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

High-resolution structure of the recombinant sweet-tasting protein thaumatin I

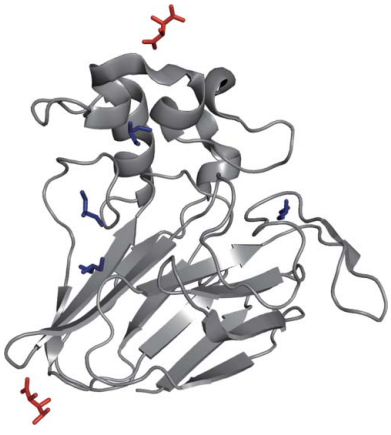
Tetsuya Masuda,[a,b]* Keisuke Ohta,[a,b] Bunzo Mikami[c] and Naofumi Kitabatake[a,b]

[a]Division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan, [b]Department of Natural Resources, Graduate School of Global Environmental Studies, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan, and [c]Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan

Correspondence e-mail: t2masuda@kais.kyoto-u.ac.jp

Received 16 February 2011
Accepted 12 April 2011

PDB References: recombinant thaumatin, 3al7; plant thaumatin, 3ald.



Thaumatin, an intensely sweet-tasting plant protein, elicits a sweet taste at a concentration of 50 n*M*. The crystal structure of a recombinant form of thaumatin I produced in the yeast *Pichia pastoris* has been determined to a resolution of 1.1 Å. The model was refined with anisotropic *B* parameters and riding H atoms. A comparison of the diffraction data and refinement statistics for recombinant thaumatin I with those for plant thaumatin I revealed no significant differences in the diffraction data. The *R* values for recombinant thaumatin I and plant thaumatin I ($F_o > 4\sigma$) were 9.11% and 9.91%, respectively, indicating the final model to be of good quality. Notably, the electron-density maps around Asn46 and Ser63, which differ between thaumatin variants, were significantly improved. Furthermore, a number of H atoms became visible in an OMIT map and could be assigned. The high-quality structure of recombinant thaumatin with H atoms should provide details about sweetness determinants in thaumatin and provide valuable insights into the mechanism of its interaction with taste receptors.

1. Introduction

Thaumatin is a sweet-tasting protein that elicits a sweet taste sensation at 50 n*M*. Since it is nearly 100 000 times sweeter than sucrose on a molar basis (Van der Wel & Loeve, 1972), thaumatin has potential use as a low-calorie sweetener for industrial applications and could be useful in clarifying the mechanisms of the perception of sweet taste (Temussi, 2002, 2006; Masuda & Kitabatake, 2006).

Thaumatin is isolated from the arils of *Thaumatococcus daniellii* Benth, a plant native to tropical West Africa. Plant thaumatin consists of five intensely sweet forms, with two major components (thaumatin I and II) and three minor components (thaumatin a, b and c). The amino-acid sequences of the major components of thaumatin have been determined by Edman degradation and the nucleotide sequences have also been determined (Iyengar *et al.*, 1979; Lee *et al.*, 1988; Edens *et al.*, 1982; Ide, Masuda *et al.*, 2007). The results showed that four or five positions (N46K, S63R, K67R, R76Q and D113N) differ between the two major components of thaumatin.

Thaumatin is frequently used in crystallization studies and 34 structures have been deposited in the Protein Data Bank to date. Although plant-sourced thaumatin contains several variants, mainly thaumatin I and thaumatin II, no further purification tends to be performed for crystallization. For this reason, some parts of the electron-density maps, especially at residues 46 and 63, are often unclear. To clarify these ambiguous regions, the determination of an atomic resolution structure by analyses using homogeneous samples would provide a number of insights (Sauter *et al.*, 2001; Howard *et al.*, 2004; Biadene *et al.*, 2007; Wang *et al.*, 2007; Sherawat *et al.*, 2007).

Although much research has been devoted to producing thaumatin using microorganisms (Faus, 2000; Masuda & Kitabatake, 2006), the production of an active sweet form is still difficult since thaumatin contains eight disulfide bonds. While some studies have achieved acceptable yields of recombinant thaumatin, there were still a few amino acids attached at the N-terminus owing to a deficient protease

Table 1
Differences in amino-acid sequence among thaumatin variants.

Thaumatin	Amino-acid position				
	46	63	67	76	113
Amino-acid sequence					
I (Iyengar *et al.*, 1979)	Asn	Ser	Lys	Arg	Asn
A (Lee *et al.*, 1988)	Asn	Ser	Lys	Arg	Asp
B (Lee *et al.*, 1988)	Lys	Ser	Lys	Arg	Asp
I (Kaneko & Kitabatake, 2001)	—	—	—	Arg	Asn
Deduced amino-acid sequence from cDNA					
I (Ide, Masuda *et al.*, 2007), AF355098	Asn	Ser	Lys	Arg	Asp
II (Edens *et al.*, 1982), J01209	Lys	Arg	Arg	Gln	Asp
II (Verrips *et al.*, 1982), A15660	Lys	Arg	Arg	Gln	Asp
II (Masuda *et al.*, 2004)	Lys	Arg	Arg	Gln	Asp

(Weickmann *et al.*, 2004; Lombraña *et al.*, 2004; Masuda *et al.*, 2004; Ide, Kaneko *et al.*, 2007). More recently, homogeneous recombinant thaumatin was successfully obtained from the yeast *Pichia pastoris* using a pre-sequence of thaumatin as a secretion signal (Ide, Masuda *et al.*, 2007; Masuda *et al.*, 2010).

In the present study, the high-resolution structure of recombinant thaumatin I produced by *P. pastoris* was determined at 1.1 Å resolution. The similarities and differences in the refined structures of recombinant thaumatin I as well as plant thaumatin I were investigated in detail. A high-resolution structural analysis with H atoms showed an improvement of the maps and could discriminate between thaumatin variants.

2. Materials and methods

2.1. Amino-acid sequence of thaumatin I

The amino-acid sequence of thaumatin I was determined by Iyengar and coworkers by Edman degradation (Iyengar *et al.*, 1979). Later, Lee and coworkers re-examined the two major components of natural thaumatin and designated them thaumatin A and thaumatin B (Lee *et al.*, 1988). The amino-acid sequences of thaumatin A and thaumatin B differed from that of thaumatin I determined by Iyengar and coworkers by one amino acid (Asp113 instead of Asn) and two amino acids (Lys46 instead of Asn and Asp113 instead of Asn), respectively.

The nucleotide sequence of thaumatin II was determined by Edens *et al.* (1982) and the deduced amino-acid sequence showed that thaumatin II differs from thaumatin I at five positions (N46K, S63R, K67R, R76Q and N113D) and from thaumatin A at four positions (N46K, S63R, K67R and R76Q). Recently, the nucleotide sequence of thaumatin I from cloned cDNA has been determined in detail (Ide, Masuda *et al.*, 2007). The deduced amino-acid sequence from this cDNA differed from that of thaumatin I reported by Iyengar and coworkers at residue 113 (Asp113 instead of Asn) and was identical to that of thaumatin A (Table 1).

2.2. Materials

Plant thaumatin was purchased from Wako Pure Chemical Industries Ltd, Osaka, Japan. Potassium sodium L-tartrate was obtained from Nacalai Tesque Inc., Kyoto, Japan. *N*-(2-Acetamido)-iminodiacetic acid (ADA) was obtained from Dojindo, Kumamoto, Japan. All other chemicals were of guaranteed reagent grade for biochemical use.

2.3. Expression and purification of recombinant thaumatin I

The expression of recombinant thaumatin was performed in *P. pastoris* as described previously (Masuda *et al.*, 2005, 2010). A 7 l fermenter (TS-M7L, Takasugi Seisakusho Co., Tokyo, Japan) was used to secrete recombinant thaumatin I into the culture medium. After centrifugation of the culture medium, the supernatant was applied onto an SP-Sephadex C-25 cation-exchange column and the bound proteins were eluted with 5 m*M* sodium phosphate buffer pH 7.0 containing 500 m*M* NaCl. The fractions containing thaumatin were combined and precipitated with 75% ammonium sulfate and purified by HW-50F gel-filtration chromatography (Tosoh Co., Tokyo, Japan). The concentrations of thaumatin were determined spectrophotometrically using a molar extinction coefficient ε_{278} of 17 000 M cm^{-1} (van der Wel & Loeve, 1972).

Table 2
Diffraction data statistics.

Values in parentheses are for the last shell.

	Recombinant	Plant
Beamline	SPring-8 BL38B1	SPring-8 BL38B1
Detector	R-AXIS V	R-AXIS V
Crystal system	Tetragonal	Tetragonal
Space group	$P4_12_12$	$P4_12_12$
Unit-cell parameters (Å, °)	a = 57.699, b = 57.699, c = 149.776, $\alpha = \beta = \gamma$ = 90	a = 57.698, b = 57.698, c = 149.846, $\alpha = \beta = \gamma$ = 90
X-ray wavelength (Å)	0.7	0.7
Temperature (K)	100	100
Resolution limits (Å)	20.0–1.10 (1.14–1.10)	50.0–1.10 (1.12–1.10)
Total reflections	1118513	1032154
Unique reflections	102368 (9251)	103182 (5055)
$R_{\rm merge}$	0.084 (0.368)	0.070 (0.417)
Completeness (%)	98.9 (90.8)	99.7 (99.5)
Multiplicity	6.5 (4.2)	5.2 (5.2)
$\langle I\rangle/\langle\sigma(I)\rangle$	27.69 (2.21)	37.74 (4.91)

2.4. Purification of plant thaumatin I

The purification of plant thaumatin I was performed as described by Kaneko & Kitabatake (1999). Briefly, thaumatin solution was applied onto an SP-Sephadex C-25 cation-exchange column and the bound proteins were eluted with a linear gradient of 20–120 m*M* NaCl in 5 m*M* sodium phosphate buffer pH 7.0. Fractions containing thaumatin I were further purified by HW-50F gel-filtration chromatography. The purity of the proteins was checked by SDS–PAGE and native PAGE.

2.5. Crystallization and data collection

The purified recombinant thaumatin and plant thaumatin were concentrated using a Centricon-10 (Millipore, Bedford, Massachusetts, USA) and the protein concentration was measured using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Rockland, Delaware, USA). Crystallization was performed using the hanging-drop vapour-diffusion method. The hanging drops were prepared by mixing 5 µl 10–100 mg ml^{-1} protein solution and 5 µl reservoir solution. The reservoir solution consisted of 0.1 *M* ADA, 0.5–1.0 *M* potassium sodium tartrate pH 6.5–6.8 and 0, 10 or 25% glycerol. All solutions were prepared with sterile water and filtrated with a MILLEXGV membrane (Millipore). Diffraction data were obtained to 1.1 Å resolution. The crystal was placed in a cold nitrogen-gas stream and X-ray diffraction images were collected using an R-AXIS V area detector (Rigaku, Tokyo, Japan) with synchrotron radiation of wavelength 0.7 Å at the BL-38B1 station of SPring-8 (Hyogo, Japan). The data obtained were processed, merged and scaled using the *HKL*-2000 program package (Otwinowski & Minor, 1997). Data-collection and structure-solution statistics are shown in Tables 2 and 3.

2.6. Structure refinement and validation

The structures of recombinant thaumatin and plant thaumatin were determined by molecular replacement using the program *MOLREP* (Vagin & Teplyakov, 2010) in the *CCP*4 suite (Winn *et al.*, 2011) with the previously reported thaumatin structure as a reference (PDB code 1rqw; Q. Ma & G. M. Sheldrick, unpublished work). Modelling was performed using *TURBO-FRODO* on a Silicon Graphics computer and *Coot* (Emsley & Cowtan, 2004). $|F_o| - |F_c|$ and $2|F_o| - |F_c|$ maps were used to locate the correct model. Several rounds of refinement were carried out to improve the model by increasing the data resolution to 1.5 Å. Water molecules were incorporated where the difference density exhibited a value of greater than 3.0σ above the mean and the $2|F_o| - |F_c|$ map showed a density of greater than 1.0σ. Refinement was performed using *SHELXL*97 (Sheldrick, 2008) with stereochemical restraints based on those of Engh & Huber (1991). All reflections were included with no σ cutoff; 5% of the data were randomly selected and omitted during refinement for cross-validation by means of the free *R* factor (Brünger, 1992). The occupancy of the major conformation was refined first; the second or third conformation was then assigned and refined based on its $|F_o| - |F_c|$ map and finally atoms of all conformations were refined. Disordered residues were taken into consideration and possible side-chain conformations were generated. The occupancies of the disordered residues were treated as free variables and were refined using SUMP and FVAR restraints. Refinement of the model with isotropic *B* factors against data to a resolution of 1.1 Å using *SHELXL* resulted in an R_{work} of 17.55% and an R_{free} of 19.54%. Subsequent anisotropic *B*-factor refinement against data to 1.1 Å resolution lowered R_{work} and R_{free} to 11.97% and 15.01%, respectively (corresponding to a fall in R_{free} of 4.53%). Next, H atoms were modelled using HFIX and R_{work} and R_{free} fell to 10.80% and 13.54%, respectively. H atoms were only included for protein atoms, not for tartrate/glycerol/solvent atoms. To finalize the refinement, the model including H atoms was refined against all data for 30 cycles using conjugate-gradient least-squares minimization, leading to an *R* factor of 10.83% for 102 055 reflections and 9.11% for 80 465 reflections with $F_o > 4\sigma$ in the resolution range 10–1.1 Å. The quality of the structure was assessed using *PROCHECK* (Laskowski *et al.*, 1993) and *WHAT IF* (Hekkelman *et al.*, 2010). Standard uncertainties were estimated with *SHELXPRO* through a block-diagonal calculation after removal of all restraints (Sheldrick, 2008). The electron-density maps and structural images were generated using *PyMOL* (DeLano, 2002).

The coordinates and observed intensities of recombinant thaumatin I and plant thaumatin I have been deposited in the PDB (accession codes 3al7 for recombinant thaumatin and 3ald for plant thaumatin).

3. Results and discussion

3.1. Crystallization of recombinant thaumatin I

In the presence of sodium potassium tartrate, thaumatin crystallizes in a tetragonal system. We purified recombinant thaumatin I from the culture medium of the yeast *P. pastoris* and attempted its crystallization. Numerous pyramid- and bipyramid-shaped crystals were obtained in a few days. We also attempted crystallization in the presence of 25%(*v*/*v*) glycerol for a high-resolution structural analysis. The first crystal appeared in only two weeks in the presence of 25%(*v*/*v*) glycerol, which was much faster than for purified plant thaumatin I and unpurified thaumatin, crystals of which appeared in three to four weeks (Charron *et al.*, 2002). It was suggested that the rate of equilibration of the protein solution by vapour diffusion was significantly reduced in the presence of glycerol and it took at least 30 d to form pyramid-shaped crystals when unpurified plant thaumatin was used (Charron *et al.*, 2002). It appeared that the homogeneous recombinant thaumatin affected the nucleation of the

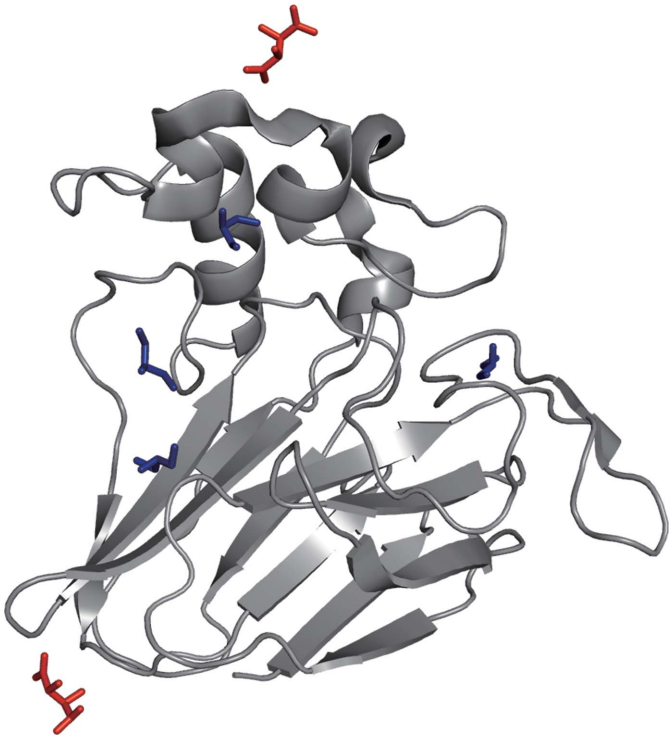


Figure 1
Overall structure of recombinant thaumatin I. Tartrate ions are shown in red and glycerol molecules are shown in blue.

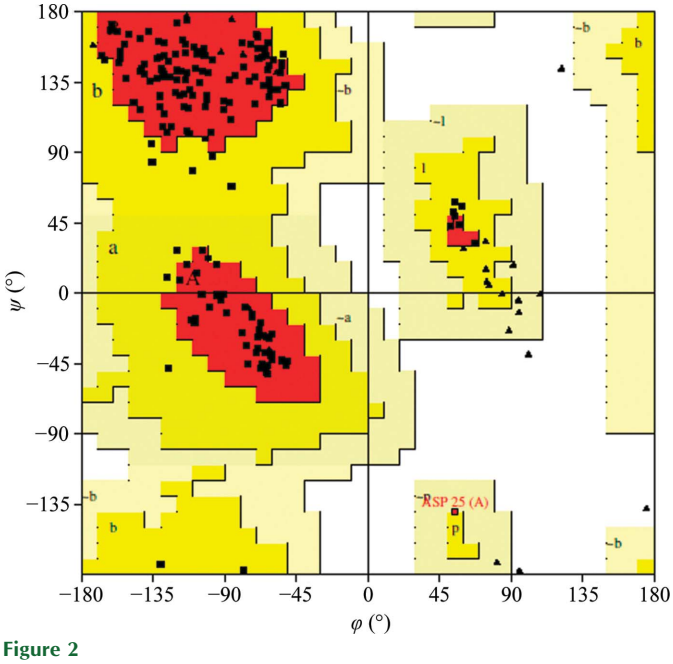


Figure 2
Ramachandran plot for recombinant thaumatin. Glycine residues are represented by triangles. 90.5% of the residues are located in most favoured regions, 8.9% in additional allowed regions and 0.6% (Asp25) in generously allowed regions.

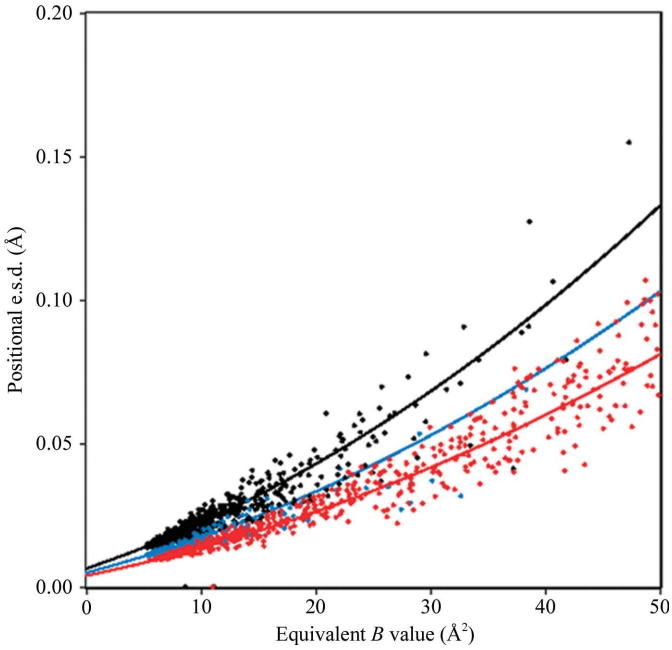


Figure 3
E.s.d. analysis of recombinant thaumatin I. Plot of positional uncertainty *versus* the thermal parameter for C atoms (black), N atoms (cyan) and O atoms (red) from the structure of recombinant thaumatin at a resolution of 1.1 Å.

Table 3
Final model statistics.

Values in parentheses are for the last shell.

	Recombinant	Plant
Resolution (Å)	10–1.10	10–1.10
Unique reflections	102055 [$F_o > 4\sigma$: 80465]	102886 [$F_o > 4\sigma$: 89235]
Isotropic R_{work}/R_{free} (%)	17.55/19.54 (15.33/17.27)	17.82/19.91 (16.38/18.53)
Anisotropic R_{work}/R_{free} (%)	11.97/15.01 (10.19/13.16)	12.19/15.37 (11.07/14.27)
Anisotropic R_{work}/R_{free} (including protein H atoms) (%)	10.80/13.54 (9.07/11.73)	10.96/13.71 (9.87/12.64)
R_{cryst} (%)	10.83 (9.11)	10.99 (9.91)
Protein atoms (non-H)	1753	1746
Protein atoms (H)	1633	1629
Tartrate atoms	20	20
Glycerol atoms	24	24
Solvent atoms	498	383
B factors (Å^2)		
Average	17.02	17.33
Protein main chain	9.82	10.53
Protein side chain	15.17	16.39
Tartrate	18.98	17.23
Glycerol	46.30	41.83
Water	37.88	37.97
Root-mean-square-deviation from ideal values		
Bonds (Å)	0.017	0.016
Angles (Å)	0.034	0.034
Ramachandran plot (%)		
Most favoured	90.5	89.3
Additional allowed	8.9	10.1
Generously allowed	0.6	0.6
Disallowed	0	0
Matthews coefficient V_M (Å^3 Da^{-1})	2.81	2.81
Solvent content (%)	56.21	56.23

crystals. The crystals obtained in this condition belonged to space group $P4_12_12$. The unit-cell parameters are listed in Table 2.

3.2. Overall structure of recombinant thaumatin I

The final model of recombinant thaumatin I consisted of 207 residues with a total of 3386 protein atoms including 1633 H atoms, two tartrate ions, four glycerol molecules and 476 water molecules (Fig. 1). A single thaumatin molecule is contained in the asymmetric unit. The *R* factor (R_{cryst}) for the final structure is 9.11% for 80 465 reflections with $F_o > 4\sigma$ and 10.83% for all data (102 055 reflections) in the resolution range 10–1.1 Å, suggesting that the final refined model is of high quality. A Ramachandran plot calculated for the final model showed that 90.5% of the residues are in most favoured regions, 8.9% are in additional allowed regions and 0.6% are in generously allowed regions. No residues are in disallowed regions (Fig. 2, Table 3). The residue located in the generously allowed region is Asp25, which is part of the thaumatin loop (Asp21-Ala22-Ala23-Leu24-Asp25-Ala26); this was not observed in other non-sweet thaumatin-like proteins (Koiwa *et al.*, 1999). The average *B* factors for C^{α}, side-chain and all protein atoms were 9.82, 15.17 and 17.02 Å^2, respectively. Positional uncertainties were plotted against thermal parameters and 73% of them were located in the region below the 0.03 Å positional e.s.d. and 30 Å^2 equivalent *B* value (Fig. 3). We used 22 204.63 Da as the molecular mass of recombinant thaumatin and plant thaumatin to calculate the Matthews coefficient V_M (Matthews, 1968). The Matthews coefficient V_M and solvent content were calculated as 2.81 Å^3 Da^{-1} and 56.21%, respectively (Table 3). OMIT maps of Phe90, Ile102 and Tyr199 clearly showed positive density (green) around these residues, suggesting the existence of H atoms (Fig. 4). Furthermore, the OMIT map of Asp113 is quite different from that of Asn104 (Fig. 5). So far, it has been quite difficult to determine whether the residue at 113 is Asp or Asn in plant thaumatin directly from the density maps. This might be because the resolution around the residue is still poor and unpurified plant thaumatin contains some variants. Since the residue in the recombi-

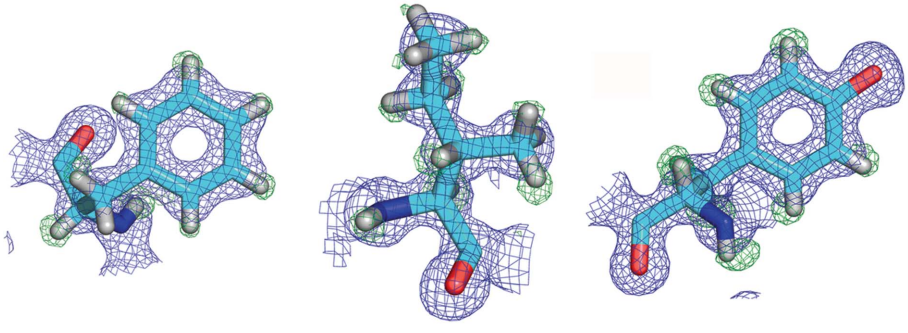


Figure 4
The model and electron density for recombinant thaumatin I around Phe90, Ile102 and Tyr199. (*a*) Phe90. (*b*) Ile102. (*c*) Tyr199. σ_A-weighted $2mF_o - DF_c$ maps contoured at 1.0σ are shown in blue and $mF_o - DF_c$ maps omitting H atoms contoured at 3.0σ are shown in green. The difference in density clearly shows numerous H atoms.

nant thaumatin is Asp, the results of the present study should help to clarify the residue at position 113 in plant thaumatin.

3.3. Amino-acid residues in multiple conformations

The quality of the 1.1 Å resolution electron-density map of recombinant thaumatin I allowed the side chains of 20 residues (Glu4, Arg8, Leu31, Glu35, Asp55, Tyr57, Asp70, Arg79, Arg82, Met112, Arg119, Arg122, Val124, Ser155, Cys159, Lys163, Glu168, Arg171, Asp179 and Arg200) to be modelled in two conformations and that of one residue (Arg76) to be modelled in three conformations. Ma and Sheldrick (PDB entry 1rqw) previously suggested that 14 residues (Glu4, Leu31, Tyr57, Ile65, Cys66, Lys67, Asp70, Met112, Cys126, Ser155, Cys159, Lys163, Asp179 and Cys193) had alternative side-chain conformations in plant thaumatin (Q. Ma & G. M. Sheldrick, unpublished work). Asherie and coworkers (PDB code 2vhk; Asherie *et al.*, 2009) also assigned 25 residues (Glu4, Arg8, Leu31, Glu35, Glu42, Asn46, Asp55, Tyr57, Ser63, Cys66, Asp70, Arg76, Arg82, Met112, Arg122, Val124, Cys126, Ser155, Cys159, Lys163, Glu168, Arg171, Asp179, Cys193 and Arg200) as having alternative side-chain conformations. Thus, previous studies have assigned some Cys residues as having alternative side-chain conformations; however, we could not identify alternative structures of Cys66, Cys126 and Cys193 in the electron-density map of recombinant thaumatin I (Fig. 6). It is known that the electron density of the disulfide bridges of thaumatin can be affected by radiation (Banumathi *et al.*, 2004; Schulze-Briese *et al.*, 2005). It has also been suggested that cleavage and/or breakage of some disulfide bonds might occur in naturally sourced plant thaumatin. Further investigations would be required in order to clarify the environments of the Cys residues.

3.4. Comparison of the structure of plant thaumatin

The differences in r.m.s. deviation on C^{α} atoms between the final refined structure of recombinant thaumatin (PDB entry 3al7) and those of purified plant thaumatin [PDB entries 2vhk (Asherie *et al.*, 2009) and 3ald] and unpurified thaumatin [PDB entries 1rqw (Q. Ma & G. M. Sheldrick, unpublished work), 1kwn (Sauter *et al.*, 2002) and 1lxz (Charron *et al.*, 2002)] were investigated. The results showed that the r.m.s.d. values against purified thaumatin (2vhk and 3ald) were 0.067 and 0.073 Å, respectively, whereas they ranged from 0.19 to 0.24 Å when compared with unpurified thaumatin (1rqw, 1kwn and 1lxz). These results indicated that recombinant thaumatin I is similar in overall structure to purified plant thaumatin I (2vhk and 3ald) but significantly different from unpurified thaumatin.

We also attempted to crystallize purified plant thaumatin I under the same conditions as used for recombinant thaumatin I. The data set for plant thaumatin had an $R_{\rm merge}$ of 7.0% and was 99.7% complete (Table 2). There were no significant differences in data collection between recombinant thaumatin and plant thaumatin.

The structure of purified thaumatin I was refined to a final R factor of 10.99% for 102 886 reflections and of 9.91% for reflections with $F_{\rm o} > 4\sigma$ (89 235 reflections) in the resolution range 10.0–1.1 Å. The average B factors for plant thaumatin were 10.53 Å^2 for C^{α} atoms, 16.39 Å^2 for side-chain atoms and 17.33 Å^2 for all atoms (Table 3). Compared with the results for recombinant thaumatin I, the final R

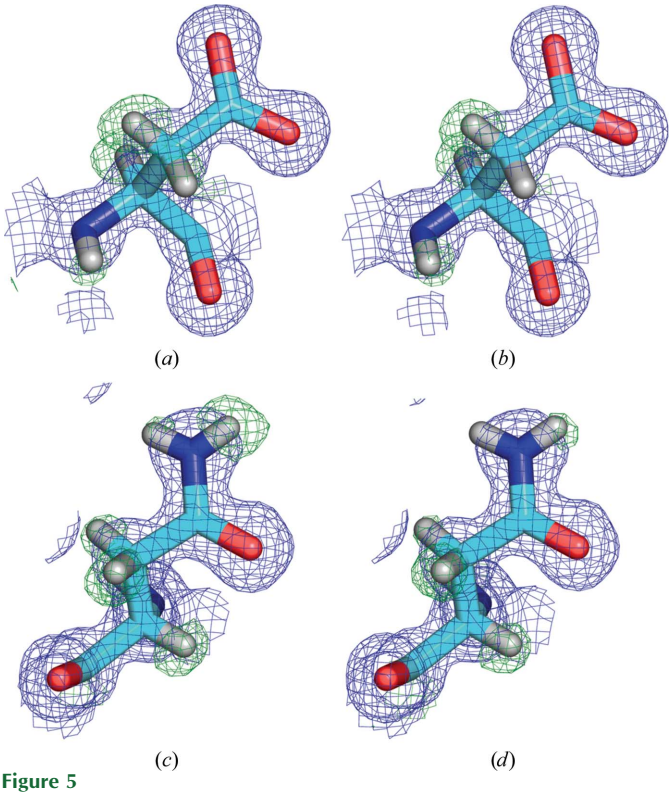


Figure 5
Comparison of the models of recombinant thaumatin I and plant thaumatin I around Asn104 and Asp113. The OMIT maps of recombinant thaumatin are shown in (*a*) for Asp113 (*B* factors: OD1, 7.74 Å^2; OD2, 9.28 Å^2) and (*c*) for Asn104 (*B* factors: ND2, 6.78 Å^2; OD1, 7.21 Å^2). In contrast, the OMIT maps of plant thaumatin I are shown in (*b*) for Asp113 (*B* factors: OD1, 8.59 Å^2; OD2, 10.19 Å^2) and (*d*) for Asn104 (*B* factors: ND2, 7.40 Å^2; OD1, 7.93 Å^2). The $\sigma_{\rm A}$-weighted $2mF_{\rm o} - DF_{\rm c}$ maps contoured at 1.0σ are drawn in blue and the $mF_{\rm o} - DF_{\rm c}$ maps omitting H atoms contoured at 3.0σ are shown in green.

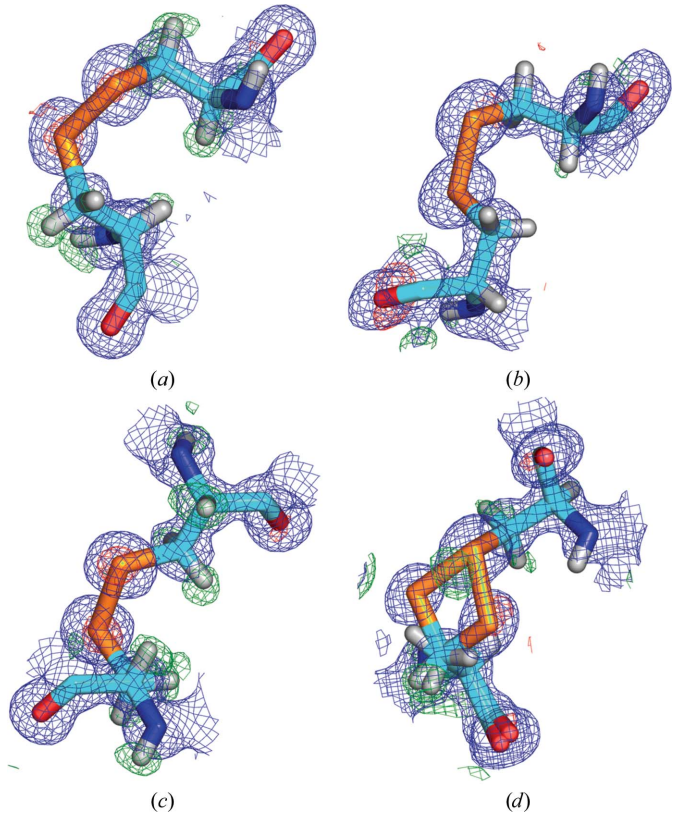


Figure 6
Models of Cys residues in recombinant thaumatin I. Four of the eight disulfide bonds are indicated: (*a*) Cys56–Cys66, (*b*) Cys121–Cys193, (*c*) Cys126–Cys177 and (*d*) Cys159–Cys164. Disulfide linkages are shown in orange. $\sigma_{\rm A}$-weighted $2mF_{\rm o} - DF_{\rm c}$ maps contoured at 1.0σ are drawn in blue and $mF_{\rm o} - DF_{\rm c}$ maps contoured at 3.0σ and −3.0σ are shown in green and red, respectively.

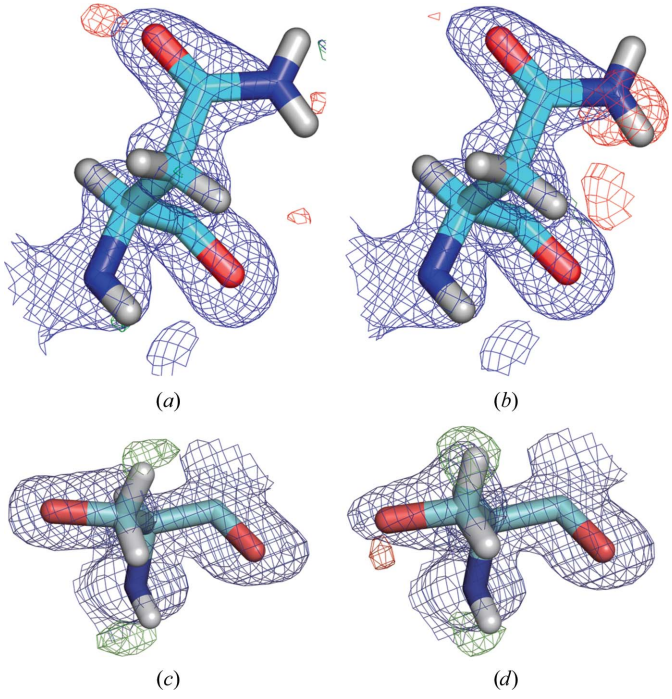


Figure 7
Comparison of the models of recombinant thaumatin I and of plant thaumatin I around Asn46 and Ser63. The electron-density maps of recombinant thaumatin I are shown for (*a*) Asn46 and (*c*) Ser63. In contrast, the electron-density maps of plant thaumatin I are indicated in (*b*) for Asn46 and (*d*) for Ser63. $2mF_o - DF_c$ maps contoured at 1.0σ are drawn in blue and $mF_o - DF_c$ maps contoured at 3.0σ and -3.0σ are shown in green and red, respectively.

factor of plant thaumatin is slightly higher than that of recombinant thaumatin. The maps of plant thaumatin showed no obvious alternative conformations of the 16 Cys residues apart from Cys159 (data not shown). In contrast, the electron-density maps around Asn46 and Ser63 differed significantly between recombinant thaumatin I and plant thaumatin I (Fig. 7). Since these two residues vary among thaumatin variants, it could be considered that plant thaumatin contains some other contaminants. It is well known that the bands for plant thaumatin smear on native PAGE on storage. It seems difficult to separate the thaumatin variants. Overall, the quality of the final model of recombinant thaumatin I is likely to be attributable to its homogeneous origins.

It had been unclear which amino-acid residue was present at position 113 of plant thaumatin I. Although the deduced sequence from thaumatin I cDNA indicated the residue to be Asp, it was difficult to determine the residue directly from the electron-density map. By comparison of the structure of recombinant thaumatin I with that of plant thaumatin I, it was clearly shown that residue 113 in plant thaumatin is not Asn but Asp (Fig. 5).

The structural requirements necessary for the sweetness of thaumatin molecules have been investigated using chemical modifications (van der Wel & Bel, 1976; Shamil & Beynon, 1990; Kaneko & Kitabatake, 2001) and site-directed mutagenesis (Kim & Weickmann, 1994; Ohta *et al.*, 2008, 2011). However, the precise structural features of the sweetness determinants remain unknown. Recently, the neutron structure of plant thaumatin has been determined to assign protonation states and solvent structure (Teixeira *et al.*, 2008, 2010). The atomic structural features of recombinant thaumatin with H atoms obtained in this study will provide important information on the perception of the sweet taste of thaumatin.

This work was supported by Grants-in-Aid for Young Scientists (B) (TM; No. 19780074) and Scientific Research (C) (TM; No. 22580105) from The Japan Society for the Promotion of Science and by the Japan Food Chemical Research Foundation. The synchrotron-radiation experiments were performed at BL38B1 and BL26B1 at SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (JASRI; proposal Nos. 2009A1096 and 2009B1379).

References

Asherie, N., Jakoncic, J., Ginsberg, C., Greenbaum, A., Stojanoff, V., Hrnjez, B. J., Blass, S. & Berger, J. (2009). *Cryst. Growth Des.* **9**, 4189–4198.
Banumathi, S., Zwart, P. H., Ramagopal, U. A., Dauter, M. & Dauter, Z. (2004). *Acta Cryst.* D**60**, 1085–1093.
Biadene, M., Hazemann, I., Cousido, A., Ginell, S., Joachimiak, A., Sheldrick, G. M., Podjarny, A. & Schneider, T. R. (2007). *Acta Cryst.* D**63**, 665–672.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Charron, C., Kadri, A., Robert, M.-C., Giegé, R. & Lorber, B. (2002). *Acta Cryst.* D**58**, 2060–2065.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Edens, L., Heslinga, L., Klok, R., Ledeboer, A. M., Maat, J., Toonen, M. Y., Visser, C. & Verrips, C. T. (1982). *Gene*, **18**, 1–12.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Faus, I. (2000). *Appl. Microbiol. Biotechnol.* **53**, 145–151.
Hekkelman, M. L., Te Beek, T. A., Pettifer, S. R., Thorne, D., Attwood, T. K. & Vriend, G. (2010). *Nucleic Acids Res.* **38**, W719–W723.
Howard, E. I., Sanishvili, R., Cachau, R. E., Mitschler, A., Chevrier, B., Barth, P., Lamour, V., Van Zandt, M., Sibley, E., Bon, C., Moras, D., Schneider, T. R., Joachimiak, A. & Podjarny, A. (2004). *Proteins*, **55**, 792–804.
Ide, N., Kaneko, R., Wada, R., Mehta, A., Tamaki, S., Tsuruta, T., Fujita, Y., Masuda, T. & Kitabatake, N. (2007). *Biotechnol. Prog.* **23**, 1023–1030.
Ide, N., Masuda, T. & Kitabatake, N. (2007). *Biochem. Biophys. Res. Commun.* **363**, 708–714.
Iyengar, R. B., Smits, P., van der Ouderaa, F., van der Wel, H., van Brouwershaven, J., Ravestein, P., Richters, G. & van Wassenaar, P. D. (1979). *Eur. J. Biochem.* **96**, 193–204.
Kaneko, R. & Kitabatake, N. (1999). *J. Agric. Food Chem.* **47**, 4950–4955.
Kaneko, R. & Kitabatake, N. (2001). *Chem. Senses*, **26**, 167–177.
Kim, S.-H. & Weickmann, J. L. (1994). *Thaumatin*, edited by M. Witty & J. D. Higginbotham, pp. 135–149. Boca Raton: CRC Press.
Koiwa, H., Kato, H., Nakatsu, T., Oda, J., Yamada, Y. & Sato, F. (1999). *J. Mol. Biol.* **286**, 1137–1145.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lee, J. H., Weickmann, J. L., Koduri, R. K., Ghosh-Dastidar, P., Saito, K., Blair, L. C., Date, T., Lai, J. S., Hollenberg, S. M. & Kendall, R. L. (1988). *Biochemistry*, **27**, 5101–5107.
Lombraña, M., Moralejo, F. J., Pinto, R. & Martín, J. F. (2004). *Appl. Environ. Microbiol.* **70**, 5145–5152.
Masuda, T., Ide, N., Ohta, K. & Kitabatake, N. (2010). *Food Sci. Technol. Res.* **16**, 585–592.
Masuda, T. & Kitabatake, N. (2006). *J. Biosci. Bioeng.* **49**, 4937–4941.
Masuda, T., Tamaki, S., Kaneko, R., Wada, R., Fujita, Y., Mehta, A. & Kitabatake, N. (2004). *Biotechnol. Bioeng.* **85**, 761–769.
Masuda, T., Ueno, Y. & Kitabatake, N. (2005). *Protein Expr. Purif.* **39**, 35–42.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Ohta, K., Masuda, T., Ide, N. & Kitabatake, N. (2008). *FEBS J.* **275**, 3644–3652.
Ohta, K., Masuda, T., Tani, F. & Kitabatake, N. (2011). *Biochem. Biophys. Res. Commun.* **406**, 435–438.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sauter, C., Lorber, B. & Giegé, R. (2002). *Proteins*, **48**, 146–150.
Sauter, C., Otálora, F., Gavira, J.-A., Vidal, O., Giegé, R. & García-Ruiz, J. M. (2001). *Acta Cryst.* D**57**, 1119–1126.
Schulze-Briese, C., Wagner, A., Tomizaki, T. & Oetiker, M. (2005). *J. Synchrotron Rad.* **12**, 261–267.
Shamil, S. & Beynon, R. J. (1990). *Chem. Senses*, **15**, 457–469.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Sherawat, M., Kaur, P., Perbandt, M., Betzel, C., Slusarchyk, W. A., Bisacchi, G. S., Chang, C., Jacobson, B. L., Einspahr, H. M. & Singh, T. P. (2007). *Acta Cryst.* D**63**, 500–507.

Teixeira, S. C. M., Blakeley, M. P., Leal, R. M. F., Gillespie, S. M., Mitchell, E. P. & Forsyth, V. T. (2010). *Acta Cryst.* D**66**, 1139–1143.
Teixeira, S. C. M., Blakeley, M. P., Leal, R. M. F., Mitchell, E. P. & Forsyth, V. T. (2008). *Acta Cryst.* F**64**, 378–381.
Temussi, P. A. (2002). *FEBS Lett.* **526**, 1–4.
Temussi, P. A. (2006). *Cell. Mol. Life Sci.* **63**, 1876–1888.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Van der Wel, H. & Bel, W. J. (1976). *Chem. Senses*, **2**, 211–218.
Van der Wel, H. & Loeve, K. (1972). *Eur. J. Biochem.* **31**, 221–225.
Verrips, C. T., Ledeboer, A. M., Edens, L., Klok, R. & Maat, J. (1982). Patent EP 0054330-A2.
Wang, J., Dauter, M., Alkire, R., Joachimiak, A. & Dauter, Z. (2007). *Acta Cryst.* D**63**, 1254–1268.
Weickmann, J. L., Blair, L. C. & Wilcox, G. L. (2004). *Thaumatin*, edited by M. Witty & J. D. Higginbotham, pp. 151–169. Boca Raton: CRC Press.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.