Acta Cryst. (1998). D**54**, 494–500

Structure of an Enantiomeric Protein, D-Monellin at 1.8 Å Resolution

LI-WEI HUNG,[a] MASANORI KOHMURA,[b] YASUO ARIYOSHI[b] AND SUNG-HOU KIM[a]*

[a]*Graduate Group in Biophysics, Department of Chemistry and Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA, and* [b]*Central Research Laboratories, Ajinomoto Co., Inc., 1-1 Suzuki-cho, Kawasaki-ku, Kawasaki 210, Japan*

(*Received 27 May 1997; accepted 16 September 1997*)

Abstract

The D-enantiomer of a potently sweet protein, monellin, has been crystallized and analyzed by X-ray crystallography at 1.8 Å resolution. Two crystal forms (I and II) appeared under crystallization conditions similar, but not identical, to the crystallization conditions of natural L-monellin. There are four molecules per asymmetric unit in crystal form I and one in crystal form II. Crystal form I is not reproducible and is equivalent to that of monoclinic L-monellin. Intermonomer contacts in crystal form II are very different from those found in natural L-monellin crystals. The backbone trace of D-monellin resembles very closely the mirror image of that of L-monellin, but the N- and C-terminus backbones as well as several side-chain conformations of D-monellin are different from those of natural L-monellin. Most of these apparent differences may be attributable to the crystal packing differences.

1. Introduction

The three-dimensional structures of a protein and its enantiomer are believed to be mirror images of each other (for a comment, see Petsko, 1992) because of the fact that L- and D-amino acids are exact mirror images of each other. This criterion was further supported by Kent and coworkers (Milton *et al.*, 1992) who synthesized the first all-D enzyme that showed the same catalytic activity as the natural all-L enzyme, but only for substrates of the D-chirality. There have been crystallographic studies of racemic mixtures of rubredoxin at 2 Å resolution (Zawadzke & Berg, 1993), Leu-enkephalin, a pentapeptide, at 0.85 Å resolution (Doi *et al.*, 1994), Trichogin A IV, an undecapeptide, at 0.9 Å resolution (Toniolo *et al.*, 1994), and a duplex of hexa-deoxynucleotides at 2.2 Å resolution (Doi *et al.*, 1993). In all cases, the racemic mixtures crystallized in centrosymmetric space groups and one enantiomer is the exact mirror image of the other. These observations suggest that a D-protein should crystallize in the same unit cell (with opposite handedness) as natural L-protein, provided the crystallization conditions are identical and not chiral specific. However, the unit-cell parameters reported for crystals of synthetic L-HIV-1 protease with inhibitor (Miller *et al.*, 1989) were very different from those of synthetic D-HIV-1 protease with the enantiomeric inhibitor (Kent *et al.*, 1995).

We have chemically synthesized D-monellin, the enantiomer of natural L-monellin. Unlike L-monellin, the D-monellin crystallized in two different unit cells from a similar condition that crystallized monoclinic L-monellin crystals. Crystal form I is equivalent to the monoclinic natural L-monellin crystal form. Crystal form II, to which most D-monellin crystals belong, has completely different intermolecular contacts from both L-monellin crystal structures (Ogata *et al.*, 1987; Bujacz *et al.*, 1997), and from an engineered single-chain L-monellin (SCM; Somoza *et al.*, 1993). In a search for the structural reasons for this observation, we have determined the crystal structure of D-monellin in the high-resolution crystal form.

Monellin is a plant protein with a molecular weight of 12 400 Da, and consists of a 44-residue *A* chain and a 50-residue *B* chain, with no modified amino acids or attached carbohydrates. Natural L-monellin is one of the most potent sweeteners known. It is approximately 70 000 times sweeter than sucrose on a molar basis (Brouwer *et al.*, 1973), or about 3000 times sweeter on a weight basis (Cagan, 1973). It was originally found in the berries of a West African plant *Dioscoreophyllum cumminsii* (Morris & Cagan, 1972; van der Wel, 1972). Experimental evidence shows that the intact three-dimensional structure of monellin is required to elicit the sweet response (*e.g.*, van der Wel, 1972; Bohak & Li, 1976; Morris & Cagan, 1972; Jirgensons, 1976). To understand the structural basis for this protein's extraordinary sweetness, the structures of natural monellin and an engineered single-chain monellin (SCM) have been determined using X-ray crystallography (Ogata *et al.*, 1987; Somoza *et al.*, 1993) and two-dimensional NMR spectroscopy (Tomic *et al.*, 1992).

To study the chiral effect on the sweet taste of monellin, D-monellin, consisting of only α-D-amino acids, has been chemically synthesized. It has an optical activity opposite to that of natural L-monellin and is devoid of any sweetness (Ariyoshi & Kohmura, 1994). It crystallized in two crystal forms under conditions similar, but not identical, to those for natural L-monellin (Table 1). Crystal form I belongs to space group $P2_1$, and

ISSN 0907-4449

Table 1. *Crystallization parameters of* D- *and* L-*monellin*

	Protein concentration in hanging drop (mg ml^{-1})	Buffer concentration in hanging drop (m*M*)	pH	Precipitant drop/reservoir (%)	Space group	Molecules/ asymmetric unit	V_m (Å^3 Da^{-1})
Natural L-monellin	6	10 sodium phosphate	7.2	4/33 PEG 8000	$P2_1$	4	2.4
D-Monellin crystal form I	5	10 sodium phosphate	7.2	14/28 PEG 8000	$P2_1$	4	2.4
D-Monellin crystal form II	5	10 sodium phosphate	7.2	14/28 PEG 8000	$P2$	1	2.2

diffracted to 2.7 Å initially, then degraded rapidly to about 4 Å resolution. Crystal form II of D-monellin is in a rare (Wukovitz & Yeates, 1995) space group, *P*2, and diffracted to 1.7 Å on a 1.08 Å synchrotron radiation source. The structures of both crystal forms have been solved by molecular replacement. The structure of crystal form II was refined at 1.8 Å resolution, but that of crystal form I was refined only as a rigid-body because of the low resolution of the diffraction data.

2. Materials and methods

2.1. *Crystallization*

The *A* and *B* chains of D-enantiomers of monellin were synthesized by the stepwise Fmoc solid-phase method. The renaturation of both chains of the D-enantiomers, and subsequent purification by hydrophobic interaction chromatography (HIC) were performed as described for the preparation of monellin analogs (Kohmura *et al.*, 1992). D-Monellin had the opposite optical activity of L-monellin when tested with circular dichroism, and was essentially tasteless (Ariyoshi & Kohmura, 1994).

D-Monellin was crystallized by the vapor-diffusion method (Hung *et al.*, 1997). The crystallization buffer was 20 m*M* sodium phosphate buffer (pH 7.2) and 28% PEG 8000. 10 mg ml^{-1} synthetic D-monellin solution was mixed with the same volume of the crystallization buffer and equilibrated with the reservoir buffer identical to the crystallization buffer. Crystals appeared only after raising the reservoir stepwise to 40% PEG 8000. Two crystal forms were observed under identical crystallization conditions. Crystals of form I are thick plate-like crystals, with cell parameters a = 39.9, b = 71.9, c = 84.8 Å, $\alpha = \gamma = 90.0$, $\beta = 99.9°$, in space group $P2_1$. However, these were not very reproducible and of poor quality. This crystal form is identical to the monoclinic L-monellin crystals (Ogata *et al.*, 1987; Somoza *et al.*, 1993; space group $P2_1$, a = 39.8, b = 87.2, c = 72.1 Å, γ = 107.3°. This unit cell, when converted to the standard convention, becomes: a = 39.8, b = 72.1, c = 84.4 Å, β = 99.5°). After 10–15 d, rod-like crystals, crystal form II, were also formed under the same conditions, with cell parameters a = 40.59, b =33.03, c = 41.35 Å, $\alpha = \gamma = 90.0$, β = 96.43°, in space group *P*2. This crystal form is very different from any L-monellin crystals reported previously. The typical crystal sizes are 300 × 150 × 80 µm for crystal form I, and 400 × 50 × 40 µm for crystal form II. Among ten D-monellin crystals from which we have collected data sets, only one was found to be in crystal form I. We have not been able to reproduce D-monellin in crystal form I, nor can we crystallize natural L-monellin in the same unit cell as that of D-monellin crystal form II.

2.2. *Data collection and reduction*

Crystal form I diffracted to 2.7 Å initially, then degraded rapidly to 4 Å resolution on an R-AXIS IIC imaging-plate system with a Cu $K\alpha$ X-ray source at 277 K. This data set was processed and integrated by programs *DENZO* and *SCALEPACK* (Minor, 1993; Otwinowski, 1993) to 4 Å resolution with a completeness of 95% and an $R_{\rm merge}$ $[\sum_{hkl,j}|I(hkl,j)-\langle I(hkl)\rangle|/\sum_{hkl}\langle I(hkl)\rangle]$ of 11%. Systematic absence on reflections $(0, 2n + 1, 0)$ suggested the space group should be $P2_1$. Crystal form II diffracted to 1.7 Å at beamline VII-1 at Stanford Synchrotron Radiation Laboratory (SSRL), using an X-ray wavelength of 1.08 Å. A complete data set of crystal form II was collected from one crystal at 277 K on a MAR Research imaging plate. The data set was processed using *DENZO* and integrated by *SCALEPACK* in space group *P*2, with an $R_{\rm merge}$ of 4.5% for all reflections between 30 and 1.8 Å. Since the resolution of crystal form II is substantially higher than crystal form I, the full refinement was performed for crystal form II, and only rigid-body refinement for crystal form I.

2.3. *Structure determination of crystal form I*

Although the data of crystal form I was not sufficient for structural refinement, we were able to study molecular packing from this 4 Å data set. Based on the similarity of this crystal form and natural L-monellin, a molecular replacement using natural L-monellin tetramer structure (Somoza *et al.*, 1993) as the initial model was carried out with the program *AMoRe* (Navaza, 1994). After fast rigid-body refinement, a solution stood out unambiguously from all others. The *R* factor was 31.5% and the correlation coefficient $[\langle I_o I_c - \langle I_o\rangle\langle I_c\rangle\rangle/[\langle I_o^2 - \langle I_o\rangle^2\rangle\langle I_c^2 - \langle I_c\rangle^2\rangle^{1/2}]$ was 73% at this

Table 2. *Summary of final refinement parameters of* D*-monellin crystal form II with natural* L*-monellin and SCM included for comparison*

	D-Monellin form II	Natural L-monellin†	SCM†
Number of non-H protein atoms	755	3136	1566
Number of water molecules	82	0	137
Final crystallographic parameters			
Resolution range (Å)	6.0–1.8	6.0–2.75	6.0–1.7
Number of reflections ($\|F\| > 0\sigma$)	18438	8853	17089
Number of reflections ($\|F\| > 2\sigma$)	17490	—	15053
R factor (%)‡	18.0	19.3	17.4
Free R factor (%)§	22.2	—	—
$\langle B \rangle$ (Å^2)			
All atoms	26.9	47.4	23.6
All protein atoms	25.5	47.4	22.3
Protein backbone atoms	22.8	46.6	19.1
Protein side-chain atoms	27.9	48.1	25.1
Water atoms	41.2	—	39.2
Root-mean-square deviations from ideal geometry			
Bond lengths (Å)	0.009	0.017	0.015
Bond angles (°)	1.61	3.58	2.86

‡ R factor $= \sum_{hkl} \big||F_o(hkl)| - |F_c(hkl)|\big| / \sum_{hkl} |F_o(hkl)|$. § Free R factor $= \sum_{hkl} \big||F_o(hkl, \text{test})| - |F_c(hkl, \text{test})|\big| / \sum_{hkl} |F_o(hkl, \text{test})|$. † Somoza *et al.* (1993).

stage. Further rigid-body refinement with rigid groups stepwise defined starting with the whole tetramer, then two dimers, and finally, each monomer, yielded an R factor of 30.3% for data from 15 to 4 Å. We did not pursue further refinement because of the low data/parameter ratio of this crystal form.

2.4. *Structure determination of crystal form II*

The structure of one of the natural L-monellin monomers (Ogata *et al.*, 1987; Somoza *et al.*, 1993) was used as the initial model to carry out molecular replacement using the program package *AMoRe* (Navaza, 1994). Diffraction data from 10 to 3 Å was used in the search. Since this crystal form has a cell volume about one-quarter of that of natural L-monellin, we assumed that there are two molecules in this D-monellin unit cell, or one molecule in the asymmetric unit. Molecular replacement was carried out in both space groups, $P2$ and $P2_1$. In space group $P2$, the rotation and translation functions revealed a unique solution which is 60% higher than the next best solution with an R factor of 40% and a correlation coefficient of 0.53. This solution also gave good contacts among symmetry-related molecules. No plausible solutions were found in space group $P2_1$.

2.5. *Refinement of crystal form II*

A rigid-body refinement was carried out using the solution of translation search in space group $P2$, and yielded $R = 35\%$ and a correlation coefficient of 0.64 for data between 10 and 3 Å. At this stage, electron-density maps using data between 10 and 2.5 Å were calculated and inspected using *FRODO* (Jones, 1985) and *O* (Jones *et al.*, 1991) for $F_o - F_c$ maps contoured at 3σ and $2F_o - F_c$ maps contoured at 1σ. Obvious discrepancies between electron densities and the model were corrected by manually refitting the model into the density. Regions with missing or weak density were omitted and the first run of positional refinement and simulated annealing (Brünger *et al.*, 1990) with data between 8 and 2 Å were carried out by *X-PLOR* (Brünger, 1992*a*). 10% of the reflections randomly selected from the data, also known as the free-R set (Brünger, 1992*b*), were left out from this and subsequent refinement. After a few cycles of manual refitting and refinement, water molecules were added according to the following criteria. (i) Water molecules are assigned to the isolated and approxi-

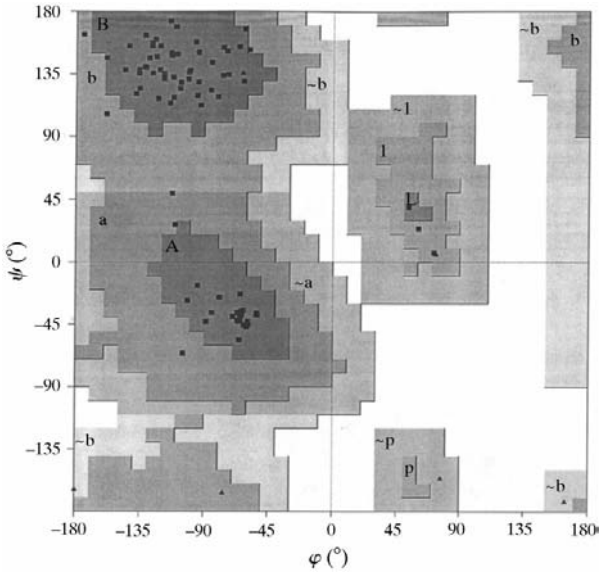

Fig. 1. Ramachandran plot of the final model of D-monellin crystal form II using its mirror image. This plot is prepared with program *PROCHECK* (Laskowski *et al.*, 1993).

mately spherical electron densities greater than 2–3σ in F_o-F_c maps and 1σ in $2F_o-F_c$ electron-density maps; (ii) they make at least two reasonable hydrogen bonds with nearby atoms. The resolution limit was then extended to 1.8 Å. The progress of refinement was monitored by the decrease of free R factor in all stages. The current model contains 755 non-H protein atoms and 82 water molecules with an R factor of 18.0% and free R factor of 22.2%. The average coordinate error estimated from a Luzzati plot (not shown) is 0.2 Å. Residues $A1$, $B49$ and $B50$ are not seen in the electron density and, therefore, were excluded from the model. The final crystallographic and geometric parameters are summarized in Table 2. A Ramachandran plot of the final refined model is shown in Fig. 1.

3. Results and discussion

The crystal structure of D-monellin crystal form II was compared with that of L-monellin in monomer and dimer states. The crystallographic statistics for the current model of D-monellin as well as for L-monellin and SCM are summarized in Table 2.

3.1. *Structural comparison of D-monellin crystal form II, L-monellin and SCM*

A topology diagram of D-monellin is shown in Fig. 2. The mirror image of the main-chain trace of D-monellin is identical to those reported previously for L-monellin and SCM (Somoza *et al.*, 1993) except for the C-terminus of the second β-strand and N-terminus of the third β-strand, corresponding to the flexible and engineered loop of L-monellin and SCM, respectively. When superposed, the mirror image of the D-monellin in crystal form II and natural L-monellin had root-mean-square differences (r.m.s.d.) between coordinates of 0.54 Å for backbone atoms and 1.3 Å for all atoms. Residues $A1$–$A4$ and $B48$–$B50$, the residues of the flexible termini, were excluded from this comparison. To estimate the accuracy of the L-monellin model, including the effects from crystal contacts, we also calculated the r.m.s.d. of backbones of the four L-monellin monomers in an asymmetric unit. The result, 0.40 Å, is not significantly smaller than that between D- and L-monellin structures. Fig. 3(*a*) shows the r.m.s.d. plot between the D- and L-monellin, and Fig. 3(*b*) shows the r.m.s.d. plot for all four L-monellin molecules in an asymmetric unit. Fig. 3(*a*) suggests that, in addition to the flexible terminus regions, the regions of residues $B25$–$B31$ appear to be different between the D- and L-monellin. Most of these residues are located in the loop connecting the α-helix and the second β-strand. Fig. 3(*b*) suggests, however, that this result is not surprising since the r.m.s.d. among four L-monellin molecules is higher around these regions also. Thus, the differences between D- and L-monellin structures are comparable to the experimental error for the L-monellin structure, which is quite large due to the poor resolution of L-monellin diffraction data. We have not been able to obtain high-

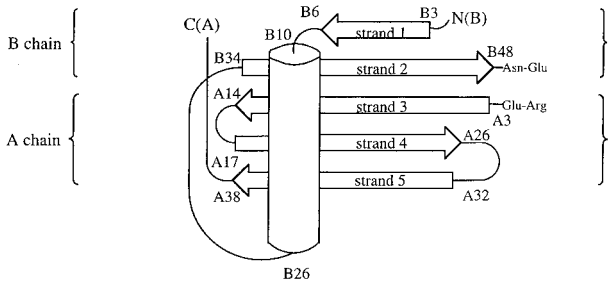

Fig. 2. Schematic drawing of the structure of D-monellin. The helix is represented by a cylinder, and the β-strands are represented by arrows. The beginning and ending residues of each secondary structure are indicated.

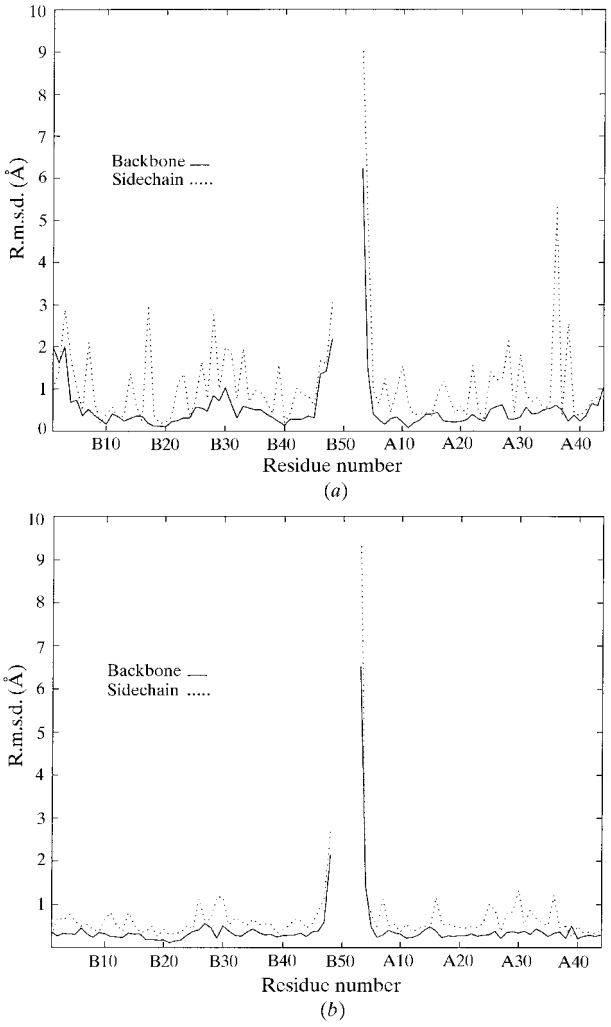

Fig. 3. Comparison of the relative r.m.s.d. of (*a*) D-monellin crystal form II *versus* L-monellin; (*b*) four L-monellin molecules in the asymmetric unit of the $P2_1$ crystal form.

resolution monoclinic L-monellin crystals under the same conditions.

SCM has an identical sequence to L-monellin, except the C-terminus of the *B* chain and N-terminus of the *A* chain of monellin are engineered to be connected covalently, forming a single-chain protein (Kim *et al*., 1989). The SCM structure was solved at a comparably high resolution, 1.7 Å, and was shown to have both structural and biochemical similarities to natural L-monellin (Somoza *et al*., 1993; Kim *et al*., 1989). Although SCM is not exactly the same as L-monellin, we consider it a close and accurate model useful for structural comparisons. The r.m.s.d between SCM and the mirror image of D-monellin is 0.48 Å for backbone atoms and 1.21 Å for all atoms. Residues 48–54, the residues near the engineered loop of SCM, and residues *A*1–*A*4 and *B*48–*B*50 of D-monellin were excluded from this comparison. The r.m.s.d. of the two SCM monomers in an asymmetric unit is 0.45 Å for backbone atoms, and 0.95 Å for all atoms. These results suggest that the backbone trace of the D-monellin resembles that of SCM, and any apparent differences are comparable to the experimental error for the two structures determined at a resolution of approximately 1.8 Å.

3.2. *Dimer interface*

Two types of monellin dimers have been observed in its crystal structures. The monomers in one type are related by a pseudo twofold screw axis as found in crystals of monoclinic L-monellin (Ogata *et al*., 1987), D-monellin crystal form I, SCM (Somoza *et al*., 1993), and the orthorhombic L-monellin crystals (Bujacz *et al.*, 1997). Those in the other type are related by an exact crystallographic twofold axis as found in D-monellin crystal form II only. Fig. 4 shows a stereo α-carbon trace of the mirror image of the dimer of D-monellin crystal form II compared to the L-monellin dimmer. Dimer interaction can also be quantitatively represented by calculating the difference in the accessible area of a molecule as a free monomer and in a dimer. A residue participates more in the dimer contact if the difference is greater. The results of the changes of accessible surfaces of D-monellin, natural L-monellin, and SCM are shown in Figs. 5(*a*), 5(*b*) and 5(*c*), respectively. In the cases of L-monellin and SCM, the majority of the dimer interactions are from contacts of the first and second β-strands (Somoza *et al*., 1993). In addition, the buried residues in the dimer interface of L-monellin and SCM are very similar. On the contrary, 70% of the dimer contacts of D-monellin are contributed by the third and fifth β-strands. This has not been observed in any other structures of the L-monellin family. Dimer contacts of L-monellin and SCM bury about 403 and 530 Å^2 accessible area in the interface, respectively. The D-monellin dimer buries about 505 Å^2.

3.3. *Differences in crystallization behavior between* D- *and* L-*monellin*

Two crystal forms of D-monellin were found under one set of crystallization conditions. These conditions are very similar to those of crystallized natural L-monellin. Crystal form I belongs to space group $P2_1$ with a unit cell equivalent to that of monoclinic L-monellin crystals (Ogata *et al*., 1987). Crystal form II is in a rare space group, *P*2, and the molecular contacts are very different from all L-monellin, and L-SCM structures reported to date. D-Monellin tends to crystallize in crystal form II, and form I is of very low reproducibility. On the other hand, under similar conditions, L-monellin always crystallizes in the unit cell identical to D-monellin form I with opposite handedness. Attempts at crystallizing L-monellin in the same unit cell as D-monellin crystal form II have not been successful. The existence of crystal form I suggests that, as expected, the structures of both enantiomers, regardless of their source and/or purification protocols, are identical or very close to mirror images of each other, especially in the regions participating in intermolecular contacts. However, the reasons for the differences in crystallization behavior under similar crystallization conditions *i.e.*, D-monellin

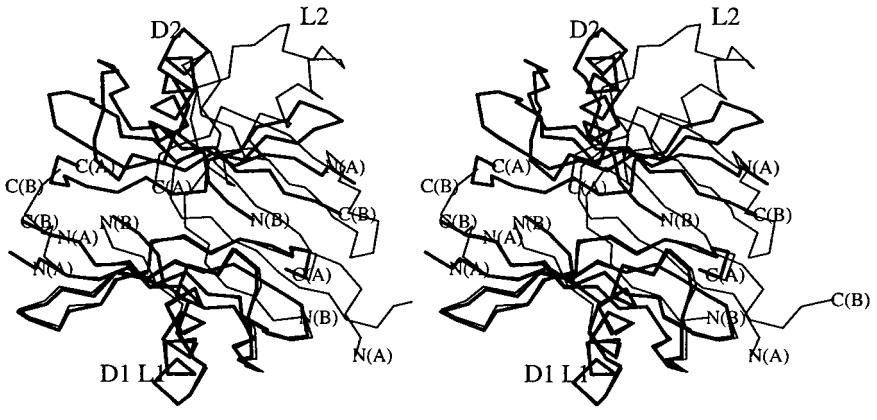

Fig. 4. Stereo Cα traces showing the differences in dimer contact between the mirror image of D-monellin crystal form II (thick lines) and L-monellin pseudo twofold dimer, which is the same as the mirror image of D-monellin crystal form I (thin lines). Molecules D1 and L1 are the ones to be superposed. The positions of D2 and L2 show the differences in molecular contact. N(*A*) and C(*A*) represent the N- and C-termini of the *A* chain, respectively. The same conventions apply to N(*B*) and C(*B*). This figure was produced using *MOLSCRIPT* (Kraulis, 1991).

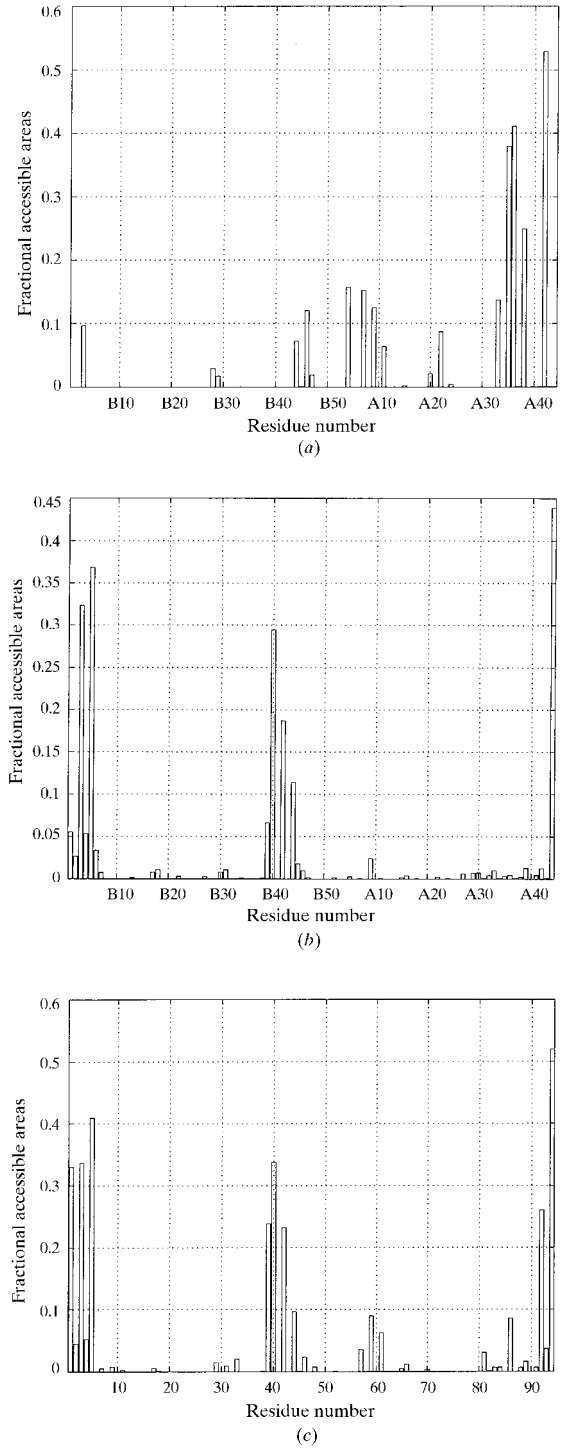

Fig. 5. Comparison of dimer contacts of (*a*) D-monellin crystal form II; (*b*) L-monellin; and (*c*) SCM. These graphs show fractional changes of the solvent accessibility of each residue of one moelcule in the presence and absence of the other molecule in a dimer. The residue numbering conventions of SCM and D- and L-monellin are different. Residues 1–50 and 51–94 of SCM correspond to *B*1–*B*50 and *A*1–*A*44 of monellin, respectively.

favors crystal form II in *P*2, while L-monellin crystallizes exclusively in $P2_1$, are not clear. One possibility is that the binding of one or more chiral 'contaminants' in either or both crystallization experiments may have caused different contact surfaces between the two enantiomers. Since no two crystallization experiments can be identical, one solution to resolve this problem is to obtain high-resolution crystals of racemic DL-monellin in which both enantiomers are under the same environments simultaneously. Under such conditions, differential effects of the binding of traces of chiral material may manifest as structural differences between two enantiomers in the same crystal. We are in the process of finding optimized achiral crystallization conditions of the racemic mixture.

We noticed that all the crystal structures of racemic mixtures so far determined (Zawadzke & Berg, 1993; Doi *et al.*, 1993, 1994; Toniolo *et al.*, 1994) were refined in a centrosymmetric space group, $P\bar{1}$, presumably based on Wilson intensity statistics (Wilson, 1949). This requires imposing exact mirror-image relationships between D- and L-enantiomers during structure refinement. However, small differences between two structures of a stereoisomer pair as a result of the binding of chiral 'contaminants' (for example, side-chain orientation, water network, *etc*.) may neither be sufficient to alter the Wilson statistics from an apparent centrosymmetric distribution nor picked up in powerful modern refinement procedures. To test if any small but significant structural differences exist between D- and L-stereoisomers, one could refine structures in the non-centric space group, *P*1, with both enantiomers in the asymmetric unit. We plan to test this possibility when high-resolution diffraction data of the racemic crystal of monellin becomes available.

We gratefully acknowledge the support of the Energy Biosciences Division, Office of Basic Energy Sciences, and the Helath Effects and Life Sciences Research Division, Office of Health and Environmental Research, US Department of Energy (DE-AC03-76SF00098) and the National Institutes of Health (DC00145), Asahi Chemical Ltd, Ohito, Japan, and Korea Research Institute of Biosciences and Biotechnology (KRIBB), Taejon, Korea.

References

Ariyoshi, Y. & Kohmura, M. (1994). *J. Soc. Synth. Org. Chem. Jpn*, **52**, 359–369.

Bohak, Z. & Li, S.-L. (1976). *Biochim. Biophys. Acta*, **427**, 153–170.

Brouwer, J. N., Hellekant, G., Kasahara, Y., van der Wel, H. & Zotterman, Y. (1973). *Acta Physiol. Scand.* **89**, 550–557.

Brünger, A. T. (1992*a*). *X-PLOR. Version 3.1. A System for X-ray Crystallography and NMR*, Yale University, Connecticut, USA.

Brünger, A. T. (1992*b*). *Nature (London)*, **355**, 472–474.

Brünger, A. T., Krukowski, A. & Erickson, J. W. (1990). *Acta Cryst.* A**46**, 585–593.
Bujacz, G., Miller, M., Harrison, R., Thanki, N., Gilliland, G. L., Ogata, C. M., Kim, S.-H. & Wlodawer, A. (1997). *Acta Cryst.* D**53**, 713–719.
Cagan, R. H. (1973). *Science*, **181**, 32–35.
Doi, M., Inoue, M., Tomoo, K., Ishida, T., Ueda, Y., Akagi, M. & Urata, H. (1993). *J. Am. Chem. Soc.* **115**, 10432–10433.
Doi, M., Ishibe, A., Shinozaki, H., Urata, H. & Ishida, T. (1994). *Int. J. Pept. Protein Res.* **43**, 325–331.
Hung, L.-W., Kohmura, M., Ariyoshi, Y. & Kim, S.-H. (1997). *Acta Cryst.* D**53**, 327–328.
Jirgensons, B. (1976). *Biochim. Biophys. Acta*, **446**, 255–261.
Jones, T. A. (1985). *Methods Enzymol.* **115**, 157–171.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kent, S. B. H., Baca, M., Elder, J., Miller, M., Milton, R., Milton, S., Rao, J. K. M. & Schnölzer, M. (1995). *Breaking the shackles of the genetic code: Engineering retroviral proteases through total chemical synthesis*. In *Aspartic Proteinases*, edited by K. Takahashi, pp. 425–438. New York: Plenum Press.
Kim, S.-H., Kang, C.-H., Kim, R., Kho, J. M., Lee, Y.-M. & Lee, T.-K. (1989). *Protein Eng.* **2**, 571–575.
Kohmura, M., Nio, N. & Ariyoshi, Y. (1992). *Biosci. Biotech. Biochem.* **56**, 1937–1942.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Miller, M., Schneider, J., Sathyanarayana, B. K., Toth, M. V., Marshall, G. R., Clawson, L., Selk, L., Kent, S. B. H. & Wlodawer, A. (1989). *Science*, **246**, 1149–1152.
Milton, R. C., Milton, S. C. & Kent, S. B. H. (1992). *Science*, **256**, 1445–1448.
Minor, W. (1993). *XDISPLAYF Program.* Purdue University, Indiana, USA.
Morris, J. A. & Cagan, R. H. (1972). *Biochim. Biophys. Acta*, **261**, 114–122.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Ogata, C., Hatada, M., Tomlinson, G., Shin, W.-C. & Kim, S.-H. (1987). *Nature (London)*, **328**, 739–742.
Otwinowski, Z. (1993). *Oscillation Data Reduction Program. In Proceedings of the CCP4 Study Weekend: Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
Petsko, G. A. (1992). *Science*, **256**, 1403–1404.
Somoza, J. R., Jiang, F., Tong, L., Kang, C.-H., Cho, J. M. & Kim, S.-H. (1993). *J. Mol. Biol.* **234**, 390–404.
Tomic, M. T., Somoza, J. R., Wemmer, D. E., Park, Y. W., Cho, J. M. & Kim, S.-H. (1992). *J. Biomol. NMR*, **2**, 557–572.
Toniolo, C., Peggion, C., Crisma, M., Formaggio, F., Shui, X. & Eggleston, D. S. (1994). *Nature Struct. Biol*. **1**, 908–914.
van der Wel, H. (1972). *FEBS Lett.* **21**, 88–90.
Wilson, A. J. C. (1949). *Acta Cryst.* **2**, 318–321.
Wukovitz, S. W. & Yeates, T. O. (1995). *Nature Struct. Biol.* **2**, 1062–1067.
Zawadzke, L. E. & Berg, J. M. (1993). *Proteins*, **16**, 301–305.