Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Isomorphous replacement combined with anomalous dispersion in the linear equations: application to a crystal containing four nonapeptide conformers

J. Konnert,[a] J. Karle,[a]* I. L. Karle,[a] K. Uma[b] and P. Balaram[b]

[a]Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, DC 20375-5341, USA, and [b]Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

Correspondence e-mail: williams@harker.nrl.navy.mil

Received 22 June 1998
Accepted 18 August 1998

The investigation of the structure of the four conformers of the nonapeptide described here has an additional purpose: to illustrate a method for combining isomorphous replacement information with anomalous dispersion information within the linear equations that have found use in the analysis of multiple-wavelength anomalous dispersion data. In the present application, isomorphous replacement data were obtained from the replacement of naturally occurring S atoms in the nonapeptide with Se atoms. Only one wavelength was used for the analysis: Cu $K\alpha$ radiation. Details of the analysis are presented, as well as the structural results obtained. It was found that the four independent molecules in the structure have similar, but not identical, conformations. The backbones fold into predominantly α-helices with one or two 3_{10}-type hydrogen bonds and have extended side chains. Three to four water molecules are associated with each of the four head-to-tail regions between the peptides. Optimal packing between hydrophobic surfaces may account for the existence of four molecules in an asymmetric unit.

1. Introduction

This investigation concerns an application of the one-wavelength anomalous dispersion technique involving, in addition, the single isomorphous replacement of sulfur by selenium. The molecule of interest is Boc–Met–Leu–Phe–Ala–Leu–Aib–Ala–Leu–Aib–OMe (I) which crystallizes in space group $P2_1$ with four independent molecules per asymmetric unit. Altogether there are the equivalent of 40 residues plus water solvent, with 310 non-H atoms whose positions need to be determined. A particular feature of this study is the demonstration that the use of a standard laboratory diffractometer with Cu radiation to collect the anomalous diffraction data is sufficiently accurate for the phase determination for structure analysis based on the linear equations for anomalous dispersion (Karle, 1980; 1989*a*,*b*). The linear equations can be written in three different equivalent forms which are related by transformations of variables (Karle, 1989*b*). The analysis involved several steps designed to optimize the accuracy of the results. They concern appropriate scaling of the data, the determination of the placement of the S (Se) atoms, the refinement of these positions and, finally, the determination of phases for the observed structure-factor magnitudes.

The rationale for studying the structure of (I) is based on the following. The tripeptide formyl–Met–Leu–Phe–OH and related sequences have been shown to induce neutrophil activation and chemotaxis (Schiffmann *et al.*, 1975; Showell *et al.*, 1976). Extensions and modifications at the C-terminus do

not have a major influence on activity (Freer *et al.*, 1982). The precise conformational requirements for activity are still unclear, with both β-turn and extended structures appearing to be capable of eliciting neutrophil responses (Prasad *et al.*, 1996). The formyl peptide chemoattractants act through a membrane-bound receptor (Becker, 1987). In a search for analogs which contain a hydrophobic helical membrane-anchoring segment we synthesized the nonapeptide sequence For–Met–Leu–Phe–Ala–Leu–Aib–Ala–Leu–Aib–OMe (II). In (II), the chemotactic tripeptide is grafted onto a well characterized hexapeptide helix (Karle *et al.*, 1989; Karle & Balaram, 1990). This sequence provides an opportunity to examine the effect of the C-terminus helical module on the conformation of the biologically active N-terminal tripeptide. Preliminary studies involving measurements of β-glucosaminidase released from rabbit neutrophil suggested that peptide (II) has activity comparable to the parent tripeptide (E. L. Becker & P. Balaram, unpublished work). The parent peptide formyl–Met–Leu–Phe–OH has been shown to adopt an extended conformation in dimethylsulfoxide solution (Becker *et al.*, 1979). In crystals of methyl and *t*-butyl ester derivatives significantly different conformations have been characterized. Both the Met and Leu residues adopt extended β-strand conformations in For–Met–Leu–Phe–OBu^t (Michel *et al.*, 1990). However, in For–Met–Leu–Phe–OMe, the Met residue is in an extended conformation, while the Leu residue adopts right-handed helical φ,ψ values (Gavuzzo *et al.*, 1989). Interestingly, an early report of the crystal structure of the parent peptide For–Met–Leu–Phe–ON (Morffew & Tickle, 1981; Cambridge Data File code FMLPLA) describes the structure of a diastereomer having D-Phe at position 3. We were interested in exploring the conformational consequences of extending the chemotactic tripeptide sequence by extension of the structure using an obligatory helical sequence (Karle & Balaram, 1990). Single crystals of (II) and the Boc-protected precursor (I) were obtained from aqueous methanol. Only (I) yielded good diffraction data to a resolution of 0.9 Å. Since the structure contained four independent molecules, we synthesized a selenomethionyl analog, Boc–Se-Met–Leu–Phe–Ala–Leu–Aib–Ala–Leu–Aib–OMe (III) to explore the possibility of structure determination using single-wavelength anomalous dispersion.

2. Experimental

2.1. Peptide synthesis

Peptides were synthesized by conventional solution-phase procedures using dicyclohexylcarbodiimide (DCC)–1-hydroxybenzotriazole (HOBt) mediated couplings. The N-terminus and C-terminus were protected using Boc and methyl ester groups. The hexapeptide Boc–Ala–Leu–Aib–Ala–Leu–Aib–OMe, synthesized as described earlier, was deprotected using formic acid (Uma, 1990). The free base, H_2N–Ala–Leu–Aib–Ala–Leu–Aib–OMe, was sequentially extended at the N-terminus using Boc–Phe, Boc–Leu and Boc–Met or Boc–Se-Met in successive couplings.

2.2. Preparation of Boc–Se-Met

100 mg (0.5 mmol) of L-Se-Met, Boc-ON (0.14 g, 0.55 mmol) and triethylamine (0.11 ml, 0.75 mmol) were taken in 50% aqueous dioxan (0.6 ml) and stirred under a nitrogen atmosphere for 24 h. The reaction mixture was diluted with water and washed with ethyl acetate. The aqueous layer was cooled to 273 K, acidified with 2 *M* HCl and extracted with EtOAc. The organic layer was dried over Na_2SO_4 and evaporated to yield Boc–Se-Met as a gum, which was stored under nitrogen at 253 K.

The nonapeptides Boc–Met/Se-Met–Leu–Phe–Ala–Leu–Aib–Ala–Leu–Aib–OMe and the peptide formyl–Met–Leu–Phe–Ala–Leu–Aib–Ala–Leu–Aib–OMe were purified by HPLC using a C_{18} (10 µm, 4 × 250 mm) column on methanol–water gradients. All peptides were shown to be homogeneous by reverse-phase HPLC on an analytical C_{18} (5 µm) column. Peptides were completely characterized by 400 MHz ^{1}H NMR spectra.

Single crystals of formyl–Met–Leu–Phe–Ala–Leu–Aib–Ala–Leu–Aib–OMe (II), Boc–Met–Leu–Phe–Ala–Leu–Aib–Ala–Leu–Aib–OMe (I) and Boc–Se-Met–Leu–Phe–Ala–Leu–Aib–Ala–Leu–Aib–OMe (III) were obtained from methanol–water solutions by slow evaporation.

2.3. Data collection

Air-dried crystals of both (I) and (III) were used at ambient temperatures to collect Friedel pairs of data with Cu $K\alpha$ radiation (λ = 1.54178 Å) on a Siemens automated four-circle diffractometer in the $\theta/2\theta$ mode. Immediately after the measurement of each reciprocal row of *hkl* reflections, the Friedel equivalents *hkl* of that row were measured. Two reflections comprising each Friedel pair were measured within several minutes of each other. The unit-cell parameters for the selenium- and sulfur-containing crystals, space group $P2_1$, are a = 21.016 (6), b = 20.248 (6), c = 31.482 (10), β = 105.04 (2)° and V = 12938 (7) Å^3, and a = 21.030 (5), b = 20.198 (6), c = 31.462 (6) Å, β = 105.16 (2)° and V = 12899 (5) Å^3, respectively. 32461 unique data were measured for the selenium-containing crystal and 21226 for the sulfur-containing crystal.

3. Solution of structure

3.1. Fundamental equations

We express the equations for the Bijvoet pairs of intensities (Karle, 1980, 1989*b*) measured for the sulfur-containing compound, as

$$\begin{aligned}|F^{\rm S}_{\pm h}|^2 = {} & |F^N_h|^2 + \alpha^{\rm S}_h |F^{A,{\rm S}}_h|^2 \\ & + \beta^{\rm S}_h |F^N_h||F^{A,{\rm S}}_h| \cos(\Phi^N_h - \Phi^{A,{\rm S}}_h) \\ & \pm \gamma^{\rm S}_h |F^N_h||F^{A,{\rm S}}_h| \sin(\Phi^N_h - \Phi^{A,{\rm S}}_h), \end{aligned} \quad (1)$$

where sulfur is assumed to be the only anomalously scattering atom, $|F^{A,{\rm S}}_h|$ and $\Phi^{A,{\rm S}}_h$ are the amplitude and phase associated with the structure of the S atoms when those atoms scatter nonanomalously and $|F^N_h|$ and Φ^N_h are the amplitude and

phase associated with the structure of the nonanomalously scattering atoms.

We express the intensities for the isomorphous selenium-containing crystal as a function of the same four parameters by recognizing that $\Phi_h^{A,S} = \Phi_h^{A,Se}$ and $|F_h^{A,Se}| = (f_h^{ASe}/f_h^{AS})|F_h^{A,S}|$ where f_h^{ASe} and f_h^{AS} are the atomic scattering factors for nonanomalous scattering, and thus obtain

$$\begin{aligned}|F_{\pm h}^{Se}|^2 = {} & |F_h^N|^2 + \alpha_h^{Se}|F_h^{A,S}|^2 \\ & + \beta_h^{Se}|F_h^N||F_h^{A,S}|\cos(\Phi_h^N - \Phi_h^{A,S}) \\ & \pm \gamma_h^{Se}|F_h^N||F_h^{A,S}|\sin(\Phi_h^N - \Phi_h^{A,S}). \qquad (2)\end{aligned}$$

The α, β and γ terms are functions of the scattering factors as defined in *Appendix A*, with average values listed in Table 1.

These values suggest that, as a consequence of experimental errors, the cosine terms in (1) and (2) would be determined much more accurately by the intensities than would the sine terms.

3.2. Structure-determination sequence

Equations (1) and (2) express the four intensity measurements for each reciprocal-lattice point as linear functions of four variables, $|F_h^N|^2$, $|F_h^{A,S}|^2$, $|F_h^N||F_h^{A,S}|\cos(\Phi_h^N - \Phi_h^{A,S})$ and $|F_h^N||F_h^{A,S}|\sin(\Phi_h^N - \Phi_h^{A,S})$, subject to the constraint $\sin^2(\Phi_h^N - \Phi_h^{A,S}) + \cos^2(\Phi_h^N - \Phi_h^{A,S}) = 1$. The sequence by which accurate values for these variables were obtained, leading to the desired phases, Φ_h^S and Φ_h^{Se}, is listed here. The sequence was as follows, and the steps are subsequently discussed further.

(i) The intensities for the sulfur- and the selenium-containing compounds were scaled to minimize the effects of systematic errors in the data and to make the intensities from the nonanomalously scattering atoms in both compounds equivalent.

(ii) Equations (1) and (2) were solved with an iterative procedure to obtain approximate values for the four unknowns, with particular attention to identifying an accurate subset of $|F_h^{A,S}|^2$ values, the intensities corresponding to the nonanomalous scattering of the S atoms. Direct methods were applied with the $|F_h^{A,S}|^2$ coefficients to determine positions for the S (Se) atoms. The positional and thermal factors for the S atoms were refined to fit the approximate $|F_h^{A,S}|^2$ data.

(iii) The parameters for the S atoms were further refined to fit a linear combination of the intensities (to be defined later). This yielded accurate values for both $|F_h^{A,S}|^2$ and $|F_h^N||F_h^{A,S}|\cos(\Phi_h^N - \Phi_h^{A,S})$ for all of the data. The difference in intensities between Bijvoet pairs in (1) and (2) yielded values for $|F_h^N||F_h^{A,S}|\sin(\Phi_h^N - \Phi_h^{A,S})$.

(iv) The Φ_h^S and Φ_h^{Se} values were obtained.

3.3. Step (i): scaling

The two main objectives of the data scaling were to correct the data sets for systematic errors which, in the extreme, can reverse the signs of the differences in intensities between Bijvoet pairs and to place the corresponding portions of data for the selenium-containing compound and the sulfur-containing compound (namely, the data for the nonanomalous scatterers) on the same scale.

Table 1
Average values for α, β and γ.

	α	β	γ
Sulfur compound	1.07	2.07	0.120
Selenium compound	5.89	4.85	0.246

Local scaling (McRee, 1993) was used to accomplish these objectives. We scaled each reflection according to the average value for the intensities in a volume of reciprocal space centered on that reflection. The following averages of the experimental intensities, $\langle|F_{h,k,l}^{S,Ex}|^2\rangle$ and $\langle|F_{h,k,l}^{Se,Ex}|^2\rangle$ were defined as

$$\langle|F_{h,k,l}^{S,Ex}|^2\rangle = \left\langle \sum_{hh=h-n}^{h+n}\sum_{kk=k-n}^{k+n}\sum_{ll=l-n}^{l+n} \Delta_{hh,kk,ll}|F_{hh,kk,ll}^{S,Ex}|^2 \right\rangle \qquad (3)$$

and

$$\langle|F_{h,k,l}^{Se,Ex}|^2\rangle = \left\langle \sum_{hh=h-n}^{h+n}\sum_{kk=k-n}^{k+n}\sum_{ll=l-n}^{l+n} \Delta_{hh,kk,ll}|F_{hh,kk,ll}^{Se,Ex}|^2 \right\rangle, \qquad (4)$$

respectively, where $\Delta_{hh,kk,ll}$ is equal to one when intensity measurements are available for Bijvoet pairs for both the sulfur- and the selenium-containing crystals and zero otherwise. The indicated averages were taken over only those terms for which $\Delta_{hh,kk,ll}$ was equal to one. These averages are of interest because they can be used to correct, to a useful approximation, the individual central reflections for the effects of some systematic errors in the data while placing the data for nonanomalously scattering atoms in the the sulfur- and selenium-containing crystals on the same scale.

This is achieved by multiplication of the central reflections by ratios of the appropriate averages. These ratios consist of $|F_{h,k,l}^{S,Ex}|^2$ divided by the individual averages associated with the central reflections being rescaled. Additionally, the data for the selenium-containing crystal must be multiplied by a factor R_h so that the contribution from the nonanomalously scattering atoms is the same for both the sulfur- and selenium-substituted molecules.

We thus have

$$|F_h^S|^2 = |F_h^{S,Ex}|^2, \qquad (5)$$

$$|F_{-h}^S|^2 = |F_{-h}^{S,Ex}|^2(\langle|F_h^{S,Ex}|^2\rangle/\langle|F_{-h}^{S,Ex}|^2\rangle), \qquad (6)$$

$$|F_h^{Se}|^2 = |F_h^{Se,Ex}|^2 R_h(\langle|F_h^{S,Ex}|^2\rangle/\langle|F_h^{Se,Ex}|^2\rangle), \qquad (7)$$

$$|F_{-h}^{Se}|^2 = |F_{-h}^{Se,Ex}|^2 R_{-h}(\langle|F_h^{S,Ex}|^2\rangle/\langle|F_{-h}^{Se,Ex}|^2\rangle), \qquad (8)$$

where

$$R_h = \left[\sum_{\substack{i=\text{atoms in}\\ \text{Se cpd}}}(f_{h,i}T_{h,i})^2\right] \Bigg/ \left[\sum_{\substack{i=\text{atoms in}\\ \text{S cpd}}}(f_{h,i}T_{h,i})^2\right]. \qquad (9)$$

The atomic scattering factors and thermal factors for atom i are $f_{h,i}$ and $T_{h,i}$ respectively. The initial scaling, when the

thermal parameters for the S and Se atoms were not known, was carried out with

$$R_h = \left(\sum_{\substack{i=\text{atoms in} \\ \text{Se cpd}}} Z_i^2 \right) \Big/ \left(\sum_{\substack{i=\text{atoms in} \\ \text{S cpd}}} Z_i^2 \right) = 1.248. \qquad (10)$$

As will be discussed later, when the Se (S) atom positions were approximately determined, a refinement of the Se (S) atomic positional and thermal parameters was carried out before the linear equations were evaluated to determine the desired phases. Use of the Se (S) thermal parameters, along with an overall thermal factor for the nonanomalously scattering atoms determined from the *K*-curve for the S crystal, provided a more accurate scaling of the data by expressing the values for R_h as functions of scattering angle.

The averages indicated in (3) and (4) were computed with $n = 4$ and contained an average of 423 terms with a minimum of 69 and a maximum of 712 terms.

3.4. Step (ii): positions of S and Se atoms

As mentioned, (1) and (2) express the four intensity measurements for each reciprocal-lattice point as linear functions of four variables $|F_h^N|^2$, $|F_h^{A,S}|^2$, $|F_h^N||F_h^{A,S}|\cos(\Phi_h^N - \Phi_h^{A,S})$ and $|F_h^N||F_h^{A,S}|\sin(\Phi_h^N - \Phi_h^{A,S})$. Our first goal was to obtain approximate values of the four unknowns with particular attention to identifying an accurate subset of $|F_h^{A,S}|^2$ values, the intensities corresponding to nonanomalous scattering of only the S atoms.

The fourth variable depends only on the difference in intensities of the Bijvoet pairs and was evaluated first:

$$\begin{aligned} |F_h^N||F_h^{A,S}|\sin(\Phi_h^N - \Phi_h^{A,S}) &= (|F_h^S|^2 - |F_{-h}^S|^2)/2\gamma_h^S \\ &= (|F_h^{Se}|^2 - |F_{-h}^{Se}|^2)/2\gamma_h^{Se}. \end{aligned} \qquad (11)$$

Only those quartets of intensities were retained for which the value indicated for $|F_h^N||F_h^{A,S}|\sin(\Phi_h^N - \Phi_h^{A,S})$ was greater than two times the standard deviation for this term computed from the errors estimated for the experimental intensities. Reference to Table 1 indicates that $|F_h^N||F_h^{A,S}|\sin(\Phi_h^N - \Phi_h^{A,S})$ is the least accurately determined parameter. Therefore, by choosing a subset of the data for which this parameter is fairly well determined, all of the other parameters for this subset should be, in general, even more accurately determined.

The solution for the three remaining variables was carried out in an iterative fashion.

The first step was to approximate $|F_h^{A,S}|^2$ with a very small value. With this parameter specified, estimates for the remaining two unknowns, $|F_h^N|^2$ and $|F_h^N||F_h^{A,S}| \times \cos(\Phi_h^N - \Phi_h^{A,S})$, were obtained by solving the following two equations which are obtained from summing the two equations (1) and the two equations (2) for a particular h:

$$\begin{aligned} |F_h^N|^2 + \beta_h^S|F_h^N||F_h^{A,S}|\cos(\Phi_h^N - \Phi_h^{A,S}) \\ = (|F_h^S|^2 + |F_{-h}^S|^2 - 2\alpha_h^S|F_h^{A,S}|^2)/2, \end{aligned} \qquad (12)$$

$$\begin{aligned} |F_h^N|^2 + \beta_h^{Se}|F_h^N||F_h^{A,S}|\cos(\Phi_h^N - \Phi_h^{A,S}) \\ = (|F_h^{Se}|^2 + |F_{-h}^{Se}|^2 - 2\alpha_h^{Se}|F_h^{A,S}|^2)/2. \end{aligned} \qquad (13)$$

Since estimates for $|F_h^N||F_h^{A,S}|\sin(\Phi_h^N - \Phi_h^{A,S})$, $|F_h^N||F_h^{A,S}|\cos(\Phi_h^N - \Phi_h^{A,S})$, $|F_h^{A,S}|$ and $|F_h^N|$ had already been obtained, the value for $\sin^2(\Phi_h^N - \Phi_h^{A,S}) + \cos^2(\Phi_h^N - \Phi_h^{A,S})$ could be computed. The iterative solution was then carried out by incrementing the value for $|F_h^{A,S}|^2$ until the solution gave the value of $\sin^2 + \cos^2$ to be 1.000 ± 0.005 or until successive values diverged from 1.0. The incremented value for $|F_h^{A,S}|^2$ was accepted if the following three conditions were satisfied: $\sin^2 + \cos^2$ converged to a suitable value, $|F_h^{N,S}|^2$ was positive and the average of the four intensities computed with (1) and (2) agreed with the average of the experimental values to within 5%. Of the 8080 quartets of experimental intensities, 1545 passed the above criteria.

Direct methods were carried out on the 1545 normalized structure-factor magnitudes, $|E_h|$, associated with these $|F_h^{N,S}|^2$. The top four peaks in the first E map proved to be the sulfur positions. Refinement of these four sulfur positions to fit the 1545 $|F_h^{N,S}|$ yielded an unweighted R of 0.327.

3.5. Step (iii): accurate refinement of S and Se atoms

After refining the S-atom parameters to fit the values obtained for the 1545 S-atom structure-factor amplitudes, more accurate S-atom parameters were obtained by refining them to fit a linear combination of the intensities. Because this refinement determines values for $|F_h^N||F_h^{A,S}|\cos(\Phi_h^N - \Phi_h^{A,S})$, $|F_h^{A,S}|^2$ and $\Phi_h^{A,S}$, its importance must be emphasized. The S-atom parameters were refined to fit the known left-hand side of the following equation which is formed by subtracting (12) from (13),

$$\begin{aligned} &\frac{(|F_h^{Se}|^2 + |F_{-h}^{Se}|^2 - |F_h^S|^2 - |F_{-h}^S|^2)}{[2(\alpha_h^{Se} - \alpha_h^S)]} \\ &= \frac{|F_h^{A,S}|^2 + (\beta_h^{Se} - \beta_h^S)|F_h^N||F_h^{A,S}|\cos(\Phi_h^N - \Phi_h^{A,S})}{(\alpha_h^{Se} - \alpha_h^S)}. \end{aligned} \qquad (14)$$

The second term on the right-hand side of (14) oscillates about zero with the result that $(\beta_h^{Se} - \beta_h^S)|F_h^N||F_h^{A,S}| \times \cos(\Phi_h^N - \Phi_h^{A,S})/(\alpha_h^{Se} - \alpha_h^S)$ may, to a good approximation, be considered as random noise added on to $|F_h^{A,S}|^2$. Thus, refinement of the 36 S-atom parameters to fit the left-hand side of (14) yields accurate estimates of $|F_h^{A,S}|^2$. This refinement of the S-atom parameters is carried out against all 8080 data.

We digress a little to discuss both the accuracy with which the S-atom parameters may be obtained using (14) and the accuracy of the associated assumption that the cosine term contributes essentially random noise to the refinement. To accomplish this, we have evaluated the terms on the right-hand side of (14) by computing values for the terms using the refined parameters for the structure. The average absolute value for the $(\beta_h^{Se} - \beta_h^S)|F_h^N||F_h^{A,S}|\cos(\Phi_h^N - \Phi_h^{A,S})/(\alpha_h^{Se} - \alpha_h^S)$ term is 1.24 times the average value for the $|F_h^{A,S}|^2$ term. However, the average value for the $(\beta_h^{Se} - \beta_h^S)|F_h^N||F_h^{A,S}|$

Table 2
Comparison of the phases of the refined structure with those obtained with the linear equations.

$\Delta\Phi$ represents average error in degrees.

$\sin\theta/\lambda$	$\Delta\Phi_h^{Se}$	$\Delta\Phi_h^{S}$	$\Delta\Phi_h^{A,S}$	Number phased	Total number	Fraction phased
0.000–0.167	29.7	32.2	13.1	494	541	0.913
0.167–0.200	27.2	30.3	13.9	359	404	0.889
0.200–0.250	32.9	35.2	15.1	728	835	0.872
0.250–0.330	30.6	36.5	15.3	1497	1985	0.754
0.330–0.420	35.8	42.6	15.0	1385	2648	0.523
0.420–0.531	45.9	51.1	15.3	454	1667	0.272
Intensity	$\Delta\Phi_h^{Se}$	$\Delta\Phi_h^{S}$	$\Delta\Phi_h^{A,S}$	Number phased	Total number	Fraction phased
0.000–44.0	31.8	48.9	13.0	581	1377	0.422
44.0–77.0	36.9	45.0	14.7	622	1370	0.454
77.0–119.0	37.9	42.4	15.8	692	1313	0.527
119.0–201.0	36.7	39.3	16.4	831	1340	0.620
201.0–439.0	29.9	32.8	14.7	1017	1333	0.763
439.0–9346.0	30.7	32.4	14.3	1174	1347	0.872

$\times \cos(\Phi_h^N - \Phi_h^{A,S})/(\alpha_h^{Se} - \alpha_h^S)$ term is only 0.04 times the average value for the $|F_h^{A,S}|^2$ term. Because of the relatively large contributions of the cosine-containing terms in (14), the unweighted R factor is 0.68 for the fit of the S-atom parameters to the left-hand side of (14). Despite the large R factor, the S-atom parameters are considerably more accurately determined than they were when the refinement was based on the estimates for 1545 values for $|F_h^{A,S}|^2$ obtained from the linear equations. In fact, the R value comparing the estimated values for $|F_h^{A,S}|$ with those computed from the refined structure is 0.21 for all 8080 data when the parameters are obtained by fitting the left-hand side of (14) and 0.44 for all 8080 data when the parameters are obtained by the preliminary fit to the 1545 data. The R values here are defined as

$$R = \left(\sum \left||F_h^{A,S}|_{Eq} - |F_h^{A,S}|_{Ref}\right|\right) / \sum |F_h^{A,S}|_{Ref}, \qquad (15)$$

where the subscript Eq refers to the amplitudes computed from the S-atom parameters obtained with the equations described earlier in steps (iii) and (ii), and Ref refers to the amplitudes computed from the refined S-atom parameters.

As mentioned, the refinement based on (14) also determines directly the values for the $|F_h^N||F_h^{A,S}|\cos(\Phi_h^N - \Phi_h^{A,S})$ terms. The fourth and final parameter was then estimated with the following equations,

$$\begin{aligned} |F_h^N|^2 &= [(|F_h^S|^2 + |F_{-h}^S|^2)/2] \\ &\quad - \beta_h^S |F_h^N||F_h^{A,S}|\cos(\Phi_h^N - \Phi_h^{A,S}) - \alpha_h^S |F_h^{A,S}|^2 \\ &= [(|F_h^{Se}|^2 + |F_{-h}^{Se}|^2)/2] \\ &\quad - \beta_h^{Se} |F_h^N||F_h^{A,S}|\cos(\Phi_h^N - \Phi_h^{A,S}) - \alpha_h^{Se} |F_h^{A,S}|^2 \end{aligned} \qquad (16)$$

3.6. Step (iv): values for Φ_h^S and Φ_h^{Se}

At this stage, approximations had been obtained for $|F_h^N|$, $|F_h^{A,S}|$, $\Phi_h^{A,S}$, $\cos(\Phi_h^N - \Phi_h^{A,S})$ and $\sin(\Phi_h^N - \Phi_h^{A,S})$. Using these previously secured values, we first obtained Φ_h^N using

$$\Phi_h^N = \tan^{-1}[\sin(\Phi_h^N - \Phi_h^{A,S})/\cos(\Phi_h^N - \Phi_h^{A,S})] + \Phi_h^{A,S} \qquad (17).$$

The desired phases could now be obtained by using the structure-factor equation, $F_h^S = F_h^N + F_h^{A,S}$, written in the form

$$\begin{aligned} |F_h^S|\exp(i\Phi_h^S) &= |F_h^N|\exp(i\Phi_h^N) \\ &\quad + r_h^S |F_h^{A,S}|\exp[i(\Phi_h^{A,S} + \delta_S)], \end{aligned} \qquad (18)$$

where r_h^S is a statistical correction so that $r_h^S|F_h^{A,S}|$ approximates $F_h^{A,S}$, the structure factor for the S atoms alone that includes anomalous scattering. The terms δ_S is the phase change associated with the anomalous scattering and is defined in *Appendix A*. Equating the real parts of (18) gives

$$|F_h^S|\cos\Phi_h^S = |F_h^N|\cos\Phi_h^N + r_h^S|F_h^{A,S}|\cos(\Phi_h^{A,S} + \delta_S) \qquad (19)$$

and equating the imaginary parts gives

$$|F_h^S|\sin\Phi_h^S = |F_h^N|\sin\Phi_h^N + r_h^S|F_h^{A,S}|\sin(\Phi_h^{A,S} + \delta_S). \qquad (20)$$

The Φ_h^S values were obtained by dividing (20) by (19) and taking the arctangent:

$$\Phi_h^S = \tan^{-1}(\sin\Phi_h^S/\cos\Phi_h^S). \qquad (21)$$

Using equations analogous to (18), (19), (20) and (21), values for Φ_h^{Se} were obtained.

Referring to (14), it can be seen that the accuracy with which the cosine in (17) is determined depends to a great extent on the accuracy with which the linear combination of intensities, $(|F_{+h}^{Se}|^2 + |F_{-h}^{Se}|^2 - |F_{+h}^S|^2 - |F_{-h}^S|^2)$, is measured. Likewise the accuracy with which the difference between the intensities of the Bijvoet pairs, $(|F_{+h}^{Se}|^2 - |F_{-h}^{Se}|^2)$, is measured determines to a great extent the accuracy with which the sine term in (17) is estimated. The estimated standard deviations in both these linear combinations of intensities were estimated using the estimated errors for the individual intensities.

The Fourier map, from which the atomic positions were evident, was obtained with a set of 4917 phases for which either $(|F_{+h}^{Se}|^2 + |F_{-h}^{Se}|^2 - |F_{+h}^S|^2 - |F_{-h}^S|^2)$ or $(|F_{+h}^{Se}|^2 - |F_{-h}^{Se}|^2)$ was greater than one standard deviation.

Table 2 summarizes the accuracy of the derived phases as both a function of $\sin\theta/\lambda$ and of the intensity for both Φ_h^S and Φ_h^{Se}. The Φ_h^{Se} were used to compute the E map from which the atomic positions were obtained. Also displayed in Table 2 is a comparison of the phases for the S (Se) atoms only, computed with the final refined atomic parameters and with the refine-

ment to fit the linear combination of intensities, (14). The average phase-angle error for 4917 reflections of the Se-containing crystal is 33.5°, the average phase-angle error for the reflections of the S-containing crystal is 38.5° and the average phase-angle error for the components of the structure factors associated with the S (Se) atoms is 14.9°.

4. Refinement of structure

Initial least-squares refinement cycles were performed on the four independent molecules of the sulfur-containing peptide using F data. The water-molecule sites were located in difference maps. Final cycles of anisotropic least-squares refinement were performed using F^2 values (program *SHELXTL* 5.03) on 2782 variables and 21226 data to 0.9 Å resolution. 694 H atoms in idealized positions riding on the C or N atom to which each is bonded were included. No other restraints were applied. The occupancy of several water molecules appeared to be near 0.5. The final R factors for the sulfur-containing crystal are R_1 = 7.4% (for F values for 12362 data observed > 4σ), wR_2 = 20.7% (for F^2 values) and S (goodness of fit) = 1.03. Least-squares refinement for the selenium-containing analog resulted in R_1 = 9.2% (for F values

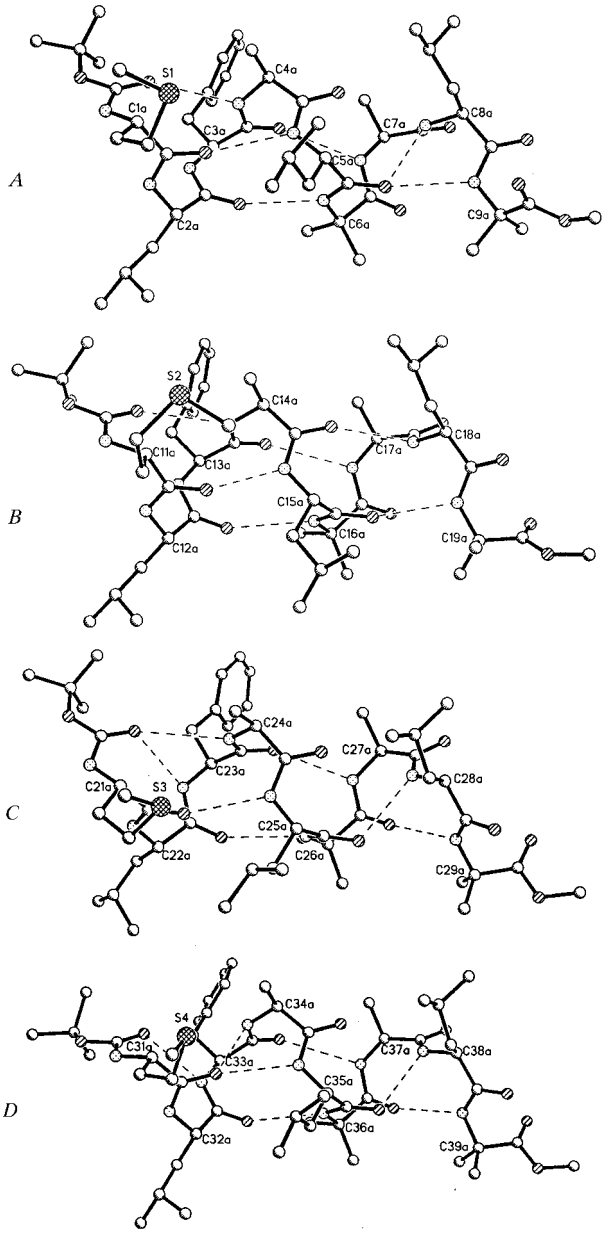

Figure 1
Comparison of the four independent conformers in the sulfur-containing crystal. Dashed lines indicate intrahelical hydrogen bonds. The helices are mixed $3_{10}/\alpha$-type, although predominantly α-type.

Table 3
Torsional angles in the four conformers.

Torsional angles φ, ψ and ω for the backbones and χ_n for the side chains follow the conventions presented in the IUPAC–IUB Commission on Biochemical Nomenclature (1970). The estimated standard deviations are ∼1.5°.

Backbone

		A	*B*	*C*	*D*
Boc	ψ_0	+169	+179	−174	−174
	ω_0	−170	−173	−167	−175
Met	φ_1	−69	−70	−64	−60
	ψ_1	−39	−25	−36	−32
	ω_1	−180	+179	+180	−178
Leu	φ_2	−64	−66	−59	−63
	ψ_2	−44	−38	−38	−26
	ω_2	+174	+176	−180	+178
Phe	φ_3	−50	−63	−72	−66
	ψ_3	−55	−49	−45	−43
	ω_3	−170	+180	+179	+176
Ala	φ_4	−64	−62	−57	−60
	ψ_4	−45	−41	−48	−46
	ω_4	−176	−179	−180	−177
Leu	φ_5	−72	−63	−64	−67
	ψ_5	−41	−46	−47	−43
	ω_5	+178	+180	−177	+178
Aib	φ_6	−54	−55	−60	−54
	ψ_6	−45	−43	−35	−40
	ω_6	−174	−176	+179	−177
Ala	φ_7	−64	−69	−66	−65
	ψ_7	−36	−32	−21	−26
	ω_7	−172	−167	−178	−177
Leu	φ_8	−86	−109	−96	−91
	ψ_8	−32	−8	+7	+2
	ω_8	−176	+178	−179	−178
Aib	φ_9	+56	−52	−52	−53
	ψ_9	−144	+138	+142	+140
	ω_9	−173	+177	+176	+177

Side chains

		A	*B*	*C*	*D*
Met1	χ_1	−170	−72	−177	+178
	χ_2	+72	−61	+68	+83
	χ_3	+68	−54	+74	+106
Leu2	χ_1	−62	−61	−75	−67
	χ_2	−59	−53	−70	−58
		+176	−179	−173	+179
Phe3	χ_1	+174	+173	−166	−178
	χ_2	−85	+81	+67	−82
		+95	−102	−124	+103
Leu5	χ_1	−61	−174	−70	−65
	χ_2	−49	+60	−65	−52
		−178	−177	+171	−179
Leu8	χ_1	−61	−61	−67	−66
	χ_2	−65	−59	−63	−68
		+173	+174	+170	+168

Table 4
Hydrogen bonds.

Type	Donor	Acceptor	N—O or O—O (Å)	H—O (Å)†	N—O═C angle (Å)
Molecule *A*					
Head-to-tail	N1	O17‡	2.941	2.05	171
Peptide–water	N2	W2‡[i]	3.198	2.30	
	N3	W3[ii]	2.943	2.10	
5→1	N4	O0	3.146	2.29	153
	N5	O1	3.043	2.17	158
	N6	O2	2.882	2.01	157
	N7	O3	3.073	2.28	151
4→1	N8§	O5	3.185	2.41	112
5→1	N9	O5	3.072	2.21	164
Water–peptide	W7	O5	3.150		
	W9	O6	2.752		
	W1	O7	2.841		
	W6	O9	2.860		
	W7	O9	2.814		
Molecule *B*					
Head-to-tail	N11	O8	2.998	2.15	128
Peptide–water	N12	W1	2.987	2.10	
	N13	W10[iii]	2.897	2.23	
5→1	N14	O012	3.308	2.44	156
	N15	O11	2.844	2.02	151
	N16	O12	3.039	2.16	151
	N17	O13	3.209	2.44	153
	N18	O14	3.099	2.46	155
4→1	N18	O15	3.084	2.32	112
5→1	N19	O15	2.876	2.19	163
Water–peptide	W13	O16[iv]	2.565		
	W2	O17	2.853		
	W4	O18[v]	3.006		
Molecule *C*					
Head-to-tail	N21	O37	3.016	2.12	176
Peptide–water	N22	W5	3.060	2.16	
4→1	N23	O022	3.113	2.40	127
5→1	N24	O022	3.282	2.40	156
	N25	O21	2.976	2.08	160
	N26	O22	3.129	2.27	141
	N27	O23	3.045	2.23	153
4→1	N28	O25	2.971	2.16	119
	N29	O26	3.256	2.38	112
Water–peptide	W6[v]	O26	2.941		
	W11[i]	O27	3.091		
	W12[vi]	O28	2.679		
Molecule *D*					
Head-to-tail	N31	O27[i]	2.852	1.98	172
	N32	W11	3.033	2.19	
4→1	N33	O031	2.971	2.14	128
	N34	O31	3.129	2.63	117
5→1	N35	O31	3.116	2.26	160
	N36	O32	2.882	2.01	149
	N37	O33	3.276	2.47	147
4→1	N38	O35	2.950	2.22	121
	N39	O36	3.208	2.34	114
Water–peptide	W4	O36	2.774		
	W5	O37	2.980		
	W3	O38	2.871		
Water–water					
	W1	W10[iii]	2.609		
	W9	W1	2.781		
	W12[ii]	W1	2.917		
	W2	W3[iii]	2.878		
	W13[vii]	W2	2.686		
	W5	W4	2.940		
	W11	W6	2.582		

† H atoms were placed in idealized positions with N—H = 0.90 Å. § N8—O4 = 3.453 Å. ‡ O atoms in water molecules (W*n*) were named O*ns* in the coordinate tables. Symmetry equivalents: (i) $x, y, 1+z$; (ii) $2-x, 0.5+y, -z$; (iii) $2-x, 0.5+y, -1-z$; (iv) $1-x, -0.5+y, -z$; (v) $1-x, -0.5+y, -1-z$; (vi) $x, y, -1+z$; (vii) $1-x, 0.5+y, -z$. W10 and W12 mutually exclusive; occupancy 0.5 each.

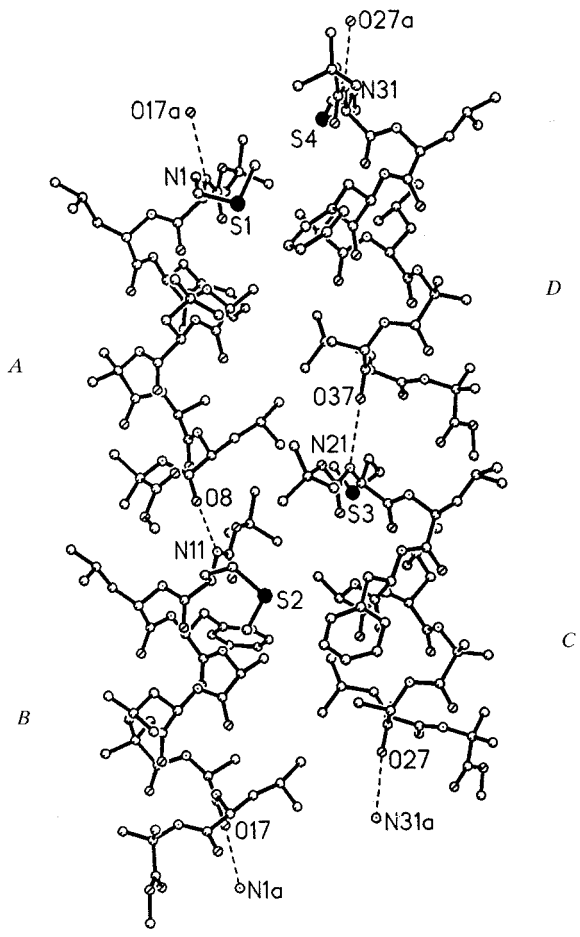

Figure 2
Stacking of conformers $A\rightarrow B\rightarrow A\rightarrow B$ and $D\rightarrow C\rightarrow D\rightarrow C$ into two distinct columns. The one direct NH···O hydrogen bond in each of the four head-to-tail regions is shown by dashed lines. Water molecules in the head-to-tail regions are not shown. One asymmetric unit is shown with *c* axis vertical and *b* axis horizontal. Lateral packing is optimized by bulges fitting into grooves.

for 18602 data observed $> 4\sigma$), $wR_2 = 27.5\%$ (for F^2 values) and $S = 0.913$.

5. Discussion and results

The four independent molecules in the structure have similar, but not identical, conformations. The backbones fold into predominantly α-helices with one or two 3_{10}-type hydrogen bonds and have extended side chains. Three to four water molecules are associated with each of the four head-to-tail regions between the peptides.

5.1. Structure of the sulfur analog

The four conformers of Boc–Met–Leu–Phe–Ala–Leu–Aib–Ala–Leu–Aib–OMe are aligned side-by-side in Fig. 1. The numbers in the labels of equivalent C, N and O atoms in the four molecules are increased by ten in *B*, 20 in *C* and 30 in *D*. Torsional angles are listed in Table 3 and hydrogen-bond parameters are listed in Table 4. The backbones of molecules *A*, *B*, *C* and *D* are quite similar to each other with a predominant α-helix folding (5→1 type intramolecular hydrogen bonding) for residues 1–7. At residue 8, molecule *B* continues the α-helix motif while molecules *A*, *C* and *D* form a 4→1 hydrogen bond characteristic of a 3_{10} helix. Carbonyl O atoms O4, O24 and O34, which do not form hydrogen bonds with N8, N28 and N38, do not participate in any intra- or intermolecular hydrogen bonding. Although most of the N···O distances involved in the intrahelical hydrogen bond occur in the usual 2.9–3.1 Å range, some are as large as 3.31 Å. The longer distances occur for N(14)···O(012) in *B*, N(24)···O(022) and N(29)···O(26) in *C* and N(37)···O(33) and N(39)···O(36) in *D*. The small distortions and differences among the four helices with the same peptide sequence are probably caused by the somewhat different environment around each molecule in the crystal.

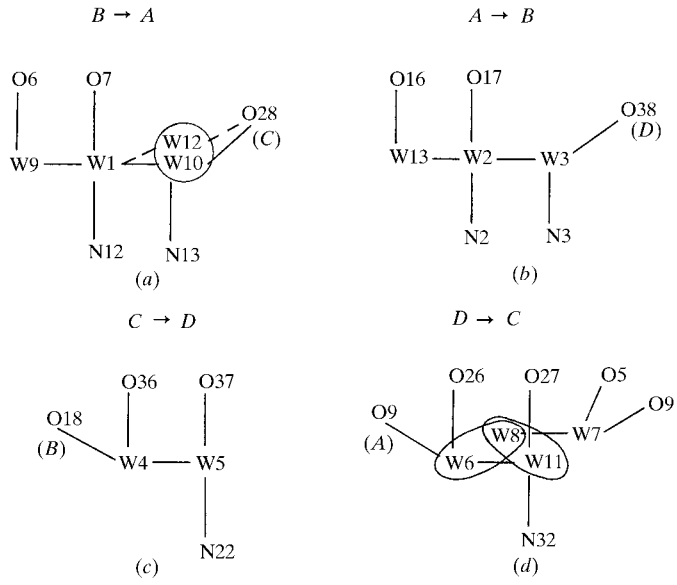

Figure 3
Schematic diagrams of hydrogen bonding involving water molecules in the four head-to-tail regions between $B\rightarrow A$, $A\rightarrow B$, $D\rightarrow C$ and $C\rightarrow D$. The circled water molecules indicate mutually exclusive sites with occupancy of 0.5 or less.

Side chains in *A*, *C* and *D* are oriented similarly. In *B* the Met1 side chain occurs in a $g^+g^+g^+$ conformation that is different than the tg^-g^- conformation[1] found in the other three conformers. Both of the above conformations for the $-CH_2CH_2SCH_3$ chain are also different than the *ttg* conformation for Met found in Boc–Met–Aib–Phe–OMe (Bardi *et al.*, 1986). Additionally, the Leu5 side chain (tg^-t) in *B* differs from the *ggt* conformation that occurs in *A*, *C* and *D*, as well as in Leu2 and Leu8 in all four molecules.

The helices assemble in a head-to-tail fashion to form two distinct sets of infinite columns. Molecules *A* and *B* alternate in one column with $A\rightarrow B\rightarrow A\rightarrow B$ head-to-tail hydrogen bonds: N(1)H···O(17) and N(11)H···O(8) for $A\rightarrow B$ and $B\rightarrow A$, respectively. The other column is formed by

[1] Where g^+ = *gauche*$^+$ = −60° torsion (Table 3); g^- = *gauche*$^-$ = +60° torsion; t = *trans* = 180° torsion.

$C\rightarrow D\rightarrow C\rightarrow D$ with head-to-tail hydrogen bonds: N(21)H$\cdots$O(37) and N(31)H$\cdots$O(27) for $C\rightarrow D$ and $D\rightarrow C$, respectively (Fig. 2). In addition to one direct NH$\cdots$O=C hydrogen bond between peptides in each of the four head-to-tail regions, one or more water molecules act as hydrogen-bond bridges between NH groups at the head of one helix and C=O groups at the bottom of a helix in the same column. Fig. 3 shows the schematic arrangement of hydrogen bonds that involve water molecules. The hydrogen-bonding scheme is quite similar for $B\rightarrow A$ and $A\rightarrow B$. Also, the scheme for $C\rightarrow D$ is similar to part of that for $D\rightarrow C$. Although there are differences in the four schemes, there are also distinct similarities, such as the O—W—W—O hydrogen-bonded loops that involve O(n6) and O(n7) (n = 0–3) for each of the helices. Each cluster of water molecules also participates in one hydrogen bond to a carbonyl O atom in a neighboring column of helices (two hydrogen bonds for W7). Further, NH moieties at the heads of helices, N2, N3, N12, N13, N22 and N32, are hydrogen-bond donors to water molecules that form bridges to carbonyl O atoms at the tail of helices, O6, O7, O16, O17, O26, O27, O36 and O37. The only water molecules in the cell occur in the head-to-tail regions.

The two distinct columns of helices are extended parallel to the c axis of the unit cell. A view from the top of the columns (Fig. 4) shows the oval-shaped projections of the columns packed with a modified herringbone motif. The columns are oriented so that the methionyl side chains face each other.

5.2. Structure of the selenium analog

The structure of the four conformers of the isomorphous Se analog is almost identical in every detail to that found for the four S conformers. Values of the torsional angles are within 2–3° of each other. Values for the comparable NH$\cdots$O=C hydrogen bonds are within 0.03 Å of each other. The minor structural deviations from ideal helices in the sulfur-containing molecules are repeated in the selenium-containing molecules. The only difference occurs in the water content between molecules $D\rightarrow C$ in the selenium-containing crystal in which W6 has almost disappeared.

The Met or SeMet residues at the N-terminus of each helix have not perturbed the folding of the helices in any significant manner. The helices have the expected α-helix type folding with some admixture of 3_{10}-helix that is usually found in Aib-

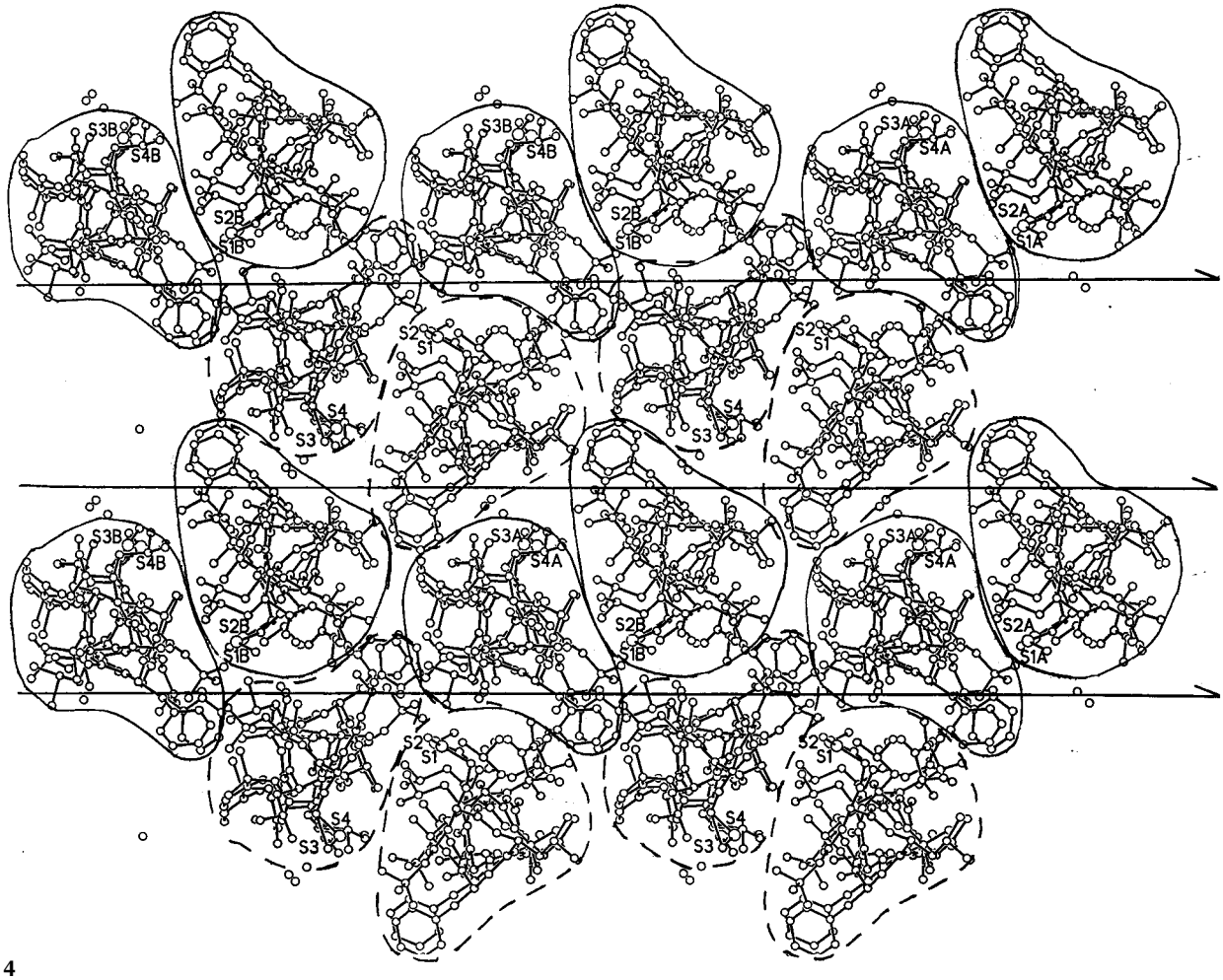

Figure 4
Packing of the two alternating columns ($A\rightarrow B$, S1 and S2; $C\rightarrow D$, S3 and S4) viewed down the c axis. The horizontal lines indicate the location of the twofold screw axes (parallel to the b axis).

containing peptides (Karle & Balaram, 1990). Optimal packing between hydrophobic surfaces may account for the existence of four molecules in an asymmetric unit.

APPENDIX *A*
Definitions of α, β and γ

With the scattering factor for the S atoms expressed as

$$f_h^{\rm S} = f_h^{A\rm S} + f_h^{'\rm S} + if_h^{''\rm S},$$

the α, β and γ terms, functions of the scattering factors, are defined as:

$$\alpha_h^{\rm S} = (r_h^{\rm S})^2,$$

$$\beta_h^{\rm S} = 2r_h^{\rm S}\cos\delta_{\rm S},$$

$$\gamma_h^{\rm S} = 2r_h^{\rm S}\sin\delta_{\rm S},$$

$$r_h^{\rm S} = f_h^{\rm S}/f_h^{A\rm S},$$

$$f_h^{A\rm S} = [(f_h^{\rm S} + f_h^{'\rm S})^2 + (f_h^{''\rm S})^2]^{1/2},$$

$$\delta_{\rm S} = \tan^{-1}[f_h^{''\rm S}/(f_h^{\rm S} + f_h^{'\rm S})].$$

The corresponding relationships for the selenium-containing compound are:

$$f_h^{\rm Se} = f_h^{A\rm Se} + f_h^{'\rm Se} + if_h^{''\rm Se},$$

$$\alpha_h^{\rm Se} = (r_h^{\rm Se})^2,$$

$$\beta_h^{\rm Se} = 2r_h^{\rm Se}\cos\delta_{\rm Se},$$

$$\gamma_h^{\rm Se} = 2r^{\rm Se}\sin\delta_h^{\rm Se},$$

$$r_h^{\rm Se} = f_h^{\rm Se}/f_h^{A\rm S},$$

$$f_h^{A\rm Se} = [(f_h^{\rm Se} + f_h^{'\rm Se})^2 + (f_h^{''\rm Se})^2]^{1/2},$$

$$\delta_{\rm Se} = \tan^{-1}[f_h^{''\rm Se}/(f_h^{\rm Se} + f_h^{'\rm Se})].$$

This work was supported in part by the Office of Naval Research and in part by the National Institutes of Health Grant GM-30902.

References

Bardi, R., Piazzesi, A. M., Toniolo, C., Antony Raj, P., Raghothama, S. & Balaram, P. (1986). *Int. J. Pept. Protein Res.* **27**, 229–238.
Becker, E. L. (1987). *Am. J. Pathol.* **129**, 16–24.
Becker, E. L., Bleich, H. E., Day, A. R., Freer, R. J., Glasel, J. A. & Visintainer, J. (1979). *Biochemistry*, **18**, 4656–4668.
Freer, R. J., Day, A. R., Muthukumaraswamy, N., Pinon, E., Wu, A., Showell, H. J. & Becker, E. L. (1982). *Biochemistry*, **21**, 257–263.
Gavuzzo, E., Mazza, F., Pochetti, G. & Scatturin, A. (1989). *Int. J. Pept. Protein Res.* **34**, 409–415.
IUPAC–IUB Commission on Biochemical Nomenclature (1970). *Biochemistry*, **9**, 3471–3479.
Karle, I. L. & Balaram, P. (1990). *Biochemistry*, **29**, 6747–6756.
Karle, I. L., Flippen-Anderson, J. L., Uma, K. & Balaram, P. (1989). *Biopolymers*, **28**, 773–781.
Karle, J. (1980). *Int. J. Quant. Chem. Symp.* **7**, 357–367.
Karle, J. (1989*a*). *Phys. Today*, **42**, 22–29.
Karle, J. (1989*b*). *Acta Cryst.* A**45**, 303–307.
McRee, D. E. (1993). *Practical Protein Crystallography*, p. 105. New York: Academic Press.
Michel, A. G., Lajoie, G. & Hassani, C. A. (1990). *Int. J. Pept. Protein Res.* **36**, 489–498.
Morffew, A. J. & Tickle, I. (1981). *Cryst. Struct. Commun.* **10**, 781–788.
Prasad, S., Rao, R. B., Bergstrand, H., Lundquist, B., Becker, E. L. & Balaram, P. (1996). *Int. J. Pept. Protein Res.* **48**, 312–318.
Schiffmann, E., Corcoran, N. A. & Wahl, S. M. (1975). *Proc. Natl Acad. Sci. USA*, **72**, 1059–1062.
Showell, H. J., Freer, R. J., Zigmond, S. H., Schiffmann, E., Aswanikumar, S., Corcoran, B. & Becker, E. L. (1976). *J. Exp. Med.* **143**, 1154–1169.
Uma, K. (1990). *J. Ind. Inst. Sci.* **71**, 395–398.