<table>
<thead>
<tr>
<th>Title</th>
<th>Crystal structure of the sweet-tasting protein thaumatin II at 1.27Å.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Masuda, Tetsuya; Ohta, Keisuke; Tani, Fumito; Mikami, Bunzo; Kitabatake, Naofumi</td>
</tr>
<tr>
<td>Citation</td>
<td>Biochemical and biophysical research communications (2011), 410(3): 457-460</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2011-07-08</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/175268">http://hdl.handle.net/2433/175268</a></td>
</tr>
<tr>
<td>Rights</td>
<td>© 2011 Elsevier Inc.; This is not the published version. Please cite only the published version. この論文は出版社版ではありません。引用の際には出版社版をご確認ください。</td>
</tr>
<tr>
<td>Type</td>
<td>Journal Article</td>
</tr>
<tr>
<td>Textversion</td>
<td>author</td>
</tr>
</tbody>
</table>
Crystal structure of the sweet-tasting protein thaumatin II at 1.27 Å

Running title: Structure of thaumatin II at 1.27 Å

Authors: Tetsuya Masuda¹,², Keisuke Ohta¹,², Fumito Tani¹,², Bunzo Mikami³, and Naofumi Kitabatake¹,²*

Affiliation: ¹ Division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University, Uji, Kyoto 611-0011, Japan
² Department Natural Resources, Graduate School of Global Environmental Studies, Kyoto University, Uji, Kyoto 611-0011, Japan
³ Division of Applied Life Sciences, Graduate School of Agriculture, Kyoto University, Uji, Kyoto 611-0011, Japan

Address correspondence to: Tetsuya Masuda,
Division of Food Science and Biotechnology, Graduate School of Agriculture, Kyoto University, Uji, Kyoto 611-0011, Japan
E-mail: t2masuda@kais.kyoto-u.ac.jp
Telephone: +81-774-38-3739; Fax: +81-774-38-3740

Keywords: thaumatin, sweet-tasting protein, amino-acid variations, positive charge
Abstract

Thaumatin, an intensely sweet-tasting protein, elicits a sweet taste sensation at 50 nM. Here the X-ray crystallographic structure of one of its variants, thaumatin II, was determined at a resolution of 1.27 Å. Overall structure of thaumatin II is similar to thaumatin I, but a slight shift of the Cα atom of G96 in thaumatin II was observed. Furthermore, the side chain of residue 67 in thaumatin II is highly disordered. Since residue 67 is one of two residues critical to the sweetness of thaumatin, the present results suggested that the critical positive charges at positions 67 and 82 are disordered and the flexibility and fluctuation of these side chains would be suitable for interaction of thaumatin molecules with sweet receptors.
1. Introduction

Thaumatin is a sweet-tasting protein isolated from the arils of

*Thaumatococcus daniellii* Benth, a plant native to tropical West Africa [1].

Thaumatin consists of at least five sweet forms; with two major components (thaumatin I and thaumatin II) and three minor components (thaumatin a, b, and c). All of these forms elicit a sweet taste at approximately 50 nM, and are 100,000-fold sweeter than sucrose on a molar basis. Unveiling the molecular mechanisms of the sweet taste of thaumatin is important not only for our basic knowledge of biological systems, but also for its possible medical implications toward lifestyle-related diseases such as hypertension, hyperlipidemia, diabetes, and obesity. Although thaumatin is a low-calorie sugar substitutes of natural origin and has the potential to replace sugars and artificial sweeteners, common features among sweet-tasting proteins have not yet been identified. Site-directed mutagenesis as well as chemical modification has shown that basicity on the cleft-containing side of thaumatin is required for sweetness and two positively charged residues, Lys67 and Arg82, play a significant role in the elicitation of sweetness [2, 3]. However, the
structural basis of the sweetness of thaumatin is not fully understood. In
spite of the feasibility of producing crystals of thaumatin, studies have
focused on the structure of thaumatin I, the most abundant component of
plant thaumatin. While various crystal forms of thaumatin I (monoclinic,
orthorhombic, tetragonal, and hexagonal) have been examined extensively
[4-7], little attention has been paid to the structure of thaumatin II. Since
thaumatin has great potential as a low-calorie sugar substitute, it is worth
obtaining structural information on thaumatin II as well as thaumatin I.
Variations at position 67, critical for sweetness (Lys67 for thaumatin I and
Arg67 for thaumatin II), are particularly interesting. Herein the structure of
thaumatin II was determined at a resolution of 1.27 Å.

2. Materials and methods

2.1. Purification of thaumatin II

Powdered thaumatin (Wako Pure Chemical Industries Ltd., Osaka,
Japan) was dissolved in 5 mM sodium phosphate buffer, pH 7.0, the
supernatant obtained by centrifugation (8,000 ×g) was applied to an
SP-Sephadex C-25 cation exchange column, previously equilibrated with 5
mM sodium phosphate buffer, pH 7.0, and the bound proteins were eluted with a linear gradient of 20 to 120 mM NaCl in 5 mM sodium phosphate buffer, pH 7.0. Fractions containing thaumatin II, detected by SDS-PAGE and native PAGE, were combined and precipitated with 75% ammonium sulfate. The precipitate was collected by centrifugation at 8,000×g, dissolved in 20 mM Hepes buffer, pH 7.0, containing 150 mM NaCl, and further purified by gel-filtration chromatography (HW50F, Tosoh Co., Tokyo, Japan). The purified thaumatin II was concentrated by a Centricon 10 (Millipore, Bedford, MA), and the protein concentration was measured with a NanoDrop® ND-1000 spectrophotometer (NanoDrop Technologies, Inc., Rockland, DE).

2.2. Crystallization and X-ray diffraction

Crystallization was performed by using the hanging-drop vapor-diffusion method. The hanging drops were prepared with 5 μL of 30-50 mg/mL protein solutions and 5 μL of a reservoir solution consisting of 0.1M N-(2-Acetanido) iminodiacetic acid, 0.5-1.0 M potassium sodium tartrate, pH 6.5, and 25% glycerol. Data were collected at 100K using an ADSC quantum CCD detector with synchrotron radiation at a wavelength
of 1.0 Å at the BL-44XU station of SPring-8 (Hyogo, Japan). Diffraction data obtained at up to 1.27 Å (Table 1) were processed, merged, and scaled from 50 to 1.27 Å using the HKL2000 program package [8].

2.3. Structure determination and refinement

The structure of thaumatin II was determined by molecular replacement using the program Shelxpro [9] with the previously reported thaumatin structure as a reference (PDB entry 3al7) [7]. The rigid body refinement and subsequent refinement were performed using the SHELXL program with the data for up to 1.5 Å. |Fo|−|Fc| and 2|Fo|−|Fc| maps were used to obtain the correct model. Water molecules were incorporated where the difference in density exhibited values of more than 3.0σ above the mean and the 2|Fo|−|Fc| map showed a density of more than 1.0σ. Model-building was performed using the Coot program [10]. The residues Asn46, Ser63, Lys67, and Arg76 were substituted with Lys, Arg, Arg, and Gln, respectively. The occupancy of the major conformation was refined first, then the second conformation was assigned and refined based on its |Fo|−|Fc| map, and finally atoms of all conformations were refined. The occupancies of the disorders were treated as free variables and refined
using the FVAR restraints. Refinement of the model with isotropic $B$-factor
against the data up to a resolution of 1.27 Å by SHELXL resulted in an
$R_{\text{work}}$ of 16.76% and an $R_{\text{free}}$ of 19.29%. Subsequent anisotropic $B$-factor
refinement against the data up to 1.27 Å lowered the $R_{\text{work}}$ and $R_{\text{free}}$ to
11.09% and 14.62%, respectively. To finalize the refinement, the model
was refined against all data for 40 cycles using conjugate-gradient least
squares minimization, leading to an $R1$ of 11.13% for 67,833 reflections
and 10.55% for 64,442 reflections with $Fo > 4\sigma(Fo)$ in the resolution range
10-1.27 Å (Table 1). The quality of the structure was assessed by using
PROCHECK [11] and WHAT IF [12]. The electron density maps and
structural images were generated using PyMOL [13]. The coordinates and
observed intensities of plant thaumatin II have been deposited in the PDBj
(accession code 3aok).

3. Results and discussion

3.1. Properties of thaumatin II

Plant thaumatin comprises a mixture of at least five variants
designated thaumatin I, thaumatin II, thaumatin a, thaumatin b, and
thaumatin c [1]. Ion exchange column chromatography and native-PAGE revealed about 60% of the plant sample to be thaumatin I and 20% is thaumatin II. The sequencing of thaumatin I and thaumatin II showed that four positions (N46K, S63R, K67R, and R76Q) varied [14, 15] (Fig. 1A).

In the presence of sodium potassium tartrate ions, thaumatin easily crystallizes as a tetragonal system. Although only four residues differed between thaumatin I and thaumatin II, the crystals of thaumatin II measured approximately 0.1 × 0.1 × 0.3 mm, much smaller than those of thaumatin I (0.3 × 0.3 ×1.0 mm) [16]. Since these four residues are located at the surface (Fig. 1A), the net charge might have affected the nucleation of the crystals.

3.2. Overall structure of thaumatin II

The final model of thaumatin II consists of 207 residues with a total of 2,215 protein atoms, 2 tartrate ions, 4 glycerol molecules, and 414 water molecules. A single thaumatin molecule is contained in an asymmetric unit. The R factor for the final structure is 10.55% for 64,442 reflections with Fo > 4σ (Fo) and 11.13% for all data (67,833 reflections). A Ramachandran plot calculated for the final model showed that 89.3% of the amino acid
residues are in most favored regions, 10.1% are in additional allowed
regions, and 0.6% are in generously allowed regions. No residues are in
disallowed regions (Table 1). The only residue located in a generously
allowed region was Asp25, part of the thaumatin loop
(D21-A22-A23-L24-D25-A26), which was not observed in other
non-sweet thaumatin-like proteins. The average $B$ factor for $\alpha$-carbon, side
chain, and all protein atoms is 11.466 Å$^2$, 15.595 Å$^2$, and 17.957 Å$^2$
respectively. The Matthews coefficient Vm [17], and the solvent content are
calculated to be 2.81 Å$^3$/Da and 56.2%, respectively (Table 1). The four
amino acid residues (Lys46, Arg63, Arg67, and Gln76) which differed from
those of thaumatin I are all located on the molecular surface (Fig. 1A). The
electron density maps around these residues are shown in Fig. 2. Most of
the side chains are well fitted compared to those of thaumatin I. Of note,
however, Arg67, which is critical for sweetness, is disordered and has two
alternative side chains (Fig. 1B and Fig. 2C). Although it has been
suggested that the basicity of the side chain of residue 67 is responsible for
elicitation of the sweet taste, the results of the present study further
confirmed that the distribution of the positively charged patches formed by
guanidino-groups is not rigid but flexible (Fig. 1B). Since the threshold
value of sweetness for thaumatin II is similar to that for thaumatin I [18],
the flexibility and fluctuation of the positively charged side chain at residue 67 would be suitable for interaction with the sweet receptors.

3.3. *Comparison of thaumatin II with thaumatin I*

Superposing the coordinates of the Cα atoms of thaumatin II onto the recombinant thaumatin I (PDB entry 3al7) more clearly illustrates the structural differences among the two variants (Fig. 3A). Amino acid residues with Cα carbons having root-mean square deviations (r.m.s.d.) that were above 0.2 Å were Y57, G96, R119, G141, G142, G143, T161, G162, and K163 (Fig. 3B). Among them, Y57, R119 and K163 were modeled in two alternative conformations and highly disordered. Furthermore, G162, K163 and R119 are considered to be influenced by solvents such as glycerol [5]. Although the three consecutive glycine residues from positions 142 to 144 are highly disordered in any thaumatin structure, the r.m.s.d. value of the G96 of thaumatin II was relatively high and this difference might be specific to thaumatin I and II. As to G96, the r.m.s.d. values were also investigated against the structures of thaumatin I (PDB entries 1rqw, 2vhk, 1lxz, and 3ald) which had been determined before at high resolutions (ranging from 0.94 to 1.25 Å). The results showed...
that r.m.s.d. values against thaumatin II were 0.28 Å, 0.35 Å, 0.39 Å, and 0.26 Å, respectively, suggesting that the difference of Cα atom in G96 might be specific to thaumatin II. Since the side chain of K46 in thaumatin II, N46 in thaumatin I, is adjacent to Q94 and Y95 (4.0 to 6.5Å away from K46), the difference in the properties of the side-chain (N and K) might influence the crystal packing and result in a slight shift of the Cα atom of G96.

The structural requirements for the sweetness of thaumatin molecules have been investigated extensively [2, 3, 19-21]. Although the three-dimensional structure of sweet receptors has not yet been elucidated, models of the binding between sweet-tasting proteins and sweet receptors have been reported [22-30]. Understanding the structural determinants of thaumatin’s sweetness will provide valuable insights into the mechanisms by which sweet taste is perceived.

Acknowledgments

This work was supported by a Grant-in-aid for Young Scientists (B) (T.M., no.19780074) and Scientific Research (C) (T.M., 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 SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (proposal number 2009A1096, 2009B1379, 2010A6538, 2010B1064).
Legends to Figures

**Figure 1.** (A) Overall structure (stereo diagram) of thaumatin II. The four residues (Lys46, Arg63, Arg67 and Gln76) which differ from thaumatin I are shown in a stick model (blue). The residue critical for sweetness (Arg82) is also shown in a stick model (green). (B and C) Comparison of thaumatin II with thaumatin I at two critical residues (67 and 82). Side chains are represented by a stick model with grey (thaumatin I) and red (thaumatin II). (B) residue 67 and (C) residue 82.

**Figure 2.** The model and electron density maps around Lys46, Arg63, Arg67 and Gln76 in thaumatin II. (A) Lys46. (B) Arg63. (C) Arg67. (D) Gln76. The σA-weighted 2mFo-DFc maps contoured at 1.0σ are shown in blue and mFo-DFc maps contoured at 3.0σ and -3.0σ are shown in green and red, respectively.

**Figure 3.** The superposition of the coordinates of the Cα atoms of thaumatin II onto the recombinant thaumatin I (PDB entry 3al7). (A) Histograms of r.m.s.d. values of Cα atoms with the residue number. (B)
The residues with r.m.s.d. values greater than 0.2 Å (Y57, G96, R119, G141, G142, G143, T161, G162, and K163) are shown in red in a ribbon model of thaumatin II.

Legends to tables

Table 1. Data collection and refinement statistics

Data on the highest shells are given in parentheses.
References


correlation to biochemical and mutational studies, in: M. Witty, J.D.
Higginbotham (Eds.), Thaumatin, CRC Press, Boca Raton Florida,
1994, pp135-149.

[22] P.A. Temussi, Why are sweet proteins sweet? Interaction of brazzein,
monellin and thaumatin with the T1R2-T1R3 receptor, FEBS Lett. 526

[23] P.A. Temussi, Natural sweet macromolecules: how sweet protein

Iwata, T. Misaka, H. Sorimachi, S. Arai, K. Abe, Crystal structure of
neoculin: insight into its sweetness and taste-modifying activity, J.

heterodimeric sweet taste receptor has multiple potential ligand

Shimizu-Ibuka, T. Misaka, K. Abe, Taste-modifying sweet protein,
neoculin, is received at human T1R3 amino terminal domain,

structure of Mabinlin II: a novel structural type of sweet proteins and
50-62.

and evaluation of new analogs of the sweet protein brazzein, Chem.

[29] F.M. Assadi-Porter, E.L. Maillet, J.T. Radek, J. Quijada, J.L. Markley,
M. Max, Key amino acid residues involved in multi-point binding
interactions between brazzein, a sweet protein, and the T1R2-T1R3

of human T1R3 is necessary for the interaction between human
T1R2-T1R3 sweet receptors and a sweet-tasting protein, thaumatin,
### Table 1. Diffraction data and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>Thaumatin II</th>
</tr>
</thead>
</table>

#### Data collection

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Beamline</td>
<td>SPRing-8 BL44XU</td>
</tr>
<tr>
<td>Detector</td>
<td>ADSC Quantum 210</td>
</tr>
<tr>
<td>Crystal system</td>
<td>Tetragonal</td>
</tr>
<tr>
<td>Space group</td>
<td>$P4_2_2_2$</td>
</tr>
<tr>
<td>Cell dimensions (Å, °)</td>
<td>$a = 57.769, b = 57.769, c = 150.126, \alpha = \beta = \gamma = 90$</td>
</tr>
<tr>
<td>X-ray wavelength (Å)</td>
<td>1.0</td>
</tr>
<tr>
<td>Temperature (K)</td>
<td>100</td>
</tr>
<tr>
<td>Resolution limit (Å)</td>
<td>50.0-1.27 (1.32-1.27)</td>
</tr>
<tr>
<td>Total reflections</td>
<td>942,463</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>68,167</td>
</tr>
<tr>
<td>$R_{\text{merge}}$</td>
<td>0.072 (0.328)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100 (100)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>9.1 (8.8)</td>
</tr>
<tr>
<td>$&lt;I&gt;/&lt;\sigma(I)&gt;$</td>
<td>48.1 (8.3)</td>
</tr>
</tbody>
</table>

#### Refinement

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution (Å)</td>
<td>10-1.27</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>67,833 (Fo&gt;4σ: 64,442)</td>
</tr>
<tr>
<td>Isotropic: $R_{\text{work}}/R_{\text{free}}$ (%)</td>
<td>16.76/19.29 (15.99/18.43)</td>
</tr>
<tr>
<td>Anisotropic: $R_{\text{work}}/R_{\text{free}}$ (%)</td>
<td>11.09/14.62 (10.51/13.85)</td>
</tr>
<tr>
<td>$R_{\text{cryst}}$ (%)</td>
<td>11.13 (10.55)</td>
</tr>
<tr>
<td>Average B factor</td>
<td>17.96</td>
</tr>
<tr>
<td>Protein main chain</td>
<td>11.47</td>
</tr>
<tr>
<td>Protein side chain</td>
<td>15.60</td>
</tr>
<tr>
<td>Tartrate</td>
<td>17.57</td>
</tr>
<tr>
<td>Glycerol</td>
<td>32.30</td>
</tr>
<tr>
<td>Water</td>
<td>35.74</td>
</tr>
</tbody>
</table>

#### R.m.s.deviation from ideal values

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>r.m.s.d bond (Å)</td>
<td>0.015</td>
</tr>
<tr>
<td>r.m.s.d angle (Å)</td>
<td>0.033</td>
</tr>
</tbody>
</table>

#### Ramachandran plot (%)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Most favored</td>
<td>89.3</td>
</tr>
<tr>
<td>Additional allowed</td>
<td>10.1</td>
</tr>
<tr>
<td>Generously allowed</td>
<td>0.6</td>
</tr>
<tr>
<td>Disallowed</td>
<td>0</td>
</tr>
<tr>
<td>Matthews coefficient Vm (Å³/Da)</td>
<td>2.81</td>
</tr>
<tr>
<td>Solvent content (%)</td>
<td>56.2</td>
</tr>
</tbody>
</table>