Solution Structure, Backbone Dynamics and Stability of a Double Mutant Single-Chain Monellin: Structural Origin of Sweetness‡

Yoon-Hui Sung§, Joon Shin§, Ho-Jin Chang§, Joong Myung Cho# and Weontae Lee§*

§Department of Biochemistry and Protein Network Research Center, College of Science, Yonsei University, Seoul, 120-740, Korea and #Biotech Research Institute, LG Chem., Research Park, P.O. Box 61, Yu-Sung, Science Town, Taejon 305-380, Korea

Running title: Structure and Dynamics of a Mutant Single Chain Monellin

*To whom correspondence should be addressed: Weontae Lee, Department of Biochemistry, College of Science, Yonsei University, Seodaemoon-Gu, Shinchon-Dong, Seoul, 120-740, Korea, (Tel) 82-2-2123-2706, (Fax) 82-2-362-9897, e-mail: wlee@biochem.yonsei.ac.kr

#Current Address: CrystalGenomics, Inc., Daeduck Biocommunity, Junmin-Dong, Yusung-Gu, Taejon, Korea

‡The atomic coordinates for 20 simulated annealing structures and the energy minimized average structure have been deposited (accession code: 1FUW) with the Protein Data Bank, Brookhaven National Laboratories, P.O. Box 5000, Upton, Long Island, New York 11973-5000
SUMMARY

Single-chain monellin (SCM), which is an engineered 94-residue polypeptide, has been characterized as being as sweet as native two-chain monellin. Data from gel-filtration HPLC and NMR has proven that SCM exists as a monomer in aqueous solution. In order to determine the structural origin of the taste of sweetness, we engineered several mutant SCM proteins by mutating Glu2, Asp7 and Arg39 residues, which are responsible for sweetness. In this study, we present the solution structure, backbone dynamics and stability of mutant SCM proteins using circular dichroism, fluorescence and NMR spectroscopy. Based on the NMR data, a stable $\alpha$-helix and five stranded antiparallel $\beta$-sheet were identified for double mutant SCM. Strands $\beta_1$ and $\beta_2$ are connected by a small bulge and the disruption of the first $\beta$-strand were observed with SCM$^{DR}$ comprising residues of Ile38-Cys41. The dynamical and folding characteristics from circular dichroism, fluorescence and backbone dynamics studies revealed that both wild type and mutant proteins showed distinct dynamical as well as stability differences, suggesting the important role of mutated residues in the sweet taste of SCM. Our results will provide an insight into the structural origin of sweet taste as well as the mutational effect in the stability of the engineered sweet protein SCM.
INTRODUCTION

The native sweet protein, monellin, which was originally isolated from the berries of the West African plant *Dioscoreophyllum cumminsii* (1, 2), consists of two separate polypeptide chains: an A chain of 45-residues, and a B chain of 50-residues. Native two-chain monellin is approximately 70,000 times sweeter than sucrose and about 300 times sweeter than the dipeptide sweetener aspartame (3, 4). Other sweet taste proteins, such as thaumatin, pentadin and mabinlin are also known (5-9). Among these sweet proteins, a curculin protein has demonstrated a sweet taste and shown taste-modifying activity (10). The crystal structure of native two-chain monellin has been determined as showing a $\beta$-sheet comprised of five antiparallel strands and a single 17-residue long $\alpha$-helix. The two chains were packed closely by hydrogen bonds and hydrophobic interactions (11). In addition, the crystal structure showed that the amino terminus of the A chain was connected to the carboxyl terminus of the B chain through intermolecular hydrogen bond networks.

Recent biochemical studies have reported that the A chain of the alcohol-denatured state of native monellin performed a structural reorganization from $\beta$-sheet to $\alpha$-helix conversion in 50% ethanol and 50% trifluoroethanol (TFE) environments (12). In addition, the conformational study for both native and mutated non-sweet analog two-chain monellin have been studied by two-dimensional nuclear magnetic resonance spectroscopy; these studies have shown that the three dimensional structures of native monellin and two thiol proteinase inhibitors, cystatin and stefin B, are very similar (13). This structural homology data indicated that monellin might play some other biological
role besides its sweetness.

An engineered 94-residue single-chain monellin (SCM), which was recently constructed by fusing the two chains to retain the topology of monellin (14), has proven to be as sweet as native monellin. Interestingly, SCM was more stable than the native two-chain monellin for both heat and acidic environments (14). Very recently, two-dimensional $^1$H NMR (15) and heteronuclear three-dimensional NMR studies (16) for recombinant SCM have been performed as a monomer conformation in solution state. The solution structure of SCM revealed that the long $\alpha$-helix was folded into the concave side of a six-stranded antiparallel $\beta$-sheet and the common residues for all sweet peptides were mostly solvent exposed. Interestingly, most of the residues involved in the sweet taste of monellin are found on the same surface of the molecule. In addition, the solution structure suggested that the relative orientation of the single $\alpha$-helix might be responsible for the global topology of the molecule. The flexibilities of the side chains were also important for both sweet taste and receptor binding (16). It is still not clear, however, whether SCM requires conformational preference for sweet taste, even though there is no doubt that the solution structure is responsible for binding as well as for sweet taste. Presented here are the solution structure, dynamical properties and protein stability of single-chain monellins by heteronuclear NMR, circular dichroism and fluorescence spectroscopy.

**EXPERIMENTAL PROCEDURES**

*Expression and Purification of Wild Type and Mutant SCM*
The recombinant SCM proteins were expressed in *Escherichia coli* strain BL21 (DE3) containing the plasmid pET21. Transformed cells were propagated in *E. coli* nitrogen base containing 5% glucose and 0.5% ammonium sulfate at 30 °C for 2 hours and grown in M9 media containing 2% glucose and 0.1% ammonium sulfate at 30 °C for 48 hours. $^{15}$N-labeled ammonium sulfate was used as the sole source of the nitrogen for uniformly $^{15}$N-labeled SCM$^{\text{DR}}$ and wild type SCM. The cells were harvested by centrifugation at 3500 rpm for 25 min. Cells were stored at -80 °C and used for purification procedures. Cell pastes were disrupted by a bead beater in 25 mM sodium phosphate, 5 mM EDTA, 150 mM NaCl, and 1 mM PMSF at pH 7.0. The cell lysates were collected by centrifugation at 12000 rpm for 215 min. After pH adjustment and centrifugation, the supernatants were diluted with 10 mM of sodium phosphate and loaded onto a CM-Sepharose column. The bound SCM was eluted with a salt gradient. The collected protein solution was dialyzed and dried with a freeze-dryer for spectroscopic measurements. The protein concentration was determined using the Bradford method.

*Fluorescence Spectroscopy*

Fluorescence spectra were measured in 50 mM potassium phosphate buffer, at pH 7.0 and 25 °C on F-4500 fluorescence spectrophotometer. Fluorescence emission spectra were recorded from 270 to 450 nm at each GdnHCl concentration using two different excitation wavelengths, 280 and 295 nm. The protein concentration in the cuvette was 30 μM and a path length of 1 cm was used. GdnHCl-unfolding experiments were carried out after the protein was incubated in solutions containing different
concentrations of the denaturant for 24 hours, at 25 °C. The refolding reaction of SCM was carried out under various conditions by diluting the denaturant concentration.

Data from the equilibrium denaturation were converted to plots of \( f_U \) versus denaturant concentration using the equation [1]:

\[
f_U = \frac{Y_0 - (Y_F + m_F[D])}{(Y_U + m_U[D]) - (Y_F + m_F[D])}
\]

where \( Y_0 \) is the observed signal at a particular GdnHCl concentration. \( Y_F \) and \( Y_U \) represent the intercepts, and \( m_F \) and \( m_U \) are the slopes of the native protein and the unfolded baselines. They were obtained by extrapolation of linear least-squares fits of the baselines.

To determine whether the two-state unfolding model was appropriate for analyzing the GdnHCl-induced denaturation data, \( f_U \) values were fitted to equation [2]:

\[
f_U = \frac{\exp[-(\Delta G_U + m_U[D]) \cdot RT]}{1 + \exp[-(\Delta G_U + m_U[D]) \cdot RT]}
\]

In eq [2], \( f_U \) is related to \( \Delta G_U \) by a transformation of the Gibbs-Helmholtz equation in which the equilibrium constant for unfolding in the folding transition zone, \( K_U \), is given by \( K_U = f_U/(1-f_U) \), for a two-state transition. It is also implicit in eq [2] that the free energy of unfolding is dependent linearly on denaturant concentration (17).

**CD Spectroscopy**

CD spectra were measured in 50mM of potassium phosphate buffer, at pH 7.0 and 25 °C on a Jasco 720 spectropolarimeter. Far-UV CD spectra were monitored from 190 to 250 nm using a protein concentration of 30 \( \mu \)M with a path length of 0.1 mm, 20
mdeg sensitivity, response time of 1 s, and scan speed of 50 nm/min. The spectra were recorded as a 6 scan average value. The molar ellipticity was determined as:

$$\theta_l = \frac{\theta_\lambda \cdot M_{ar}}{c \cdot l} \tag{3}$$

where $c$ is the protein concentration (in g/ml), $l$ the light path length in the cell (in mm), $\theta_\lambda$ the measured ellipticity (in degrees) at wavelength $\lambda$, and $M_{ar}$ the mean molecular mass of amino acid of the protein determined from its amino acid sequence.

The temperature scanning CD measurements were carried out with a Jasco 715 spectropolarimeter from 25 ºC to 100 ºC using a cuvette of 0.2 mm for 222 nm wavelength. The heating rates were 30 ºC/hour with a step interval of 0.1 ºC. Full CD spectra were collected at 20, 40, 60, 70 and 80 ºC for far-UV region. Noise reduction was applied to thermal scan profiles of CD spectra and for determination of the transition midpoint temperatures ($T_m$), transitions analyzed on the basis of the two-state approximation were fitted to the following relation derived from the van’t Hoff equation,

$$f_U = \frac{\exp[-\Delta H_m^{\text{van}} / R(1/T - 1/T_m)]}{1 + \exp[-\Delta H_m^{\text{van}} / R(1/T - 1/T_m)]} \tag{4}$$

Reversibility was examined by comparing the transition curves of a sample that was briefly heated to a temperature where the protein was completely unfolded.

**NMR Spectroscopy**

All NMR spectra were acquired on a Bruker DRX-500 spectrometer in quadrature detection mode, equipped with a triple-resonance probe with an actively shielded pulsed field gradient (PFG) coil. All two-dimensional experiments were performed at 298 K. Pulsed-field gradient techniques were used for all H$_2$O experiments,
resulting in good suppression of the solvent signal. $^{15}$N-$^1$H HSQC (18, 19) spectra were recorded with a uniformly $^{15}$N-labelled sample with 2048 complex data points in $t_2$ and 256 $t_1$ increments. Two-dimensional nuclear Overhauser effect spectroscopy (NOESY) (20, 21) spectra were recorded with mixing times with 100 ms and 150 ms in H$_2$O and D$_2$O solution. 2D total correlation spectroscopy (TOCSY) (22) spectrum was acquired in H$_2$O solution with a mixing time of 69.668 ms using MLEV17 spin lock pulses. $^{15}$N-edited NOESY-HSQC (23) with a mixing time of 100 ms and $^{15}$N-edited TOCSY-HSQC (24) with a mixing time of 69.668 ms spectra for uniformly $^{15}$N-labelled SCM$^{DR}$ were recorded. $^{15}$N-edited HNHA (25) and double quantum-filtered (DQF) COSY (26) spectra were collected to get vicinal coupling constant values. The $^1$H spectra were referenced to the water resonance at 4.76 ppm. Longitudinal ($R_1$) and transversal ($R_2$) relaxation data for the backbone $^{15}$N nuclei of wild type SCM were recorded as 2048 x 64 data sets with 64 scans per point using 1 s of a recovery delay. Seven different values for the relaxation time were used; $T_1$ = 5, 65, 145, 246, 366, 527 and 757 ms and $T_2$ = 8.3, 25.1, 41.8, 58.6, 75.3, 108.8 and 142.3 ms. To permit estimation of noise levels, duplicate spectra were recorded for $T = 246$ ms (T$_1$ spectra) and $T = 56.8$ ms (T$_2$ spectra). In order to eliminate the effects of cross correlation between $^{15}$N-$^1$H dipolar and $^{15}$N CSA relaxation mechanisms, $^1$H 180° pulses were inserted during the relaxation time according to the published methods (27, 28). $^{15}$N-$^1$H steady-state heteronuclear NOE (XNOE) (29, 30) data was also obtained using a relaxation delay of 5 s.

NMR Data Processing and Analysis

The NMR data were processed using the nmrPipe/NMRDraw software packages.
(Biosym/Molecular Simulations, Inc.) and analyzed by the Sparky 3.60 software. The experimental data were extended by linear prediction and zero-filled to give 2048 x 512 data matrices and processed using gaussian multiplication and a shifted ($\pi/3$) sine bell function prior to Fourier transformation. The peak intensities in the two-dimensional spectra were measured by peak heights using the Sparky program. The XNOE value for a given residue was calculated as the intensity ratio ($I/I_0$) of the $^{15}$N-$^1$H correlation peak in the presence ($I$) and absence ($I_0$) of 3 s proton saturation. The standard deviations of these values were measured background noise levels. Relaxation rates were determined by nonlinear fits of the time dependence of the peak intensities. In addition, Monte Carlo simulations were performed to estimate the uncertainty of the relaxation parameters.

**Experimental Constraints and Structure Calculations**

Structures were generated using hybrid distance geometry and dynamical simulated annealing protocol with the CNS 1.0 program on a SGI Indigo$^2$ workstation. Our methodology was similar to that used by Clore and Gronenborn (31, 32) and their coworkers. Distance geometry (DG) substructures were generated using a subset of atoms in the peptide and followed a refinement protocol described in Lee et al (33). The target function for molecular dynamics and energy minimization consisted of a covalent structure, van der Waals repulsion, NOE and torsion angle constraints (34). The torsion angle and NOE constraints were represented by square-well potentials. Based on cross peak intensities in the NOESY spectra with mixing times of 100 and 150 ms, the
distance restraints were then classified as strong, medium or weak corresponding to upper distance bounds of 2.7 Å, 3.3 Å, and 5.0 Å, respectively. An additional 1 Å was added to upper distance bounds for pseudoatom involving non-stereospecifically assigned methylene protons, methyl groups and the ring protons of phenylalanine residue (35). A lower distance bound of 1.8 Å was used for all NOE-derived distance restraints. Structures were calculated using 296 intra-residues, 245 sequential, 124 medium range and 331 long-rang NOE restraints. A total of 76 hydrogen bond restraints were also included in the calculations. Potential hydrogen bond donors were assigned from a $^1$H-$^{15}$N HSQC spectrum recorded immediately after dissolving lyophilized H$_2$O sample to 100% D$_2$O solution. Hydrogen bonds were further identified from characteristic NOE patterns that were observed for residues in regular secondary structure, together with the solvent exchange data. From 2D DQF- COSY (36, 37) and $^{15}$N-edited HNHA spectra, 65 torsion angle restraints were also derived for backbone $\Phi$-angles within elements of secondary structure based on $^3$J$_{HN\alpha}$ coupling-constants ($^3$J$_{HN\alpha}$ $>$ 8, 120 (±50)$^\circ$, $^3$J$_{HN\alpha}$ $<$ 6, 60(±45)$^\circ$). Distance geometry (DG) substructures were generated using a subset of atoms in the peptide, and followed a refinement protocol described in Lee et al (33). All modeling calculations were performed within the InsightII program (Biosym/Molecular Simulations, Inc.) on a SGI Indigo$^2$ workstation.

**RESULTS AND DISCUSSION**

_Circular Dichroism_

Circular dichroic spectra of both wild-type and double mutant proteins in the far-
UV region were collected in 50 mM of sodium phosphate buffer solution at pH 7.0. The spectra suggest that the global folding of the two proteins are similar, showing a major β-strand and minor α-helical contents. However, a small difference at 217 nm was clearly detected, indicating that the structural change of the β-sheet region was due to double mutation of SCM. The additional minima observed at 206-208 nm have been ascribed from the contributions of aromatic side chains (38, 39) (Fig. 1).

**Solution Structures and Sweet Taste**

Spin system assignments were easily made by homonuclear 2D TOCSY and $^{15}$N-edited 3D TOCSY-HSQC spectrum. All identified spin systems served as a starting point for complete sequence-specific resonance assignment procedure. Fig. 2 shows 2D $^1$H-$^{15}$N HSQC spectrum with the assignments. A total of 50 substructures generated from distance geometry algorithms were used as starting structures in the simulated annealing stage. After simulated annealing calculations, the 20 structures ($<$SA$>_k$) showed no constraint violations greater than 0.5 Å for distances and 5° for torsional angles. These structures were used for detailed structural analysis. Table I shows the energy and structural statistics of the final 20 simulated annealing structures related to experimental constraints. The final $<$SA$>_k$ structures were well converged with a root-mean-square deviation of 0.96 Å for all backbone atoms, 0.61 Å for backbone atoms except for disordered regions (Gly1-Trp3, Glu48-Ile53 and Asp76-Arg82) and 0.49 Å for residues involved in secondary structures. The REM average structure ($<$SA$>_{kr}$) from 20 final structures exhibited a root-mean-square deviation of 0.99 Å for backbone atoms with respect to 20 $<$SA$>_k$ structures and 0.64 Å for backbone atoms excluding
unstructured regions (Gly1-Trp3, Glu48-Ile53, Asp76-Arg82). Table II summarizes the structural statistics associated with 20 final $<\text{SA}>_k$ structures of SCM$^{\text{DR}}$. The deviations from idealized geometry are also very small and satisfy ideal geometry. The $<\text{SA}>_k$ structure clearly demonstrates the relative orientations of its major secondary structures, showing that the twisted $\beta$-sheet partially wrapped the beginning of the $\alpha$-helix. Especially, the bend of three strands ($\beta_2$, $\beta_3$ and $\beta_4$) enable to have a close contacts with the $\alpha$-helix. A best-fit backbone superposition of all final $<\text{SA}>_k$ structures with an average REM structure is displayed in Fig. 3. The angular order parameter for $\Phi$, $\Psi$ which indicates the degree of dihedral heterogeneity of structures showed that the $\Phi$, $\Psi$ values of all residues except prolines and residues of Gly30-Thr33 and Ile8 in loop regions are close to 1, suggesting the well defined backbone angles of SCM$^{\text{DR}}$ (data not shown). The backbone torsion angles $\Phi$, $\Psi$ for the $<\text{SA}>_k$ ensured that all of the $\Phi$, $\Psi$ values of final 20 structures were distributed in energetically favorable regions. The ribbon diagram of the REM average structure clearly shows that the general topology consisted of both a single long $\alpha$-helix and a five-stranded antiparallel $\beta$-sheet (Fig. 4). Five loops including an engineered one were also characterized. The side chain orientations of Tyr63 and Asp66 which are common to all sweet peptides and correspond to Phe and Asp residues of aspartame, were observed on the opposite side to the H1 helix and those that were mostly exposed to solvent.

Fig. 5 displays a best-fit superposition of REM average solution structure of SCM$^{\text{DR}}$ over that of the wild type. A high flexibility around an engineered loop and structural rearrangements of $\beta$-sheet due to double mutation were observed, even though
the overall three dimensional structure of the SCM$^{\text{DR}}$ is similar to that of the wild type.

**Stability and Folding of SCM$^{\text{DR}}$ based on Fluorescence Data**

The emission spectra of the SCM at an excitation wavelength of 280 nm in the various GdnHCl concentrations are shown in Fig 6A. The emission spectrum of the SCM was dominated by the tryptophan fluorescence. In the unfolding process, spectra exhibited two emission peaks, corresponding to the contribution of tyrosine and tryptophan residues based on the tyrosine-tryptophan energy transfer mechanism. The efficiency of energy transfer can be used to estimate the distances between the donor and acceptor. The fluorescence energy transfer transition curve was measured using denaturant GdnHCl and the ratio of fluorescence intensities at 298 nm corresponding to tyrosine fluorescence and at 350 nm to tryptophan fluorescence. It can be seen that GdnHCl-induced denaturation begins at 2.0 M concentration and fully denatures at 3.4 M of GdnHCl. Fig. 6B shows that the fractional change of unfolding suggests a two-state model for GdnHCl-induced denaturation of both SCM$^{\text{DR}}$ and wild-type protein. However, the transition midpoint of the double mutant SCM has shifted slightly towards the left, implying that the double mutation induces destabilization of the protein.

In each case, the reversibility of the unfolding reaction was confirmed by obtaining a refolding curve through dilution of the protein from high GdnHCl concentration. The denaturation and renaturation curves are determined to be exactly superimposable.

**Backbone Dynamics of SCM$^{\text{DR}}$**
The rotational diffusion anisotropy was estimated for each residue, yielding a ratio of the principal moments of inertia of SCM\textsuperscript{DR} as 1.00:0.86:0.40. This data suggests that the global shape of SCM\textsuperscript{DR} is a prolate ellipsoid. Using data from \(R_2/R_1\) ratios and structural coordinates of 66 selected residues, axial diffusion tensors, \(D_{\text{z}}/D_1\), were also calculated. The starting values of the parameters \(\tau_m\) and \(D_{\text{z}}/D_1\) were 4.9 ns and 0.65 at 500 MHz, respectively. The standard values of \(R_1\), \(R_2\), and \(^{15}\)N-{\(^{1}\)H} NOEs were fitted in the model-free system selected by an F-test. Fig. 7 demonstrates that the regions of the secondary structures showed higher \(S^2\) values than those of loops, as we expected. The average values of \(S^2\) are above 0.94 for most residues in \(\beta\)-strands and above 0.98 for \(\alpha\)-helical region. More than 80% of the residues in SCM\textsuperscript{DR} have order parameter \(S^2\) values greater than 0.8, indicating that the protein in general is relatively rigid. Five residues of Glu4, Ile5, Thr33, Tyr47, and Arg82 belonged to the loop regions showed relatively higher \(R_2/R_1\) ratios, originated from a significant chemical exchange contribution (Fig. 7A). In addition to this exchange contribution, residues of Glu4 and Ile6 exhibited a significant reduction in the \(^{15}\)N-{\(^{1}\)H} NOEs and \(S^2\) values of 0.7 and 0.69 (Fig. 7B). Thus, it can be supposed that both Glu4 and Ile5 residues might be involved in dynamical motions. Four residues of Ile5, Tyr47, Arg51, and Asp66 located in the loop regions contained \(\tau_e\) contributions, indicating enhanced backbone dynamics on a fast time scale.

In the previous report, we determined that SCM exists as a monomer conformation based on NMR and gel-filtration experiments. Solution structure suggested that Arg70, Arg86 residues have a close correlation with the degree of
sweetness of SCM protein. A comparison of the secondary structure of SCM\textsuperscript{DR} with wild type protein indicates that the structural difference between the two proteins can be observed mainly in β-strands. The first short β-strand composed of residue Glu2-Ile5 found in the wild type SCM was not detected in SCM\textsuperscript{DR}. Amide hydrogen exchange data, backbone-backbone NOEs, and C\textsubscript{α}H chemical shift indices did not provide any evidence of support β-strand in this region. Therefore, it can be supposed that a mutation of Asp7 disrupts the first short strand, perturbating the stability of the network of β-strand. The main structural differences are the following: each β-strand encompasses six residues from Cys41 to Ile46 for a second β-strand, nine residues from Gly55 to Tyr63 for a third, and six residues from Arg70 to Glu75 for a fourth β-strand, whereas seven residues from Cys41 to Tyr47, eleven residues from Lys54 to Ala64 and seven residues from Phe69 to Glu75 constitute each third, fourth, and fifth β-strand in wild type SCM. Several reports have already proposed that a sulfhydryl group of Cys41 in the beginning of β3 strand could be critical for sweetness. In our solution structures of both wild type SCM and SCM\textsuperscript{DR}, the side chain of Cys41 is located on the hydrophobic interface between β-sheet and α-helix. We might think that the side chain of Cys41 plays a role for sweetness because it maintains a bulge between Ile38-Cys41 responsible for structural organization, especially the H1 helix orientation and, furthermore, the tertiary structure of single chain monellin. We also proposed that hydrophobic and/or side chain-side chain interaction related to tertiary structure of SCM is in part responsible for sweetness. Mutational studies of SCM proteins have supported our structural data, showing that the size of Asp7 residue in the loop A is important for
sweetness (Fig. 8). Our structure showed that Asp7 and Arg39 residues which locate on
the same surface of the protein are involved in charge-charge interaction, causing
structural instability from mutations of these residues. This instability is also proven
from the data of thermal unfolding experiments of SCM proteins (Fig 9). The midpoint
of denaturation temperature ($T_m$) of SCM$^{DR}$ was determined to be significantly less than
that of the wild type. In addition, our data suggest that the secondary structures and
stability of tertiary structure were closely correlated with mutations of Asp7 and Arg39
residues. We thought that the H1 helix would be responsible for the general topology as
well as the side chain orientations of the key residues involved in sweetness. Therefore,
we can conclude that even though amino acids located in the loop regions are mainly
involved in biological activity of monellin, both the relative orientations of those side
chains based on tertiary structure and protein stability are of importance for the sweet
taste of the SCM protein.
ACKNOWLEDGEMENTS

The authors would like to thank TMSI Korea for the use of the molecular simulation programs (Molecular Modeling Tools, Molecular Simulations, Inc.). We also thank Dr. Dan Garrett of NIH for providing the programs PIPP and CAPP.
REFERENCES


Abbreviations used: SCM, single-chain monellin; NMR, nuclear magnetic resonance; 1D, one-dimensional; 2D, two-dimensional; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlated spectroscopy; DQF-COSY, double-quantum-filtered correlated spectroscopy; HSQC, heteronuclear single quantum coherence; SA, simulated-annealing
FIGURE LEGENDS

Fig. 1. CD spectra of mutant SCM<sup>DR</sup> and wild type SCM at far-UV region from 190 to 250 nm at pH 7.0. The symbols represent SCM<sup>DR</sup> (○), and wild type SCM(●).

Fig. 2. The 2D <sup>1</sup>H-<sup>15</sup>N HSQC spectrum of uniformly <sup>15</sup>N-labelled double mutant SCM<sup>DR</sup>. Two boxes are indicated as mutated residues. Indole NH cross-peak from Trp3 residue is marked as W(NH) and cross-peaks connected by dotted lines correspond to side chain NH<sub>2</sub> groups of Gln and Asn residues.

Fig. 3. Steroview of the backbone superposition of the energy-minimized average structure ($<SA>$, thick line) over the family of 20 final simulated annealing structures ($<SA>$, thin line).

Fig. 4. Ribbon diagram of the REM average structure displaying ordered secondary structure elements and relative orientation of secondary structures. The side chain atoms for mutated residues are also displayed. The figure was generated with MOLSCRIPT.

Fig. 5. Comparison of the backbone conformation for the SCM<sup>DR</sup> (red) and wild type SCM (yellow).

Fig. 6. Fluorescence emission spectra of SCM<sup>DR</sup> on excitation at 295 nm with
various GdnHCl concentrations (A). Equilibrium denaturation curves of SCM<sup>DR</sup>(○) and wild type SCM(●) are represented in (B). All amplitudes are relative to a value of 0 to <em>f</em><sub>U</sub>. The solid lines are non-linear extrapolation of the equilibrium unfolding data to two-state <em>F</em><em>U</em> transitions.

**Fig. 7.** 15N NMR relaxation parameters and model-free analysis of SCM<sup>DR</sup> (A-E) and wild type SCM (F). (A) R<sub>2</sub>/<R<sub>1</sub> (B) Heteronuclear NOE (XNOE) values plotted against the number of residue. (C) Generalized order parameter S<sup>2</sup>, (D) R<sub>ex</sub> exchange broadening (E) effective correlation times, τ<sub>e</sub> derived from model-free fits. (F) Generalized order parameter S<sup>2</sup> of wild type SCM. The α-helical region is represented by a black bar, the β-strand region by a gray bar and the loop region by a white bar.

**Fig. 8.** Relative sweet activities of various mutant SCM proteins.

**Fig. 9.** The thermal unfolding transitions of wild type SCM (T<sub>m</sub>=359K,●) and SCM<sup>DR</sup> (T<sub>m</sub>=348K,○) monitored by the decrease of the CD signal at 222 nm are displayed. The results of the analysis based on a two state model are shown by the solid lines.
Table I: Structural Statistics for the 20 Final Simulated Annealing Structures of SCM_Dk

<table>
<thead>
<tr>
<th>Parameter</th>
<th>(&lt;\text{SA}&gt;_{k})</th>
<th>(&lt;\text{SA}&gt;_{w})</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Rms deviations from experimental distance restraints(Å)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>all(996)</td>
<td>0.022 ± 0.002</td>
<td>0.022</td>
</tr>
<tr>
<td>sequential (</td>
<td>i-j</td>
<td>= 1 ) (245)</td>
</tr>
<tr>
<td>medium-range(1&lt;</td>
<td>i-j</td>
<td>&lt;5) (124)</td>
</tr>
<tr>
<td>long-range (5 ≤</td>
<td>i-j</td>
<td>) (331)</td>
</tr>
<tr>
<td>intraresidue (296)</td>
<td>0.023 ± 0.003</td>
<td>0.011</td>
</tr>
<tr>
<td>hydrogen bond (76)</td>
<td>0.018 ± 0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>B. rms deviations from experimental dihedral restraints(deg) (65)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.041 ± 0.024</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>C. E_L-J energies(kcal mol(^{-1}))(^a)</td>
<td>-207.00</td>
<td>-217.48</td>
</tr>
<tr>
<td>D. Deviations from idealized covalent geometry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bonds (Å)</td>
<td>0.002 ± 0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>angles(deg)</td>
<td>0.361 ± 0.023</td>
<td>0.309</td>
</tr>
<tr>
<td>impropers(deg)</td>
<td>0.263 ± 0.026</td>
<td>0.178</td>
</tr>
</tbody>
</table>

\(^a\)E_L-J is the Lennard-Jones/van der Waals potential calculated using the CHARMM empirical energy function
Table II: Atomic rms Deviations for the Final Simulated Annealing Structures of SCM^dr (91 Residues)

<table>
<thead>
<tr>
<th>parameters</th>
<th>backbone atoms (Å)</th>
<th>all atoms (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;SA&gt;_k vs &lt;SA&gt;_k</td>
<td>0.96(0.61)^a</td>
<td>1.87(1.48)</td>
</tr>
<tr>
<td>&lt;SA&gt;_kr vs &lt;SA&gt;_k</td>
<td>0.76(0.58)</td>
<td>1.45(1.01)</td>
</tr>
<tr>
<td>&lt;SA&gt;_k vs &lt;SA&gt;_kr</td>
<td>0.99(0.64)</td>
<td>1.93(1.53)</td>
</tr>
</tbody>
</table>

^a Atomic rms deviations were calculated for residues excluding disordered regions (residues 1-3, 48-52 and 78-81).
Fig. 7
Fig. 9

A graph showing the relationship between temperature (K) and the value of $f_U$. The graph includes two curves and data points, indicating changes in $f_U$ with temperature.
Solution structure, backbone dynamics and stability of a double mutant single-chain 
monellin: structural origin of sweetness 
Yoon-Hui Sung, Joon Shin, Ho-Jin Chang, Joong Myung Cho and Weontae Lee 

J. Biol. Chem. published online March 7, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M100930200

Alerts:
  • When this article is cited
  • When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts