter vinelandii*,[5] the marine bacterium *Pseudomonas stutzeri*,[6] and the purple sulfur photosynthetic bacterium *Thiocapsa roseopersicina*[7] have been purified and characterized. It has been reported that *T. ferrooxidans* cyt c_4 functions as an electron carrier between copper protein rusticyanin and aa_3-type cyt *c* oxidase,[8] but in other bacteria that have no rusticyanin, the function of this protein has been not yet determined.

Vibrio parahaemolyticus is a gram-negative facultatively anaerobic halophilic bacterium. Although the complete genome sequences of *V. parahaemolyticus* strain RIMD2210633 have been published[9] and various metabolic process can be predicted using this information, its respiratory electron transfer chain remains poorly understood. *Vibrio* genera have the gene of diheme cyt, annotated cyt c_4, and do not contain the gene of rusticyanin in their genome DNA. Considering that the type of the terminal oxidase of the *Vibrio* genera is the cbb_3-type, the mechanism of protein interaction between cyt c_4 and terminal oxidase is predicted to be different from that of *T. ferrooxidans* cyt c_4 and terminal oxidase. Here, to characterize *V. parahaemolyticus* cyt c_4 (VPc4), we expressed VPc4 in *Escherichia coli* and characterized for the first time the physicochemical properties of the cyt c_4 from a facultatively anaerobic bacterium.

The gene encoding VPc4 (GenBank Accession no. BA000031) was amplified from *V. parahaemolyticus* strain RIMD2210633 genome DNA using two primers: Vpc4-NdeI, 5′-ACCCATATGAAGAAATTAGCGCTAATTTTG-3′, and Vpc4-BamHI, 5′-AAAGGATCCTCAGTGCAGACCGCCAACGTA-3′. The product was cloned in the *Nde* I-*Bam* HI site of pET22b (+) (Novagen, Madison, WI, USA) to construct plasmid pET22bVPc4. For overproduction of VPc4, the expres-

† To whom correspondence should be addressed. Tel/Fax: +81-466-84-3950; E-mail: oku@brs.nihon-u.ac.jp

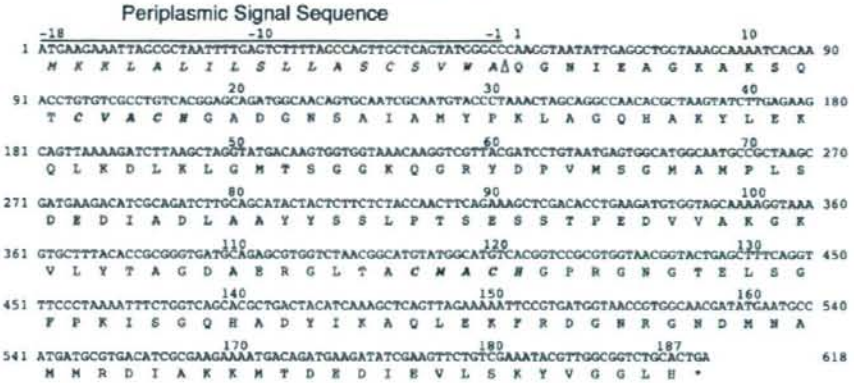


Fig. 1. Nucleotide Sequence of the VPc4 and Deduced Amino Acid Sequence (GenBank Accession no. VP0110).
The signal sequence is underlined. Amino acid residue in bold face italic type corresponds to the Heme *c* motif (Cys-X-X-Cys-His). The open arrowhead indicates the periplasmic signal sequence cleavage site.

sion vector was co-introduced with pSTV28*ccmA-H* into *E. coli* BL21 (DE3).[10] The cultures were grown for 48 h at 125 rpm at 30 °C. Cells were harvested by centrifugation at 6,000 rpm. Periplasmic extraction was performed as described previously.[11] The periplasmic suspension was fractionated with ammonium sulfate (30–70% saturation). The precipitate was dissolved and dialyzed against 10 mM acetate buffer (pH 4.0). The dialyzed sample was centrifuged at 10,000 rpm, and the supernatant was collected. The supernatant was dialyzed against 10 mM Tris–HCl buffer (pH 7.0), and was applied to a HiTrp Q XL column (GE Healthcare Bio-science, Piscataway, NJ, USA) (5 ml) equilibrated with the same buffer. The proteins were eluted using a liner gradient of NaCl (0–350 mM) in 10 mM Tris–HCl buffer (pH 7.0), and the fractions of VPc4 were collected. VPc4 was dialyzed against 10 mM Tris–HCl buffer (pH 8.0), and was applied to a HiTrp Q XL column (5 ml) equilibrated previously with same buffer. VPc4 was eluted using a liner gradient of NaCl (0–350 mM) in 10 mM Tris–HCl buffer (pH 8.0). The sample obtained by this chromatography was used as purified recombinant VPc4. The final yield of purified recombinant VPc4 was approximately 0.1 µmol (2.12 mg)/l. This expression yield was a little higher than that of recombinant *Pseudoalteromonas haloplanktis* cyt c_4 (0.08 µmol (1.66 mg)/l).[12] The N-terminal amino acid sequence of VPc4 was determined to be QGNIEAGKAKSQT, cleaved at the N-terminal 18 amino acid residues, by amino acid sequence analysis using a Perkin Elmer Biosystems model Procise 49X HT protein sequencer (Fig. 1). This results indicates that the N-terminal 18 amino acid residues functioned as a periplasmic signal sequence. The protein consisted of 187 amino acids and two *c*-type hemes, and its molecular mass was calculated to be 21,167 Da. This value is in agreement with that obtained by SDS–PAGE analysis. In addition, the results of MALDI-TOF mass spectrometry using an Applied Biosystem 4800 plus MALDI-TOF/TOF analyzer of the protein was closely in agreement with the theoretical value. Based on these results, VPc4 was confirmed to be the covalent attachment of the two *c*-type heme groups to the protein.

The optical spectra of VPc4 were measured with a Hitachi U3310 spectrophotometer (Hitachi, Tokyo) (Fig. 2A). In the reduced form, the α-, β-, γ (soret)-, and δ-absorption maxima peaks appeared at 552.0 nm, 523.0 nm, 416.0 nm, and 316.0 nm respectively, and the ratio of the α-peak and the β-peak (α/β) was 1.12. The α/β ratio of cyt c_4 is lower than that of the other *c*-type cyts. This lower ratio is one of characteristics of diheme cyts c_4. For the oxidized form of VPc4, the absorption maxima peaks of $\alpha+\beta$, γ (soret), and δ were 526.0 nm, 411.0 nm, and 360.0 nm respectively, and a shoulder peak at 697.0 nm, indicating His-Fe-Met coordination, was observed (Fig. 2A inset). In addition, the shoulder at 630 nm suggests that a high spin component was also present (Fig. 2A, inset). The high spin component has been observed in other cyts c_4[4–7] and is derived from N-terminal domain heme, which is engaged in low- and high-spin equilibrium.[13] Figure 2B and C shows the soret band of oxidized VPc4 in a pH range of 7.0–0.5. The soret band was shifted from 411.0 nm to 395.0 nm, and the two places of the isosbestic point of VPc4 were observed. This blue shift is indicative of conversion of the heme toward the high-spin form with disruption of heme six coordinate. The first isosbestic point at 402.0 nm appeared in a pH range of 7.5–4.0, and a second isosbestic point at 400.8 nm was found in a pH range of 3.6–0.5. Considering that N-terminal domain heme is a mixture of low- and high-spin forms in equilibrium, this result suggests that a disruption of the N-terminal domain six coordinated occurs in a pH range of 7.5–4.0 and that disruption of C-terminal domain six coordinate occurs at pH <3.6.

The circular dichroism (CD) spectrum of VPc4 was recorded with a Jasco J600 spectropolarimeter (Jasco,

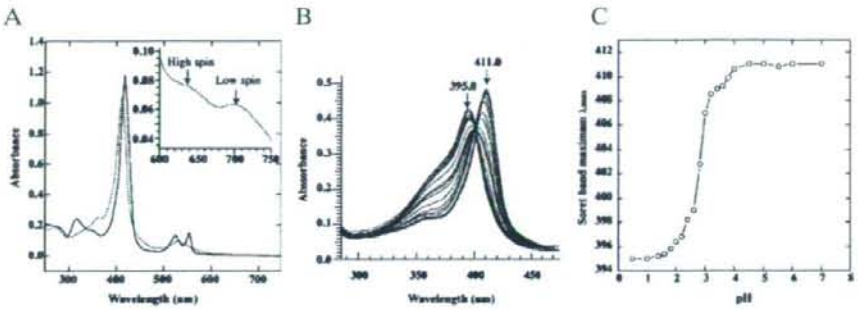


Fig. 2. UV/Visible Spectra of Reduced (Solid Line) and Oxidized (Broken Line) Forms of Recombinant VPc4 (A), Absorption Spectra (B), and Soret Bandshift (C) of VPc4 in the Soret Band Region at Different pH Levels.

A, absorption spectra of 4 μM cyt c_4 were measured in 10 mM sodium phosphate (pH 7.0) at 25 °C. The inset shows the 700 nm band of the oxidant form at 40 μM. Sodium dithionite and potassium ferricyanide were used as oxidant and reductant, respectively. B and C, absorption spectra of 2 μM oxidized cyt c_4 were measured in 20 mM buffers of different pH values between 7.5 and 0.5 at a constant ionic strength 100 mM (NaCl). The buffers were Tris (pH 7.5 and 7.0), MES (pH 6.5–5.5), acetate (pH 5.0–4.0), citrate (pH 3.6–2.6), and glycine (pH 2.4–0.5).

Tokyo) (Fig. 3A). The CD spectrum of VPc4 showed negative bands at 208 nm and 222 nm, and the α-helical content was estimated to be 45.6% the basis of an ellipticity of −32,000 deg cm^2/dmol for 100% helicity.[14] This α-helical content of the protein closely approximates those of *P. stutzeri* (47%) and *T. ferrooxidans* (50%) cyts c_4.[2,3]

Redox titrations were performed as described previously.[10] The redox data were analyzed with a theoretical curve based on the Nernst equation. The data points were fitted by a superposition of two $n = 1$ Nernst curves using $E^1{}_m = 230$ mV and $E^2{}_m = 320$ mV. The redox potentials of VPc4 were 237 mV (N-terminal domain heme) and 314 mV (C-terminal domain heme) at pH 7.0 (Fig. 3B). These redox potentials are equal to those of other cyts c_4, which have a characteristic high redox potentials of approximately 200 mV and 300 mV.[4–7] The biological function of VPc4 has been not yet determined, but the high redox potentials are indicative of a position close to the cbb_3-type terminal oxidase, because the redox potential of the terminal oxidase was higher than that of cyts c_4.[15]

In this study, we expressed a VPc4 of unknown function in *E. coli*, and determined physicochemical properties of the protein such as molecular mass, UV/vis, and CD spectra, and the redox potentials of the N-terminal and C-terminal domain hemes. This might facilitate the elucidation of the bacterial respiratory electron transfer chain and its regulation.

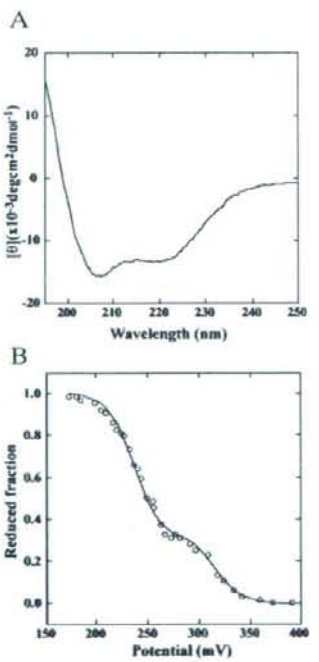


Fig. 3. Circular Dichroism Spectrum (A) and Redox Titration (B) of Recombinant VPc4.

A, CD spectrum was recorded at 25 °C in 10 mM sodium phosphate buffer (pH 7.0) with a Jasco J600 spectropolarimeter using 1-mm path length rectangular quartz cuvettes. B, The redox state of the protein was determined by changing the α-band absorption spectrum (552.0–558.6 nm) using a Hitachi U-3310 spectrophotometer. The smooth curve was drawn by the Nernst equation giving the best fit to the data using a Sigma Plot program.

Acknowledgment

We thank Dr. Yuji Haishima of the Division of Medical Devices of the Japanese, National Institute of Health Sciences for measurements in MALDI-TOF mass spectrometry. This work was supported in part by Nihon University College of Bioresource Sciences Research Fund for 2008.

References

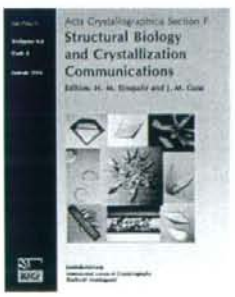
1) Thöny-Meyer, L., Biogenesis of respiratory cytochromes in bacteria. *Microbiol. Mol. Biol. Rev.*, **61**, 337–376 (1997).
2) Kadziola, A., and Larsen, S., Crystal structure of the dihaem cytochrome c_4 from *Pseudomonas stutzeri* determined at 2.2 Å resolution. *Structure*, **5**, 203–216 (1997).
3) Abergel, C., Nitschke, W., Malarte, G., Brushci, M., Claverie, J. M., and Giudici-Orticone, M. T., The structure of *Acidithiobacillus ferrooxidans* c_4-cytochrome: a model for complex-induced electron transfer tuning. *Structure*, **11**, 547–555 (2003).
4) Cavazza, C., Guidici-Orticoni, M. T., Nitschke, W., Appia, C., Bonnefoy, V., and Brusch, M., Characterisation of a soluble cytochrome c_4 isolated from *Thiobacillus ferrooxidans*. *Eur. J. Biochem.*, **242**, 308–314 (1996).
5) Tissieres, A., Purification, some properties and the specific biological activity of cytochromes c_4 and c_5 from *Azotobacter vinelandii*. *Biochem. J.*, **64**, 3286–3290 (1956).
6) Conrad, L. S., Kalsson, J., and Ulstrup, J., Electron transfer and spectral α-band properties of the di-heme protein cytochrome from c_4 from *Pseudomonas stutzeri*. *Eur. J. Biochem.*, **231**, 133–141 (1995).
7) Branca, R. M., Bodó, G., Várkonyi, G., Debreczeny, M., Ösz, J., and Bagyinka, C., Oxygen and temperature-dependent structural and redox changes in a novel cytochrome c_4 from the purple sulfur photosynthetic bacterium *Thiocapsa roseopersicina*. *Arch. Biochem. Biophys.*, **467**, 174–184 (2007).
8) Malarte, G., Leroy, G., Lojou, E., Abergel, C., Bruschi, M., and Giudici-Orticone, M. T., Insight into molecular stability and physiological properties of the diheme cytochrome CYC41 from the acidophilic bacterium *Acidithiobacillus ferrooxidans*. *Biochemistry*, **44**, 6471–6481 (2005).
9) Makino, K., Oshima, K., Kurokawa, K., Yokoyama, K., Uda, T., Tagomori, K., Iijima, Y., Najima, M., Nakano, M., and Yamashita, A., Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism distinct from that of *V. cholerae*. *Lancet*, **361**, 743–749 (2003).
10) Satoh, T., Itoga, A., Isogai, Y., Kurihara, M., Yamada, S., Natori, M., Suzuki, N., Suruga, K., Kawachi, R., Arahira, M., Nishio, T., Fukazawa, C., and Oku, T., Increasing the conformational stability by replacement of heme axial ligand in *c*-type cytochrome. *FEBS Lett.*, **531**, 543–547 (2002).
11) Pollock, W. B., and Voordouw, G., Aerobic expression of the cyf gene encoding cytochrome *c*-553 from *DesulfoVibrio Vulgaris Hildenborough* in *Escherichia coli*. *Microbiology*, **140**, 879–887 (1994).
12) Giulia, D. R., Gianantonio, B., Marco, B., Francesca, D. R., Antonio, R., Maria, L. T., and Marco, S., Cloning, expression and physicochemical characterization of a di-heme cytochrome c_4 from the psychrophilic bacterium *Pseudoalteromonas haloplanktis* TAC 125. *J. Biol. Inorg. Chem.*, **13**, 789–799 (2008).
13) Andersen, N. H., Nørgaard, A., Jensen, T. J., and Ulstrup, J., Sequential unfolding of the two-domain protein *Pseudomonas stutzeri* cytochrome c_4, *J. Inorg. Biochem.*, **88**, 316–327 (2002).
14) Pace, C. N., Shirly, B. A., and Thompson, J. A., "Protein Structure: A Practical Approach," IRL, Oxford, pp. 311–330 (1989).
15) Veríssimo, A. F., Sousa, F. L., Baptista, A. M., Teixeira, M., and Pereira, M. M., Thermodynamic redox behavior of the heme centers of cbb_3 heme-copper oxygen reductase from *Bradyrhizobium japomicum*. *Biochemistry*, **46**, 13245–13253 (2007).

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091
Editors: **H. M. Einspahr and M. S. Weiss**

Structure of the ligand-binding domain of rat VDR in complex with the nonsecosteroidal vitamin D_3 analogue YR301

Shinji Kakuda, Kazuhisa Okada, Hiroshi Eguchi, Kazuya Takenouchi, Wataru Hakamata, Masaaki Kurihara and Midori Takimoto-Kamimura

Acta Cryst. (2008). F**64**, 970–973



Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structure of the ligand-binding domain of rat VDR in complex with the nonsecosteroidal vitamin D_3 analogue YR301

Shinji Kakuda,[a]* Kazuhisa Okada,[a] Hiroshi Eguchi,[a] Kazuya Takenouchi,[a] Wataru Hakamata,[b] Masaaki Kurihara[b] and Midori Takimoto-Kamimura[a]

[a]Teijin Institute for Biomedical Research, Japan, and [b]Division of Organic Chemistry, National Institute of Health Sciences, Japan

Correspondence e-mail: s.kakuda@teijin.co.jp

Received 16 May 2008
Accepted 19 August 2008

PDB Reference: vitamin D receptor ligand-binding domain–YR301 complex, 2zfx, r2zfxsf.

Vitamin D receptor (VDR) is a ligand-inducible hormone receptor that mediates $1\alpha,25(OH)_2D_3$ action, regulating calcium and phosphate metabolism, induces potent cell differentiation activity and has immunosuppressive effects. Analogues of $1\alpha,25(OH)_2D_3$ have been used clinically for some years. However, the risk of potential side effects limits the use of these substances. LG190178 is a novel nonsecosteroidal ligand for VDR. (2*S*)-3-[4-(3-[4-[(2*R*)-2-hydroxy-3,3-dimethylbutoxy]-3-methylphenyl]pentan-3-yl)-2-methylphenoxy] propane-1,2-diol (YR301) is the only one of the four evaluated stereoisomers of LG190178 to have strong activity. To understand the strong activity of YR301, the crystal structure of YR301 complexed with the rat VDR ligand-binding domain (VDR LBD) was solved at 2.0 Å resolution and compared with the structure of the VDR LBD–$1\alpha,25(OH)_2D_3$ complex. YR301 and $1\alpha,25(OH)_2D_3$ share the same position and the diethylmethyl group occupies a similar space to the *C* and *D* rings of $1\alpha,25(OH)_2D_3$. YR301 has two characteristic hydroxyl groups which contribute to its potent activity. The first is 2′-OH, which forms hydrogen bonds to the NE2 atoms of both His301 and His393. The other is 2-OH, which interacts with Ser233 OG and Arg270 NH1. These two hydroxyl groups of YR301 correspond exactly to 25-OH and 1-OH, respectively, of $1\alpha,25(OH)_2D_3$. The terminal hydroxyl group (3-OH) of YR301 is directly hydrogen bonded to Arg270 and also interacts indirectly with Tyr232 OH and the backbone NH of Asp144 *via* water molecules. Additional derivatization of the terminal hydroxyl group using the positions of the water molecules might be useful for the design of more potent compounds.

1. Introduction

The vitamin D receptor (VDR) is a member of the nuclear receptor superfamily, with which it shares structural and functional similarity (Mangelsdorf *et al.*, 1995). It contains three principal domains, a variable N-terminal domain that contains a ligand-independent activation function, a central highly conserved DNA-binding domain and a large C-terminal ligand-binding domain (VDR LBD), and a short linker between the DNA-binding domain and the VDR LBD. VDR is a ligand-dependent transcriptional regulator that acts in several physiological processes and is an important target for a wide spectrum of clinical applications such as osteoporosis, psoriasis and secondary hyperparathyroidism.

The function of VDR is regulated by the binding of vitamin D, $1\alpha,25(OH)_2D_3$, which induces a conformational change of LBD that allows the recruitment of a coactivator from the p160 (Kalkhoven *et al.*, 1998) or the DRIP/TRAP (Rachez *et al.*, 1999) families. $1\alpha,25(OH)_2D_3$ regulates calcium and phosphate metabolism, induces potent cell differentiation activity and has immunosuppressive effects (Bouillon *et al.*, 1995; Lieberherr, 1987; Hendy *et al.*, 2006). Most analogues of $1\alpha,25(OH)_2D_3$ have been synthesized with the goal of improving the biological profile of the natural ligand in order to decrease its hypercalcaemic effects for therapeutic application (Stein & Wark, 2003).

Vitamin D_3 analogues can be classified as steroidal and nonsteroidal based on their structures. The endogenous $1\alpha,25(OH)_2D_3$ and modified analogues have some shortcomings that limit their general use, mainly their potential for side effects and their mode of admini-

stration (Mizwicki & Norman, 2003). Nonsecosteroidal vitamin D_3 analogues have been under investigation for many years.

YR301 is a member of a novel class of VDR ligands that have the potential to provide improved therapeutic benefits while reducing the risk of side effects. In *in vitro* transcriptional assays and VDR affinity assays, only the (2*S*,2′*R*)-isomer (YR301) exhibited potent transcriptional activity (Hakamata *et al.*, 2008). To gain further insight into the structure–activity relationships of vitamin D and to determine the basis of the observed selectivity effect on a molecular level, we solved the crystal structure of rat VDR LBD in complex with YR301. A comparison of this structure with the structure of rat VDR LBD complexed with $1\alpha,25(OH)_2D_3$ (Vanhooke *et al.*, 2004) provides clues to achieving tissue selectivity and obtaining more effective potential therapeutic agents.

2. Material and methods

Rat VDR LBD (residues 116–423) lacking the flexible insertion region (residues 165–207) was amplified by polymerase chain reaction technology using the appropriate primers and inserted between the *Nde*I site and the *Not*I site of pET28a (Novagen). The resulting fusion protein contained a His tag and an N-terminal thrombin cleavage site. The constructs were produced as described by Vanhooke *et al.* (2004) and were then transformed into *Escherichia coli* strain BL21 (DE3) (Novagen). Expression was carried out in LB medium containing 20 μg ml^{-1} kanamycin with 500 μ*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) at 293 K for about 16 h before harvesting by centrifugation. The cell pellets were suspended in 20 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl, 10% glycerol, 5 m*M* 2-mercaptoethanol (2-ME) and and were lysed by 20 mg ml^{-1} lysozyme. The fused protein was purified by affinity chromatography using an Ni-Sepharose column. The rat VDR LBD was eluted using a buffer consisting of 20 m*M* Tris–HCl pH 8.0, 10% glycerol, 100 m*M* NaCl, 5 m*M* 2-ME, 400 m*M* imidazole. Cleavage with 50 units of thrombin per millilitre was performed overnight at 277 K. After thrombin treatment, an LBD consisting of residues 116–423 with a tag remnant of sequence GSHM attached to the N-terminus was produced. The rat VDR LBD was further purified using a gel-filtration column (GE Biosciences) equilibrated with 10 m*M* Tris–HCl pH 7.0, 100 m*M* NaCl and 10 m*M* DTT (buffer I) before crystallization.

YR301 was synthesized as previously described by Hakamata *et al.* (2008). Crystallization experiments were performed using the hanging-drop vapour-diffusion method. The ligands were added to aliquots of the purified protein in a fivefold molar excess. We have cocrystallized the rat VDR LBD with YR301 and a synthetic peptide containing the L*XX*LL sequence of the coactivator DRIP 205. The synthetic peptide, with amino-acid sequence KNHPMLMNLLKDN-NH_2, was added to the rat VDR LBD/YR301 in a fivefold molar excess over the protein. Crystallization conditions were similar to those used for the VDR LBD–$1\alpha,25(OH)_2D_3$ complex crystals; 1 μl of protein solution (10 mg ml^{-1} in buffer I) was mixed with an equal

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Space group	$P2_12_12$
Unit-cell parameters (Å)	
a	44.22
b	47.33
c	136.43
Data collection	
Beamline	PF NW12
Wavelength (Å)	1.000
Resolution (Å)	50.00–2.00
Total No. of reflections	292936
Unique reflections	20439
R_{merge}†	0.088 (0.341)
Completeness (%)	100.0 (99.7)
Multiplicity	7.7 (7.2)
Average $I/\sigma(I)$	31.2 (6.49)
Refinement statistics	
R factor‡ (%)	20.6
R_{free}‡ (%)	25.8
R.m.s. deviation from ideal values	
Bond lengths (Å)	0.017
Bond angles (°)	1.59

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean intensity of i reflections with intensities $I_i(hkl)$ and common indices h, k and l. ‡ R factor $= \sum_{hkl} ||F_{obs}| - k|F_{calc}|| / \sum_{hkl} |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors. R_{free} is calculated for a randomly chosen 5% of reflections and R factor is calculated for the remaining 95% of reflections.

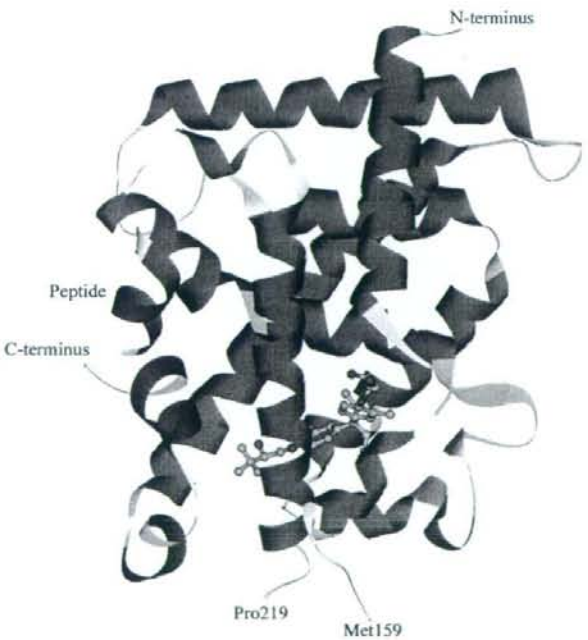


Figure 1
Overall structure of the VDR LBD–YR301 complex structure, showing the position of YR301 (shown as a ball-and-stick model).

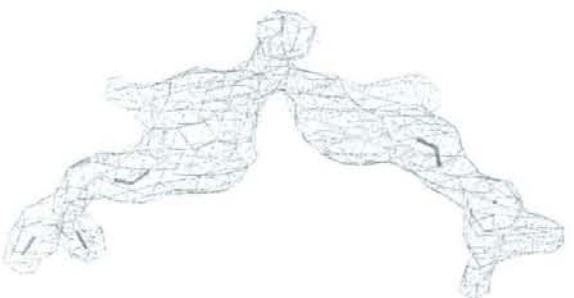


Figure 2
Conformation of the bound YR301. YR301 is shown in the $F_o - F_c$ electron-density OMIT map contoured at 1.0σ.

volume of the reservoir solution, which contained 100 m*M* MOPS pH 7.0, 200 m*M* ammonium citrate, 20% PEG 4000 and 4% 2-propanol, and equilibrated against 1 ml reservoir solution (Vanhooke *et al.*, 2004). Single crystals grew to suitable dimensions in 2–4 d. Crystals were cryoprotected in 30% glycerol and cooled at 79 K and X-ray data were collected on beamline NW12 at Photon Factory. The data were processed using the *HKL*-2000 software package (Otwinowski & Minor, 1997).

We carried out molecular replacement using *MOLREP* (Vagin & Teplyakov, 1997) from *CCP*4 (Collaborative Computational Project, Number 4, 1994) with the coordinates of the VDR LBD–$1\alpha,25(OH)_2D_3$ complex [PDB code 1rk3; the solvent molecules and $1\alpha,25(OH)_2D_3$ were removed] as the initial model. Refinement was carried out using the program *REFMAC* (Murshudov *et al.*, 1997). A sample containing a random 5% of the total reflections in the data set was excluded for R_{free} calculations. After rigid-body refinement, electron density for $1\alpha,25(OH)_2D_3$ and the YR301 ligand was clearly constructed using *Coot* (Emsley & Cowtan, 2004). In the final refinement at 2.0 Å, the crystallographic R factor and R_{free} were 20.6% and 25.8%, respectively, with good stereochemistry. Statistics of the data collection and final structure are summarized in Table 1. Figures were produced using *DS Visualizer* (Accelrys; http://accelrys.co.jp/).

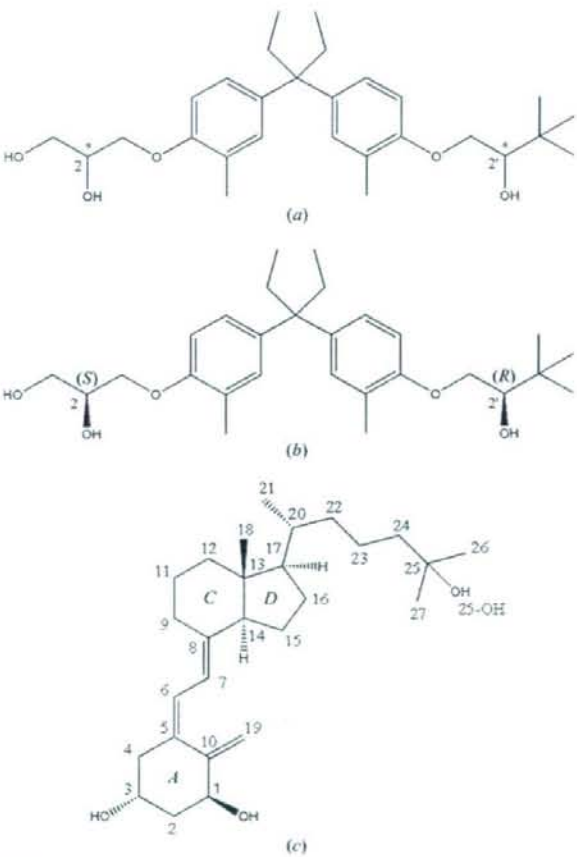


Figure 3
Chemical structures of (*a*) LG190178, (*b*) YR301 [(2*S*)-3-[4-(3-[4-[(2*R*)-2-hydroxy-3,3-dimethylbutoxy]-3-methylphenyl]pentan-3-yl)-2-methylphenoxy]propane-1,2-diol] and (*c*) $1\alpha,25(OH)_2D_3$ ($1\alpha,25$-dihydroxy vitamin D_3).

3. Results and discussion

The overall structure of the VDR LBD–YR301 complex is very similar to that of the VDR LBD–$1\alpha,25(OH)_2D_3$ complex and we obtained a good electron-density map of the protein from Leu116 to Ser423 without the residues in the loop between 160 and 218 at 2.0 Å resolution. When compared with the VDR LBD–$1\alpha,25(OH)_2D_3$ complex, the root-mean-squared deviation (r.m.s.d.) on C^{α} atoms is 1.25 Å for the VDR LBD–YR301 complex. The two protein backbones are superposable. The structure of the VDR LBD–YR301 complex and the electron density for YR301 are shown in Figs. 1 and 2, respectively.

YR301 is a nonsecosteroidal vitamin D_3 analogue which embeds in the ligand-binding pocket in the same position as $1\alpha,25(OH)_2D_3$ in the VDR LBD–$1\alpha,25(OH)_2D_3$ complex. The chemical structures of YR301 and $1\alpha,25(OH)_2D_3$ are shown in Fig. 3. YR301 and $1\alpha,25(OH)_2D_3$ share the same position and the diethylmethyl group occupies a similar space to the *C* and *D* rings of $1\alpha,25(OH)_2D_3$, as shown in Fig. 4. The hydrophobic pocket is filled by the diethylmethyl group, which interacts with Trp282.

The interactions between the receptor and the ligand involve both hydrophobic and electrostatic contacts. A schematic representation of the interactions is shown in Fig. 5. The 2′-OH group of YR301 hydrogen bonds to His301 NE2 (2.71 Å) and His393 NE2 (2.74 Å) of VDR LBD. The 2-OH group of YR301 interacts with Ser233 OG (2.79 Å) and Arg270 NH1 (2.76 Å). In comparison, the 25-OH group of $1\alpha,25$-$(OH)_2D_3$ interacts with His301 NE2 (2.85 Å) and His393 NE2 (2.70 Å) and the 1-OH group interacts with Ser233 OG (2.76 Å) and Arg274 NH1 (2.83 Å). The 2′-OH and 2-OH groups of YR301 play exactly the same roles as the 25-OH and 1-OH groups of the natural ligand. Although the 3-OH group of $1\alpha,25(OH)_2D_3$ forms

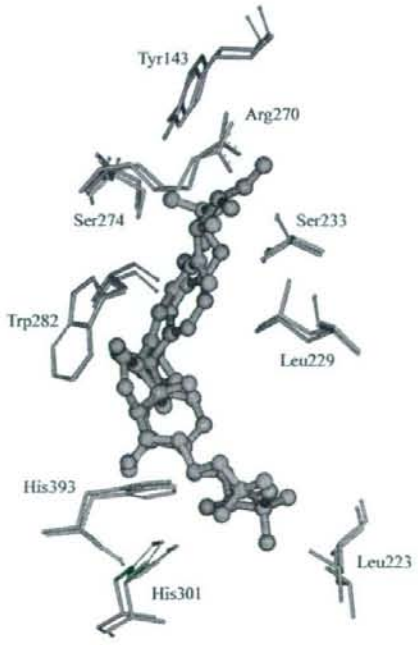


Figure 4
Ligand-binding pocket of the structure of the VDR LBD–YR301 complex superimposed on the structure of the VDR LBD–$1\alpha,25(OH)_2D_3$ complex. YR301 is shown in red and $1\alpha,25(OH)_2D_3$ is shown in blue.

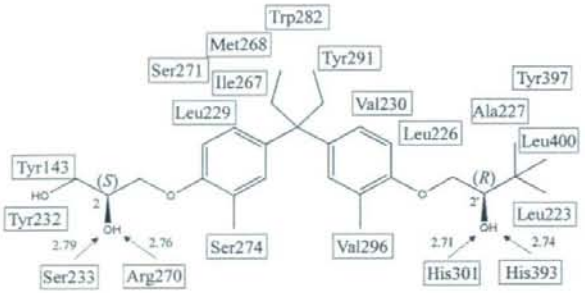


Figure 5
Schematic representation of the interactions between YR301 and VDR at a distance cutoff of 4.0 Å. Arrows correspond to hydrogen bonds between ligand and protein residues. The numbers show the distances of these hydrogen bonds in Å.

a hydrogen bond to Tyr139 OH and Ser274 OG, YR301 does not interact with those amino acids. It is probable that additional interactions with those amino acids might increase the binding affinity (Hakamata *et al.*, 2008). In the present study, we find that the terminal hydroxyl group of YR301 forms three direct hydrogen bonds to Arg270 NH1 and interacts indirectly with Tyr232 OH and Asp144 NH *via* mediated water molecules (W55, W57 and W58), despite the disorder of W57 as shown in Fig. 6. We think that the hydrogen bonds of the terminal hydroxyl groups anchor the ligands to the receptor in VDR LBD–YR301. The nonsecosteroidal vitamin D_3 analogue YR301 is a mimic of the natural ligand $1\alpha,25(OH)_2D_3$. Additional derivatization of the terminal hydroxyl group using the positions of the water molecules might be useful for the design of more potent compounds.

In this communication, we present the crystal structure of the VDR LBD–YR301 complex and compare it with the structure of the VDR LBD–$1\alpha,25(OH)_2D_3$ complex. YR301 has been proven to be a selective vitamin D receptor with reduced effects on bone compared with other receptors and the structural information described here provides further details of its ligand-binding characteristics and provides some new insights into its ligand-binding requirements, which may aid in the discovery of more selective and effective potential therapeutic agents.

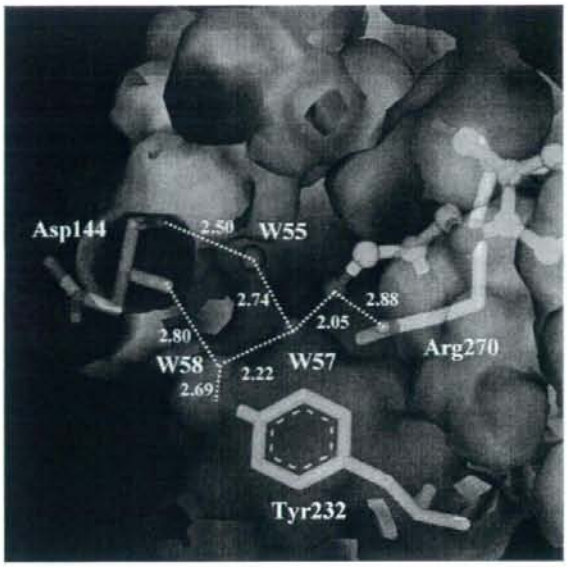


Figure 6
Interaction between the terminus of YR301 and VDR LBD. Negative charge is shown in red and positive charge in blue. The C atoms of VDR side chains are shown in grey and those of YR301 are shown in green. Hydrogen bonds are denoted by dotted lines and the numbers show their distances in Å.

References

Bouillon, R., Okamura, W. H. & Norman, A. W. (1995). *Endocr. Rev.* **16**, 200–257.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Hakamata, W., Sato, Y., Okuda, H., Honzawa, S., Saito, N., Kishimoto, S., Yamashita, A., Sugiura, T., Kittaka, A. & Kurihara, M. (2008). *Bioorg. Med. Chem. Lett.* **18**, 120–123.

Hendy, G. N., Hruska, K. A., Mathew, S. & Goltzman, D. (2006). *Kidney Int.* **69**, 218–223.

Kalkhoven, E., Valentine, J. E., Heery, D. M. & Parker, M. G. (1998). *EMBO J.* **17**, 232–243.

Lieberherr, M. (1987). *Biol. Chem.* **262**, 13168–13173.

Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schütz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P. & Evans, R. M. (1995). *Cell*, **83**, 835–839.

Mizwicki, M. T. & Norman, A. W. (2003). *J. Bone Miner. Res.* **18**, 795–806.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Rachez, C., Lemon, B. D., Suldan, Z., Bromleigh, V. & Gamble, M. (1999). *Nature (London)*, **398**, 824–828.

Stein, M. S. & Wark, J. D. (2003). *Expert Opin. Invest. Drugs.* **12**, 825–840.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Vanhooke, J. L., Benning, M. M., Bauer, C. B., Pike, J. W. & DeLuca, H. F. (2004). *Biochemistry*, **43**, 4101–4110.

最近の研究から

輸送レクチン VIP36 による高マンノース型糖鎖認識の構造基盤

佐藤匡史[1]，Nathan P. Cowieson[2]，袴田　航[3]，井手尾浩子[4,5]，福島慶子[4,5]，栗原正明[3]，加藤龍一[1]，山下克子[4,5]，若槻壮市[1]

[1]高エネ研・物構研・構造生物学研究センター，[2]Queensland 大学・Institute for Molecular Bioscience，[3]国立医薬品食品衛生研究所・有機化学部，[4]東京工業大学・イノベーション研究推進体，[5]科学技術振興事業機構 CREST

Structural Basis for Recognition of High Mannose Type Glycan by Transport Lectin VIP36

Tadashi Satoh[1], Nathan P. Cowieson[2], Wataru Hakamata[3], Hiroko Ideo[4,5], Keiko Fukushima[4,5], Masaaki Kurihara[3], Ryuichi Kato[1], Katsuko Yamashita[4,5] and Soichi Wakatsuki[1]

[1]Structural Biology Research Center, Photon Factory, Institute of Materials Structure Science, High Energy Accelerator Research Organization [2]Institute for Molecular Bioscience, University of Queensland, Australia [3]Division of Organic Chemistry, National Institute of Health Sciences [4]Innovative Research Initiatives, Tokyo Institute of Technology [5]CREST, Japan Science and Technology Agency

1．はじめに

細胞の中には膜によって仕切られた細胞内小器官（オルガネラ）があり，それぞれのオルガネラは専門の機能を担っている。これらオルガネラは「輸送小胞」を介して，相互に積み荷（タンパク質や脂質）のやりとりをしている。この輸送小胞による積み荷タンパク質の輸送は，厳密にコントロールされていて，送り手側のオルガネラに留まるべきタンパク質と受け手側のオルガネラへと送り出されるタンパク質の選別が行われている。細胞質領域をもつ膜タンパク質は，膜の細胞質側に局在している輸送小胞と直接結合でき，選別・輸送される [1]。一方，細胞質側の領域を持たない可溶性のタンパク質の選別・積み込みは，積み荷タンパク質と輸送小胞を同時に結合して繋ぐ膜貫通型の積み荷タンパク質レセプターがこれを担っている [2]。

小胞体で生合成されてくるタンパク質の多くは *N* 型糖鎖修飾を受けるが，輸送小胞への選別と *N* 型糖鎖には密接な関係があることがわかっている。すなわち，マンノース 6 リン酸受容体（MPR），VIP36，ERGIC-53，Emp46/47p などが積み荷タンパク質の *N* 型糖鎖を荷札として，選別・積み込みを行う積み荷タンパク質レセプター（輸送レクチン）として機能していることが明らかにされている [3-5]。その中でも P-type（mannose-6-phosphate）レクチンファミリーに分類されるカチオン依存性および非依存性 MPR の構造生物学的研究は進んでおり，それぞれ糖鎖結合型および非結合型の構造が明らかにされている [6-9]。一方，L-type（leguminous-like）レクチンファミリーでは，これまでにラット由来 ERGIC-53 の Ca^{2+} 結合型および非結合型，Emp46p の K^+ 結合型および非結合型，Emp47p の金属非結合型の構造が明らかにされているが [10-12]，積み荷糖タンパク質および糖鎖との複合体構造は解析されていない。また近年，ERGIC-53 の遺伝子変異により，糖タンパク質である血液凝固因子（第 V 因子および第 VIII 因子）の小胞体からゴルジ体への輸送に機能的障害を生じ，血友病と同様の出血性症状を呈する疾病が引き起こされることが明らかにされた [13]。従って，これら輸送レクチンの立体構造研究は，糖タンパク質輸送に関わる疾患についての理解を深め，最終的にはその治療の道へと繋げるための重要な課題の一つである。高エネルギー加速器研究機構・物質構造科学研究所の若槻壮市教授を中心とする構造生物学研究センターのグループは，東京工業大学の山下克子教授，国立医薬品食品衛生研究所の袴田航主任研究官（（現）日本大学生物資源科学部，専任講師）のグループとの共同研究で，VIP36 と糖鎖との複合体の高分解能 X 線結晶構造解析に成功した [14]。本稿では，高分解能を示す単結晶が得られるまでの経緯も併せて，Ca^{2+} 依存性の輸送レクチン VIP36 について紹介する。

2．VIP36 の結晶化，回折データ収集および構造決定

イヌ由来の VIP36 は，グルタチオン -S- トランスフェラーゼ（GST）との融合タンパク質として大腸菌で発現させ，各種クロマトグラフィーにより精製を行った。結晶化はハンギングドロップ蒸気拡散法で行った。まず最初に，ERGIC-53 および Emp46/47p の結晶構造 [10-12] を基に糖鎖認識ドメイン（CRD）に対応する 51-278 残基を用いたコンストラクトを作成した。結晶は 20℃ において短時間で様々な結晶化条件で得られたものの（Fig. 1a），いずれもモザイク性が高く多結晶であったため，回折データ収集に適したものではなかった。このコンストラクトを用い，さらに他の結晶化条件の検討を行ったが，いずれも全く改善は認められなかった。我々はこの実験と並行して，VIP36 が C 末端に短い stalk ドメインを持つことから，CRD に加えて様々な長さの stalk ドメインを持つコンストラクト（51-286，51-295，51-301，51-322 残基）を作成し，結晶化を試みていた。しかしながら，数週間に渡り 4℃ および 20℃ で仕込んだ結晶化プレートの観察を続けたが，結晶を確認することが出来なかった。4℃ で仕込んだ結晶

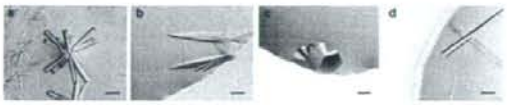


Figure 1
Crystals of VIP36. (a) The crystals of VIP36 CRD (residues 51-278). (b) The crystals of Ca^{2+}-bound VIP36 CRD with stalk domain (residues 51-301). (c) The crystals of Ca^{2+}/Man_2-bound VIP36 CRD with stalk domain (residues 51-301). (d) The crystals of Ca^{2+}/$Man_3GlcNAc$-bound VIP36 CRD with stalk domain (residues 51-301). A black bar indicates 0.1 mm.

化プレート 20 枚，計 1920 条件を定期的に観察するのはハードな実験であったことと，CRD（51-278 残基）の結晶化条件の検討に重点を置いていたため，仕込んでから 1 ヶ月後の観察を怠っていた。その後，筆者は一向に改善されない CRD の結晶に嫌気がさし，4℃ で仕込んだ結晶化プレートを約 3 ヶ月ぶりに観察した。その結果，51-301 残基のコンストラクトのみにおいて，全く新しい条件で結晶らしき物体を発見し，その後結晶化条件の最適化を行ったところ，数日で美しい結晶を得る事が出来た（Fig. 1b）。ドメインの結晶化を進めるに当たって，筆者は教訓として「コンストラクトは削るだけでなく，延ばしたものも試す」，「4℃ での結晶観察は怠らない」の 2 点を再認識させられた。ちなみに現在，高エネ研・構造生物学研究センターには低温室での結晶観察自動化装置が導入されており，タンパク質結晶の量産に役立っている。

VIP36 の Ca^{2+} 結合型の結晶は，PF-AR NW12A ビームラインを用いて 1.80 Å 分解能で回折データを収集した。構造解析は，ERGIC-53 の Ca^{2+} 結合型の CRD をサーチモデルとして分子置換法で行った。空間群は *C*2 で，非対称単位中に 2 分子の VIP36 が存在していた。Ca^{2+} 非結合型の結晶は，Ca^{2+} 非存在下では結晶が全く得られなかったことから，Ca^{2+} 結合型の結晶に 10 mM EDTA を浸漬して調製した。回折データは PF-AR NW12A を用いて 2.10 Å 分解能で収集した。これまでに VIP36 は生化学的実験から *N* 型糖鎖の高マンノース型糖鎖，特に Man-α1,2-Man 結合を多く持つ D1 アームを認識することが明らかにされている [15,16]。そこで我々は，高マンノース型糖鎖の D1 アームおよびその近傍の糖鎖構造に一致する様々な長さと種類の糖鎖を用いて複合体の解析を行った。Ca^{2+}/Mannose（Man）結合型は Ca^{2+} 結合型の結晶に 50 mM Man を浸漬して調製した。Ca^{2+}/Man-α1,2-Man（Man_2）および Ca^{2+}/Man-α1,2-Man-α1,3-Man-β1,4-GlcNAc（$Man_3GlcNAc$）結合型は，共結晶化（モル比 1 : 10）により調製した（Fig. 1c, d）。一方，Man-α1,2-Man-α1,2-Man，Man-α1,2-Man-α1,3-Man および Man-α1,3-Man-β1,4-GlcNAc 糖鎖を用いての共結晶化も試みたが，回折データの収集に適した結晶を得る事が出来なかった。Ca^{2+}/Man_2 結合型の格子常数は Ca^{2+} および Ca^{2+}/Man 結合型のものとほぼ同じであった。一方，Ca^{2+}/$Man_3GlcNAc$ 結合型の空間群および格子常数は Ca^{2+} および Ca^{2+}/Man_{1-2} 結合型のものと全く異なっており，空間群は $P2_12_12_1$ で非対称単位中に 5 分子の VIP36 が存在していた（Table 1, Appendix 参照）。Ca^{2+}/ 糖鎖結合型の回折データの収集は全て PF BL-5A で行った。Table 1 に結晶学的データ，回折データおよび精密化の統計値を示す。

3．糖鎖認識ドメイン（CRD）と stalk ドメイン

VIP36 の CRD の全体構造は，L-type レクチンでよく見られる凹型および凸型 β シートからなる β サンドイッチ構造から構成されていた（Fig. 2a）。ERGIC-53 および Emp46/47p の CRD と比較してみると，アミノ酸配列の同一性は 18-34% とさほど高くないが，その全体構造はよく似ていた（root mean square deviation = 0.8-1.2 Å）。一方 stalk ドメインは，大部分はループ構造からなり，二次

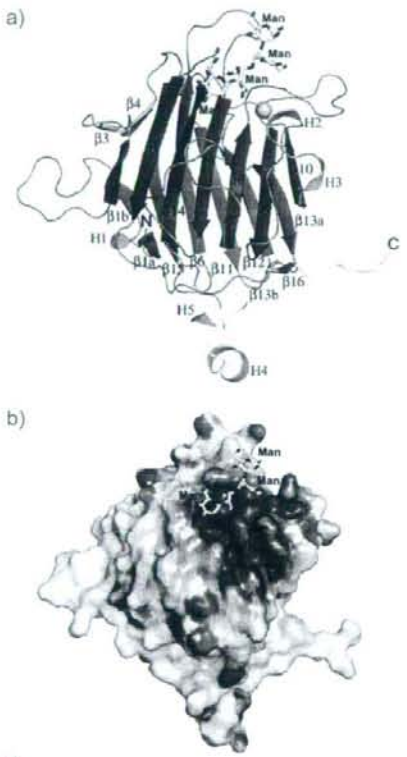


Figure 2
Overall structure of the exoplasmic/luminal domain of VIP36. The secondary structures are highlighted (β-strands belonging to the concave β-sheets, *red*; β-strands belonging to convex β-sheets, *blue*; β-strands belonging to β-hairpin, *cyan*; β-strands belonging to the short β-sheet formed between the stalk domain and one of the loops of the CRD, *magenta*; helices, *green*), and the loops of the CRD and stalk domain are colored *gray* and *white*, respectively. The bound Ca^{2+} is shown as a *pink sphere*. The bound oligomannoses are superimposed from the VIP36 complex structures with Man-α-1,2-Man (Man_2) and Man-α-1,2-Man-α-1,3-Man (Man_3) and are shown as a *yellow stick* model. The reducing-end mannose residue in the Man_3-bound form is omitted because its position is almost the same as that of the $Man_3GlcNAc$-bound form. The surface models (*b*) are shown in the same orientations as in (*a*) and colored according to the electrostatic surface potential (*blue*, positive; *red*, negative; scale from -10 to +10 *kTle*).

構造はαヘリックス（H4）に続き，3_{10}ヘリックス（H5），CRDのβ13bと逆平行βシートを形成するβストランド（β16）が存在していた。またCa^{2+}およびCa^{2+}/Man_{1-2}結合型において，stalkドメインは結晶学的な対称分子と水素結合を介して相互作用し，主に逆平行βシート（297-299残基）を形成していた。恐らくこのstalkドメインを介した結晶中での相互作用がVIP36結晶化成功の鍵であったと思われる。

ERGIC-53およびEmp46/47pのstalkドメインは約140-200残基からなり，coiled-coil構造を持つと推測され，多量体を形成することが明らかになっている[17,18]。一方，VIP36のstalkドメインは約40残基でERGIC-53やEmp46/47pのものと比較すると約100-160残基短く，coiled-coil構造を持たないと予想されている。また，超遠心分析法および化学架橋剤を用いた実験から，VIP36は生体内において単量体で存在すると推測されている[19]。実際，我々もゲルろ過クロマトグラフィーによりVIP36が単量体で存在し，結晶構造解析によりstalkドメインの半分にあたる279-301残基はcoiled-coil構造を形成しないことを明らかにした。以上の結果から，生体内においてVIP36はERGIC-53やEmp46/47pとは異なり単量体で機能することが強く示唆された。

4．Ca^{2+}依存的に糖鎖と結合

VIP36のCa^{2+}結合部位は，豆科レクチンやERGIC-53と同じ様に凹んだβシート上の負に荷電したポケットに存在していた（Fig. 2b）。しかしながら，豆科レクチンやERGIC-53とは異なりVIP36の金属結合部位はCa^{2+}の1箇所のみであった。豆科レクチンやERGIC-53では，VIP36で見られたCa^{2+}結合部位に加えて，2つ目の金属結合部位が存在する（Fig. 3）。豆科レクチンではMn^{2+}（S1），ERGIC-53ではCa^{2+}（M1）がさらに結合し，1つ目のCa^{2+}結合部位（S2もしくはM2）の安定化に寄与しているが，それらの金属結合部位は豆科レクチンとERGIC-53間で保存されていない。豆科レクチンでは，1つ目のCa^{2+}（S2）が糖鎖リガンドと相互作用するアミノ酸側鎖を固定させていることが結晶構造解析により明らかにされている[20]。VIP36は1分子のCa^{2+}としか結合しない新奇のL-typeレクチンであると言える。

これまでに生化学的実験により，ERGIC-53およびVIP36はCa^{2+}依存性の輸送レクチンであることが示唆されているが[16,21]，それらの詳細なメカニズムは明らかにされていない。そこで我々は，Ca^{2+}結合型に加えてCa^{2+}非結合型の構造解析を行い，Ca^{2+}が結合することによりCa^{2+}と（直接もしくは水分子を介して）結合しているAsp131，Asn166，His190の側鎖が動くことを明らかにした（Fig. 4）。後で詳しく述べるが，Ca^{2+}/糖鎖結合型の構造では豆科レクチンで報告された構造の様に，Ca^{2+}が糖鎖リガンドと相互作用するアミノ酸側鎖（Asp131，Asn166，His190）を固定させており（Fig. 5），VIP36のCa^{2+}依存性の糖鎖認識メカニズムが詳細に明らかになった。

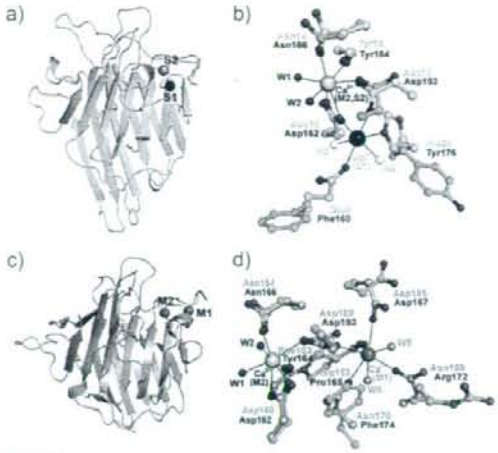


Figure 3
(a) Overall structure of ConA monomer. The bound Mn^{2+} (S1) and Ca^{2+} (S2) are shown as *large black* and *gray spheres*, respectively. (b) Comparison between VIP36 (*green*) and ConA (*cyan*) metal-binding site structures. Residues of VIP36 and ConA are labeled in *black* and *cyan*, respectively. The Ca^{2+} in VIP36 is shown as a *large pink sphere*. Because the position of Ca^{2+} at the S2 site in ConA is almost the same as in VIP36, it is not shown. Water molecules found in the S1 site of ConA are shown as *small white spheres* and are labeled *W3* and *W4*. (c) Overall structure of ERGIC-53 CRD. The bound Ca^{2+} are shown as *large magenta spheres* (M1 and M2). (d) Comparison between VIP36 (*green*) and ERGIC-53 (*purple*) Ca^{2+}-binding site structures. Because the position of Ca^{2+} at the M2 site in ERGIC-53 is almost the same as in VIP36, it is not shown. Water molecules found in the M1 site of ERGIC-53 are shown as *small orange spheres* and are labeled *W5* and *W6*. Residues involved in the metal binding are shown as *ball-and-stick models*.

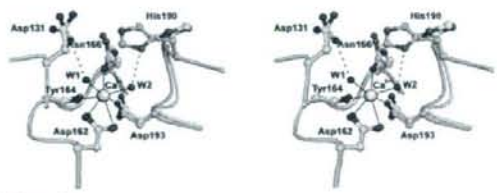


Figure 4
Ca^{2+}-binding site and its conformational changes upon Ca^{2+} binding of VIP36. The Ca^{2+}-bound and metal-free structures are shown in stereo and colored in *green* and *cyan*, respectively. Residues coordinating Ca^{2+} and those with notable conformational changes are shown in *ball-and-stick* models. Water molecules are labeled *W1* and *W2*. Ca^{2+}-coordinating bonds are *solid lines*, and hydrogen bonds are *dotted lines*.

5．高マンノース型糖鎖のD1アームを特異的に認識

これまでの生化学的実験により，VIP36が高マンノース型糖鎖のD1アームを認識することが明らかにされている[15,16]。そこで我々は高マンノース型糖鎖のD1アームおよびその近傍の糖鎖構造に一致するMan，Man_2および$Man_3GlcNAc$結合型の立体構造解析を行った。$Man_3GlcNAc$結合型では，VIP36はGlcNAc残基を認識していないため，電子密度は観測されなかった。糖鎖リガンドは，予