

# Rotational Isomeric State Treatment of the Cystine Residue. Configuration Partition Function and Its Relationship to the Optical Activity Exhibited by the Disulfide Bond

Wayne L. Mattice

Contribution from the Department of Biochemistry, Louisiana State University, Baton Rouge, Louisiana 70803. Received July 14, 1976

**Abstract:** The configuration partition function for two polypeptide chains cross-linked via a single cystinyl residue has been formulated. This objective is achieved using rotational isomeric state theory, in the form appropriate for branched molecules, to combine the configuration partition function for the cystinyl residue with appropriate representations for the configuration partition function of uncross-linked polypeptide chains. Rotational states, as well as preliminary estimates of the corresponding conformational energies, are obtained from consideration of semiempirical energy computations. The sequences -Cys-X- and -Cys-Pro- (X  $\neq$  Pro) lead to different results. The a priori probabilities computed for the two rotational states of the disulfide bond are found to be in excellent agreement with the sign, magnitude, and temperature dependence of the optical activity exhibited by the lowest energy electronic transition in simple derivatives of L-cystine.

Rotational isomeric state theory<sup>1-4</sup> has been successfully applied to homopolypeptides,<sup>5-15</sup> oligopeptides,<sup>16</sup> and both random<sup>17-19</sup> and sequential<sup>20</sup> copolypeptides. It has also been applied to denatured proteins in which the disulfide bonds have been reduced so that covalent cross-links are absent.<sup>21,22</sup> Recent extensions<sup>23,24</sup> of rotational isomeric state theory to encompass the treatment of branched macromolecules now permit a study of disordered proteins containing an intact interchain disulfide bond. The first requirement is the formulation of the appropriate configuration partition function for the two cysteine residues which form the disulfide bond. The configuration partition function is presented in this paper. Subsequent communications will present the unperturbed dimensions computed for denatured proteins containing an interchain disulfide bond.

**Geometry and Potential Functions.** Figure 1 shows a representation of two polypeptide chains cross-linked via the disulfide bond formed by two cysteine residues. The C-S and S-S bond lengths are assigned the values 1.87 and 2.04 Å, respectively.<sup>25</sup> The C-C-S bond angle was 113°,<sup>26</sup> and the C-S-S bond angle was 107°.<sup>27</sup> All other bond lengths and bond angles were identical with those used by Brant et al.<sup>9</sup> in their treatment of the L-alanyl residue. Virtual bonds<sup>5</sup> are shown between successive  $\alpha$ -carbon atoms. Each  $C^\alpha$ ,  $C'$ ,  $C^\beta$ , N, and S atom, as well as the intervening bonds and virtual bonds, is assigned a pre- and postsubscript according to the scheme described in ref 23. The presubscript denotes the branch to which that atom, bond, or virtual bond is assigned. The atoms in a given branch are numbered sequentially, as denoted by the postsubscript. The bond represented by bond vector  ${}_3\mathbf{1}_i$  extends from atom  $i-1$  to atom  $i$  in branch 3, while the virtual bond represented by  ${}_j\mathbf{1}_i$  ( $j = 1, 2, 4, 5$ ) extends from  $\alpha$ -carbon atom  $3i-3$  to  $\alpha$ -carbon  $3i$  in branch  $j$ . The statistical weight matrix for bond (or virtual bond)  $i$  in branch  $j$  is denoted by  ${}_j\mathbf{U}_i$ , and  $n_j$  is the number of virtual bonds in branch  $j$ .

Nonbonded energies were computed using the 6-12 functions described by Brant et al.<sup>9</sup> The radius, polarizability, and effective number of electrons for the sulfur atom were taken from Scheraga.<sup>26</sup> The electrostatic interaction of the peptide groups was calculated in the monopole approximation as described by Brant et al.<sup>9</sup> Torsional potentials for rotation about the N-C $^\alpha$  and C $^\alpha$ -C' bonds were also taken from Brant et al.,<sup>9</sup> while the torsional potential for the C $^\alpha$ -C $^\beta$  bond was that of Abe et al.<sup>28</sup> The torsional potential for the C $^\beta$ -S bond was threefold, with minima at 0 and  $\pm 120^\circ$  and a barrier height of 1.5 kcal/mol.<sup>29</sup> (We adopt here the convention that a trans

placement corresponds to a dihedral angle of 0°.) For the S-S bond, the torsional potential is that used by Allinger et al.<sup>30</sup> Peptide bonds are maintained in the planar trans conformation.

**Statistical Weight Matrices.** The independence of the rotations about the N-C $^\alpha$  and C $^\alpha$ -C' bonds at adjacent  $\alpha$ -carbon atoms<sup>5</sup> permits immediate assignment of statistical weight matrices for virtual bonds which do not involve the  $\alpha$ -carbon atoms in the cysteine residues. A  $1 \times 1$  matrix whose sole element is unity is denoted by eq 1.

$${}_1\mathbf{U}_i = {}_2\mathbf{U}_j = {}_4\mathbf{U}_j = {}_5\mathbf{U}_j = [1] \quad i < n_1, j > 1 \quad (1)$$

Atoms involved in the interactions pertinent to the assembly of  ${}_1\mathbf{U}_{n_1}$  are shown in Figure 2. The statistical weights for the rotational states about the  ${}_1C^\alpha{}_{3n_1-3}C^\beta{}_1$  bond will depend on the values of  $\varphi$  and  $\psi$ . Consequently it is necessary for  ${}_1\mathbf{U}_{n_1}$  to be a row with the number of elements identical with the number of states for  $\varphi$  and  $\psi$ . Evaluation of the number of states, and the statistical weights, is based on the conformational energies for the fragment in Figure 2, evaluated at  $10^\circ$  intervals for  $\varphi$  and  $\psi$ . The result is shown in Figure 3. The convention adopted is that  $\varphi, \psi = 0^\circ, 0^\circ$  for the fully extended chain. Figure 3 is, of course, identical with the conformational energy map obtained for the L-alanyl residue by Brant et al.<sup>9</sup> The minimum energy is obtained when  $\varphi, \psi$  are  $100^\circ, 330^\circ$ .<sup>9</sup> This minimum and six local minima are denoted by the X's in Figure 3. They are listed in Table I in order of increasing energy.

A preliminary evaluation was made of the feasibility of using seven states for rotation about the last virtual bond in branch 1. Statistical weights for the states were assigned as shown in eq 2.

$$\frac{\sum_{\varphi} \sum_{\psi} \exp[-E(\varphi, \psi)/RT]}{\sum_{\varphi} \sum_{\psi} \exp[-E(\varphi, \psi)/RT]} \quad (2)$$

The summations are taken at  $10^\circ$  intervals for  $\varphi$  and  $\psi$ , ignoring locations where the energy exceeds the minimum by more than 5 kcal/mol. The range for the double sum in the numerator is over that section of the conformational energy map assigned to the state in question, while the range for the double sum in the denominator extends over the entire conformational energy map. The boundaries for the states are provided by the 5 kcal/mol contour line and the dashed lines in Figure 3. The dashed lines approximate the location of the

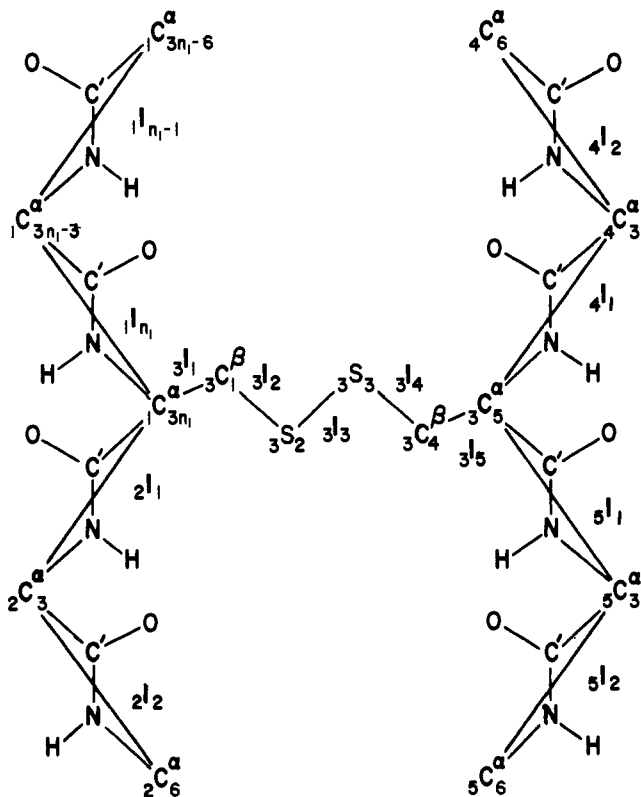


Figure 1. Representation of two polypeptide chains cross-linked by the disulfide bond formed by two cysteine residues.

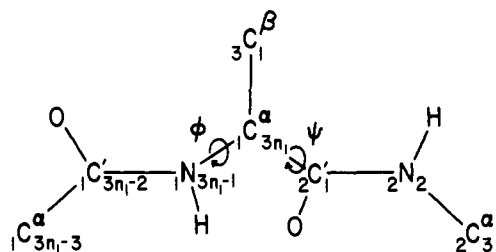


Figure 2. Atoms considered in the formulations of  ${}_1U_{n_1}$ .

energy barriers between the minima. Average values of  $\varphi$  and  $\psi$  were determined for each state, using eq 3 and 4. The summations extend

$$\langle \varphi \rangle = \frac{\sum_{\varphi} \sum_{\psi} \varphi \exp[-E(\varphi, \psi)/RT]}{\sum_{\varphi} \sum_{\psi} \exp[-E(\varphi, \psi)/RT]} \quad (3)$$

$$\langle \psi \rangle = \frac{\sum_{\varphi} \sum_{\psi} \psi \exp[-E(\varphi, \psi)/RT]}{\sum_{\varphi} \sum_{\psi} \exp[-E(\varphi, \psi)/RT]} \quad (4)$$

over the region of the conformational energy map assigned to the state in question. Results are summarized in Table I for the case where  $RT = 0.6$  kcal/mol.

The final two states listed in Table I were eliminated from further consideration because of their small combined weight. An additional simplification is provided by combination of the fourth and fifth entries to yield a state with  $\langle \varphi \rangle$ ,  $\langle \psi \rangle = 86^\circ$ ,  $126^\circ$  and a weight of 0.043. This procedure yields the four states denoted by the squares in Figure 3. The statistical weight matrix  ${}_1U_{n_1}$  is shown in eq 5.

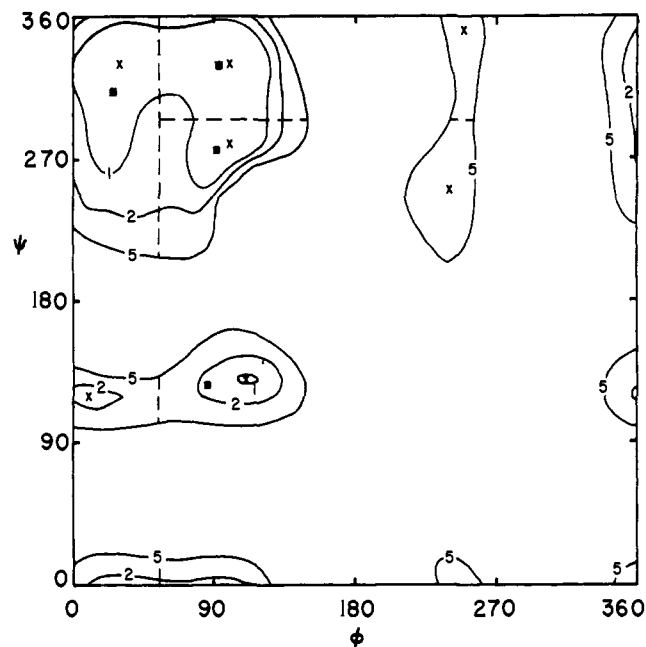


Figure 3. Conformational energy map obtained for the atoms in Figure 2. Contours are drawn at 1, 2, and 5 kcal/mol above the energy minimum, which occurs at  $\varphi, \psi = 100^\circ, 330^\circ$ . Minima are denoted by X's. The squares represent the rotational states used (see text).

Table I. Minima in the L-Alanyl Residue Conformational Energy Map

Energy, kcal/mol	$\varphi$ , deg	$\psi$ , deg	$\langle \varphi \rangle$ , deg	$\langle \psi \rangle$ , deg	Statistical wt
0.00	100	330	94	328	0.415
0.21	100	280	92	275	0.173
0.27	30	330	26	312	0.363
0.91	110	130	107	128	0.033
1.33	10	120	15	119	0.010
2.34	230	240	236	248	0.004
2.38	250	350	247	348	0.002

Table II. States Used for  $\varphi, \psi$

Designation	$\langle \varphi \rangle$ , deg	$\langle \psi \rangle$ , deg	Statistical wt	${}_1U_{n_1}$
$\beta_1$	94	328	$w_1$ (0.418)	0.426
$\beta_2$	26	312	$w_2$ (0.365)	0.380
$\beta_3$	92	275	$w_3$ (0.174)	0.154
$\alpha_1$	86	126	$w_4$ (0.043)	0.040

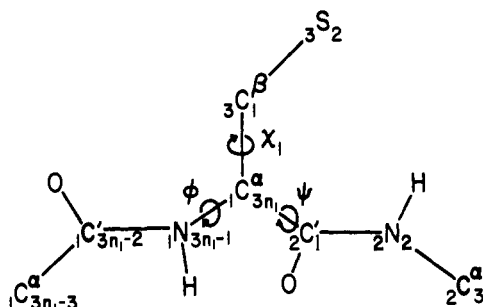
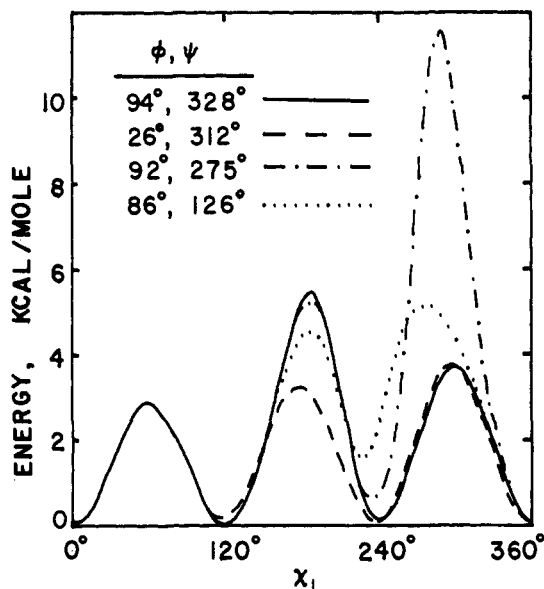
$${}_1U_{n_1} = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \alpha_1 \\ w_1 & w_2 & w_3 & w_4 \end{bmatrix} \quad (5)$$

Table II provides a summary in which the  $w_i$  have been adjusted so that their sum is unity.

The dimensions of  ${}_1U_{n_1}$  and  ${}_2U_2$  require that  ${}_2U_1$  be a column consisting of four elements. Independence of events at neighboring  $\alpha$ -carbon atoms requires that these four elements be identical.

$${}_2U_1 = \text{col}(1, 1, 1, 1) \quad (6)$$

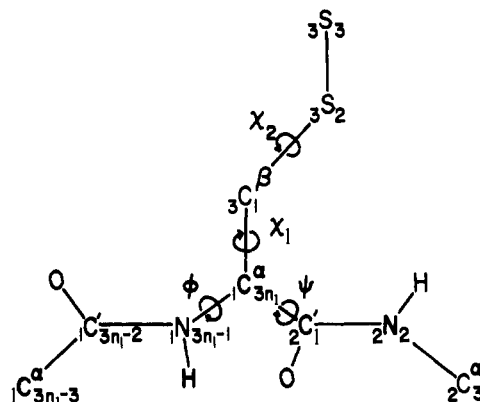
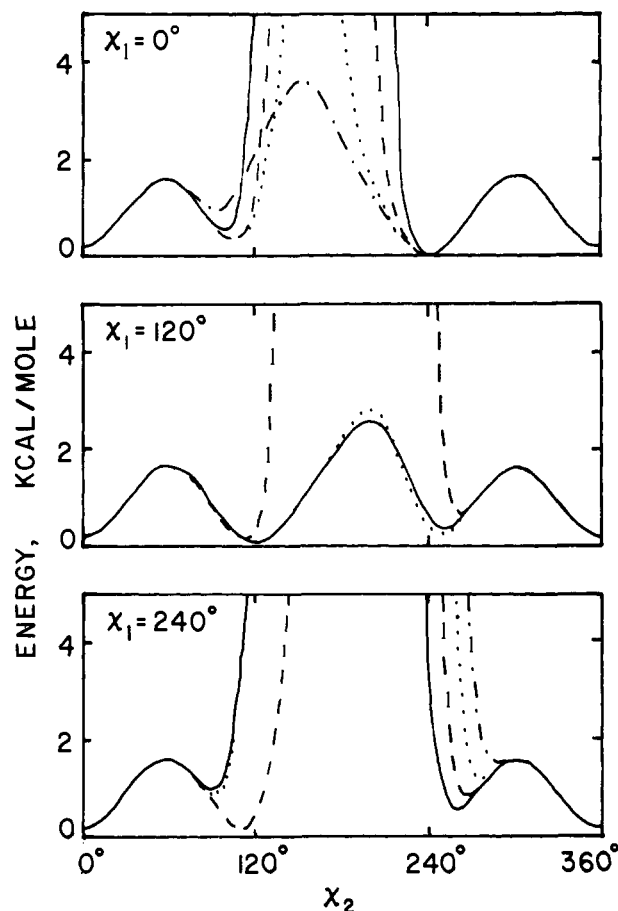
The dihedral angle,  $\chi_1$ , about the  ${}_1C^{\alpha}_{3n_1}-C^{\beta}_1$  bond specifies the position of  ${}_3S_2$ . This angle will be considered to be  $0^\circ$  for a trans placement of  ${}_1N_{3n_1-1}-{}_1C^{\alpha}_{3n_1-3}-C^{\beta}_1-{}_3S_2$ . Interactions contributing to the statistical weights represented in  ${}_3U_1$  are the torsional potential about the  ${}_1C^{\alpha}_{3n_1-3}-C^{\beta}_1$  bond and the

Figure 4. Atoms considered in the formulation of  ${}_3U_1$ .Figure 5. Conformational energy as a function of  $\chi_1$  for the four indicated combinations of  $\phi, \psi$ .

nonbonded interactions between  ${}_3S_2$  and all other atoms in Figure 4 except  ${}_1C^{\alpha}_{3n1}$  and  ${}_3C^{\beta}_1$ . The resulting energies are shown in Figure 5 for the four states assigned by  $\phi, \psi$ . This energy is essentially identical for all four combinations of  $\phi, \psi$  when  $\chi_1 = 0^\circ$ . When  $\chi_1 = 120^\circ$  and  $\phi$  is  $26^\circ$ , the  ${}_3S_2$  atom is 2.96 Å from the oxygen atom bonded to  ${}_1C^{\alpha}_{3n1-2}$ . However, the energy of this interaction is less than 0.1 kcal/mol and can be alleviated by a slight increase in  $\phi$ , so it will be ignored. The  ${}_3S_2$  atom interacts with the oxygen atom bonded to  ${}_2C^{\alpha}_1$  when  $\chi_1 = 240^\circ$  and  $\psi = 275^\circ$ . The estimated energy is about 0.6 kcal/mol, and the statistical weight is  $s_1$ . This interaction is relieved at larger values of  $\psi$  and is practically nonexistent when  $\psi$  is in the  $\beta_1$  or  $\beta_2$  states. A repulsive interaction occurs between  ${}_2S_2$  and  ${}_2N_2$ , plus its attached hydrogen atom, when  $\chi_1 = 240^\circ$  and  $\psi = 126^\circ$ . The S-H distance is only 2.23 Å. The energy of this interaction is about 2 kcal/mol, and the statistical weight is  $s_2$ . Weak three-bond interactions, represented by  $s_3$ , occur between  ${}_3S_2$  and both  ${}_1N_{3n1-1}$  and  ${}_2C^{\alpha}_1$  when  $\chi_1 = 240^\circ$ . The energy is 0.2 kcal/mol. The resulting statistical weight matrix is shown in eq 7. The dimensions of this matrix

$${}_3U_1 = \begin{matrix} & \beta_1 & \beta_2 & \beta_3 & \alpha_1 \\ & t & g^+ & g^- & t & g^+ & g^- & t & g^+ & g^- \\ \beta_1 & \begin{bmatrix} 1 & 1 & s_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & s_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & s_1 s_3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} & \alpha_1 & \begin{bmatrix} 1 & 1 & s_2 s_3 \end{bmatrix} \end{matrix} \quad (7)$$

are  $4 \times 12$ , rather than  $4 \times 3$ , because we wish to be able to recognize the dependence of rotation about the  ${}_3C^{\beta}_1-{}_3S_2$  bond on the state adopted by  $\phi, \psi$ . The rectangular matrix denoted

Figure 6. Atoms considered in the formulation of  ${}_3U_2$ .Figure 7. Conformational energy as a function of  $\chi_2$  for the three values of  $\chi_1$  and four values of  $\phi, \psi$ . The  $\phi, \psi$  are given by the same type as lines used in Figure 5.

${}_2U_1 \ominus {}_3U_1$ , and defined in ref 23, is identical with  ${}_3U_1$ .

Specification of the dihedral angle,  $\chi_2$ , about the  ${}_3C^{\beta}_1-{}_3S_2$  bond determines the position of  ${}_3S_3$ . Interactions contributing to the statistical weights embodied in  ${}_3U_2$  are the torsional potential about the  ${}_3C^{\beta}_1-{}_3S_2$  bond and the nonbonded interaction between  ${}_3S_3$  and all other atoms in Figure 6 except  ${}_3C^{\beta}_1$  and  ${}_3S_2$ . The resulting energies are shown in Figure 7 for the four states assigned to  $\phi, \psi$  and the three states for  $\chi_1$ . The energy in the vicinity of  $\chi_2 = 0^\circ$  is seen to be essentially independent of  $\chi_1$  and the state assigned by  $\phi, \psi$ .

The three-bond interaction between  ${}_3S_3$  and  ${}_1C^{\alpha}_{3n1}$  is slightly attractive (0 to -0.1 kcal/mol) when  $\chi_2$  is in either of the gauche states. This interaction is so small that it will be ignored. Repulsive interactions involving  ${}_3S_3$  and atoms in the amide units may occur when  $\chi_2 = \pm 120^\circ$ . These repulsions

cause the minima for the gauche states to frequently be displaced from  $\pm 120^\circ$ .

Interaction occurs between  $3S_3$  and  $2C'_1$  (separation 3.04 Å) when  $\chi_1, \chi_2 = 0^\circ, 110^\circ$ . An additional interaction occurs between  $3S_3$  and the oxygen atom attached to  $2C'_1$  (separation 2.64 Å) when  $\psi = 275^\circ$ . The  $3S_3$  and  $2N_2$  atoms are separated by only 2.75 Å when  $\psi = 328^\circ$ . These latter two interactions can be relieved somewhat by an appropriate adjustment of  $\psi$ . The relevant statistical weights are denoted by  $s_4s_5, s_4, s_4s_6$ , and  $s_4$  for the cases where  $\varphi, \psi$  are in the  $\beta_1, \beta_2, \beta_3$ , and  $\alpha_1$  states, respectively. The energies corresponding to  $s_4-s_6$  are estimated at 0.2, 0.7, and 1.0 kcal/mol, respectively.

No repulsive interactions occur when  $\chi_1, \chi_2 = 0^\circ, 250^\circ$ . The energy in this conformation may actually be negative by as much as 0.2 kcal/mol due to favorable nonbonded interactions. The statistical weight is denoted by  $s_7$ . No significant interactions occur when  $\chi_1, \chi_2 = 120^\circ, 110^\circ$  unless  $\varphi, \psi$  are in the  $\beta_2$  state. A slightly higher energy (up to 0.2 kcal/mol) may be appropriate for  $\beta_2g^+g^+$  due to the interaction between  $3S_3$  and the oxygen atom attached to  $1C'_{3n_1-2}$ . The statistical weight is  $s_8$ .

A weak repulsive interaction occurs between  $3S_3$  and the hydrogen atom attached to  $1N_{3n_1-1}$  when  $\chi_1 = 120^\circ, \chi_2 = 250^\circ$ , and  $\varphi = 90 \pm 4^\circ$ . This interaction is alleviated by a slight decrease in  $\varphi$ . A stronger interaction occurs between  $3S_3$  and  $1C'_{3n_1-2}$  (separation 2.70 Å) when  $\chi_1 = 120^\circ, \chi_2 = 250^\circ$ , and  $\varphi = 26^\circ$ . The statistical weights are denoted by  $s_9$  and  $s_{10}$  and the corresponding energies are estimated to be 0.2 and 1 kcal/mol, respectively.

Severe interaction occurs between  $3S_3$  and the hydrogen atom attached to  $1N_{3n_1-1}$  (separation 2.13–2.14 Å) when  $\varphi = 90 \pm 4^\circ, \chi_1 = 240^\circ$ , and  $\chi_2 = 110^\circ$ . This interaction is reflected in  $s_{11}$ , with an energy of 2–3 kcal/mol.

Repulsive interactions exist whenever  $\chi_1, \chi_2$  are in the  $g^-g^-$  state. Interaction occurs between  $3S_3$  and the oxygen atom attached to  $2C'_1$  when  $\psi = 300 \pm 28^\circ$ , with the interaction becoming of increasing severity as  $\psi$  decreases through this range. The statistical weights are denoted by  $s_{12}-s_{14}$  with estimated energies of 0.5, 1–2, and  $>5$  kcal/mol. The interactions are between  $3S_3$  and both  $2N_2$  and its hydrogen atom when  $\psi$  is  $126^\circ$ . The estimated energy is  $>3$  kcal/mol. The statistical weight matrix  $3U_2$  is shown in eq 8.

$$3U_2 = \begin{matrix} & & t & g^+ & g^- \\ \beta_1 & t & 1 & s_4s_5 & s_7 \\ & g^+ & 1 & 1 & s_9 \\ & g^- & 1 & s_{11} & s_{12} \\ \beta_2 & t & 1 & s_4 & s_7 \\ & g^+ & 1 & s_8 & s_{10} \\ & g^- & 1 & 1 & s_{13} \\ \beta_3 & t & 1 & s_4s_6 & s_7 \\ & g^+ & 1 & 1 & s_9 \\ & g^- & 1 & s_{11} & s_{14} \\ \alpha_1 & t & 1 & s_4 & s_7 \\ & g^+ & 1 & 1 & s_9 \\ & g^- & 1 & s_{11} & s_{15} \end{matrix} \quad (8)$$

Specification of the dihedral angle,  $\chi_3$ , about the  $3S_2-3S_3$  bond determines the position of  $3C^{\beta_4}$ . Interactions contributing to the statistical weights embodied in  $3U_3$  are the torsional potential about the  $3S_2-3S_3$  bond and the nonbonded interactions between  $3C^{\beta_4}$  and  $3C^{\beta_1}$  and between  $3C^{\beta_4}$  and  $1C^{\alpha_{3n_1}}$ . The resulting energies are shown in Figure 8 for the three states of

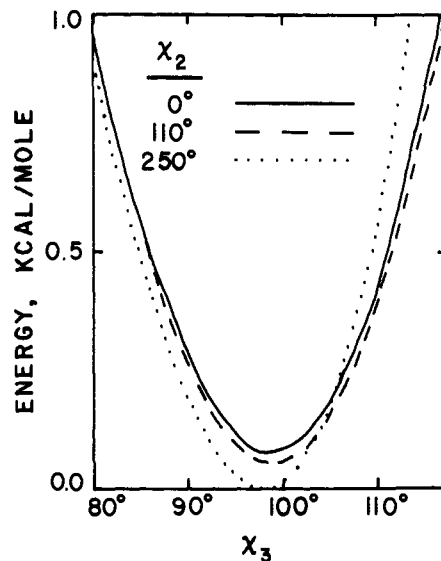


Figure 8. Conformational energies as a function of  $\chi_3$  for the indicated values of  $\chi_2$ .

Table III. Energies Associated with the  $s_i$

Statistical wt	Estd energy, kcal/mol	Statistical wt	Estd energy, kcal/mol
$s_1$	0.6	$s_9$	0.2
$s_2$	2	$s_{10}$	1
$s_3$	0.2	$s_{11}$	2
$s_4$	0.2	$s_{12}$	0.5
$s_5$	0.7	$s_{13}$	1.5
$s_6$	1	$s_{14}$	5
$s_7$	-0.2	$s_{15}$	3
$s_8$	0.2	$s_{16}$	-0.1

the  $3C^{\beta_1}-3S_2$  bond. Essentially the same result is found when the gauche states for  $\chi_2$  are located at  $\pm 110^\circ$  or  $\pm 120^\circ$ . Minima occur near  $\pm 100^\circ$  for  $\chi_3$  in Figure 8. The lowest energies are obtained when  $\chi_2$  and  $\chi_3$  adopt gauche placements of opposite signs, due to favorable nonbonded interactions generated between the  $3C^{\beta_4}$  and  $1C^{\alpha_{3n_1}}$  atoms in this conformation. The energy associated with  $s_{16}$  in  $3U_3$  is about -0.1 kcal/mol. The energy in the  $g^+g^+$  and  $g^-g^-$  states is calculated to be only about 0.02 kcal/mol lower than that in the  $tg^\pm$  states, and consequently it will be assigned a statistical weight of unity.

$$3U_3 = \begin{matrix} & t & \begin{bmatrix} g^+ & g^- \\ 1 & 1 \\ 1 & s_{16} \\ s_{16} & 1 \end{bmatrix} \\ g^+ & & \\ g^- & & \end{matrix} \quad (9)$$

The interactions embodied in  $3U_4$  and  $3U_5$  have already been discussed in connection with preceding statistical weight matrices. The resulting matrices are shown in eq 10 and 11.

$$3U_4 = \begin{matrix} & t & g^+ & g^- \\ g^+ & 1 & 1 & s_{16} \\ g^- & 1 & s_{16} & 1 \end{matrix} \quad (10)$$

$$3U_5 = \begin{matrix} & t & \beta_1 & & \beta_2 & & \beta_3 & & \alpha_1 \\ & & g^+ & g^- & g^+ & g^- & g^+ & g^- & g^+ & g^- \\ t & w_1 & w_1 & s_3w_1 & w_2 & w_2 & s_3w_2 & w_3 & w_3 & s_1s_3w_3 & w_4 & w_4 & s_2s_3w_4 \\ g^+ & s_4s_5w_1 & w_1 & s_3s_{11}w_1 & s_4w_2 & s_8w_2 & s_3w_2 & s_4s_6w_3 & w_3 & s_1s_3s_{11}w_3 & s_4w_4 & w_4 & s_2s_3s_{11}w_4 \\ g^- & s_7w_1 & s_9w_1 & s_3s_{12}w_1 & s_7w_2 & s_{10}w_2 & s_3s_{13}w_2 & s_7w_3 & s_9w_3 & s_1s_3s_{14}w_3 & s_7w_4 & s_9w_4 & s_2s_3s_{15}w_4 \end{matrix} \quad (11)$$

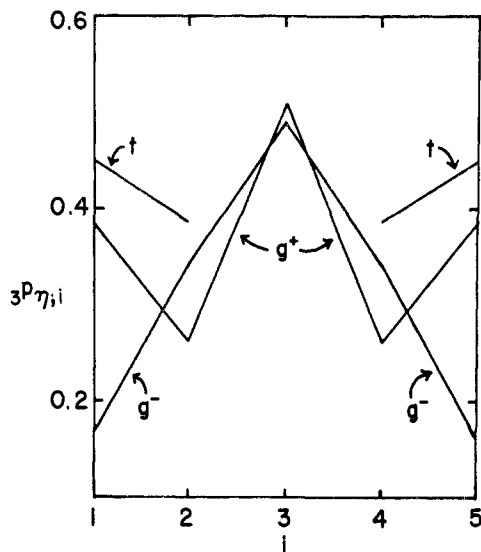


Figure 9. A priori probabilities for the bonds in branch 3 when  $RT = 0.6$  kcal/mol.

A summary of the preliminary estimates of the energies associated with the various  $s_i$  is presented in Table III. The dimensions of  ${}_3U_5$ ,  ${}_4U_2$ , and  ${}_5U_2$  require that  ${}_4U_1$  and  ${}_5U_1$  each be a column consisting of 12 elements. Independence of events at adjacent  $\alpha$ -carbon atoms<sup>5</sup> requires that all of the elements be identical.

$${}_4U_1 = {}_5U_1 = {}_4U_1 \otimes {}_5U_1 \\ = \text{col}(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1) \quad (12)$$

**Configuration Partition Function and average Bond Conformations.** The configuration partition function,  $Z$ , for the molecule in Figure 1 is obtained from the statistical weight matrices via eq 13.

$$Z = {}_1U_1^{(n_1)}({}_2U_1 \otimes {}_3U_1)[({}_2U_2^{(n_2-1)}) \\ \otimes [{}_3U_2^{(4)}({}_4U_1 \otimes {}_5U_1)[({}_4U_2^{(n_4-1)}) \otimes ({}_5U_2^{(n_5-1)})]]] \quad (13)$$

Symbolism of the type  ${}_1U_1^{(n_1)}$  denotes the product of  $n_1$  successive matrices, commencing with  ${}_1U_1$ , and  $\otimes$  denotes the direct product.<sup>4</sup> The a priori probability that bond  $i$  in branch  $j$  is in state  $\eta$ , denoted  ${}_jP_{\eta,i}$ , and the a priori probability that this bond is in state  $\eta$  and the preceding bond is in state  $\xi$ , denoted  ${}_jP_{\xi\eta,i}$ , were calculated as  $Z'/Z$ .<sup>23</sup> The procedure for calculating  $Z'$  is identical with that used for  $Z$  except that appropriate elements of one of the statistical weight matrices are replaced by zeros.<sup>3,23</sup> Table II shows the  ${}_jP_{\eta,n_j}$  calculated using the energies in Table III, the  $w_i$  in Table II, and  $RT = 0.6$  kcal/mol. They differ only slightly from the  $w_i$ . This result indicates that atoms beyond  ${}_3C\beta_1$  have only a minor effect upon the conformational preferences for  $\varphi$ ,  $\psi$ , a result in harmony with the experimental unperturbed dimensions of several homopolypeptides containing  $-\text{CH}_2\text{R}$  side chains.<sup>5-7,12</sup>

A priori probabilities for the bonds in branch 3 are shown in Figure 9. Symmetry dictates that  ${}_3P_{\eta,i} = {}_3P_{\eta,6-i}$ . All bonds except the disulfide bond favor the trans conformation. The a priori probability that the disulfide bond will be in the  $g^+$  state is slightly greater than the a priori probability that it will be in the  $g^-$  state.

It is of interest to compare the results obtained from the present study with those predicted<sup>31,32</sup> from theoretical studies of the electronic transitions in the disulfide group. The wavelength of the lowest energy electronic transition is predicted, and observed, to be extremely sensitive to the dihedral angle about the disulfide bond.<sup>31</sup> If  $\chi_3$  is  $\pm 100^\circ$ , as suggested by Figure 8, this band is predicted to be located at 260–265 nm.<sup>31</sup>

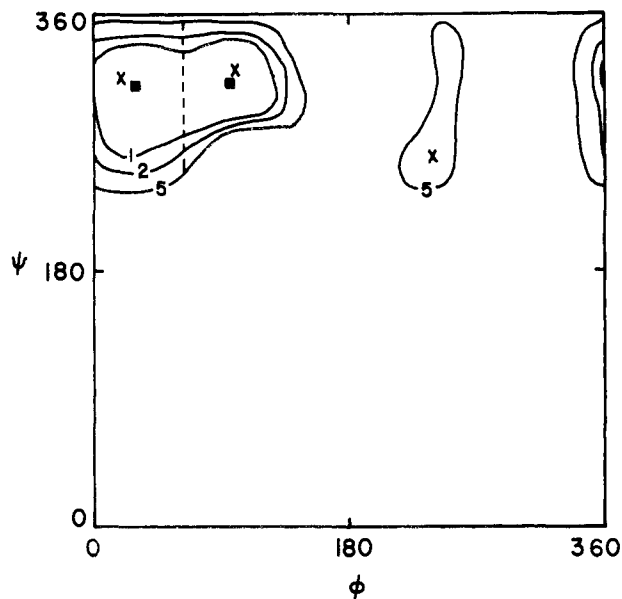


Figure 10. Conformational energy map for L-cysteine followed by L-proline. Contours are drawn at 1, 2, and 5 kcal/mol above the energy minimum, which occurs at  $\varphi, \psi = 100^\circ, 320^\circ$ . Minima are denoted by 'x's. The squares represent the rotational states used (see text).

Coleman and Blout<sup>33</sup> find that the lowest energy disulfide transition is at  $262 \pm 3$  nm for  $N,N'$ -diacetyl-L-cystine bis(methylamide) in aqueous solution.

The lowest energy electronic transition is predicted<sup>31,32</sup> to have negative rotatory strength when  $\chi_3 = 100^\circ$  and positive rotatory strength when  $\chi_3 = -100^\circ$ . Experimentally, this band exhibits negative circular dichroism for solutions of L-cystine,<sup>33</sup>  $N,N'$ -diacetyl-L-cystine,<sup>33</sup>  $N,N'$ -diacetyl-L-cystine bis(methylamide),<sup>33</sup> oxidized glutathione,<sup>33,34</sup>  $S$ -ethylthio-L-cysteine,<sup>34</sup>  $S$ -ethylthio-L-cysteine,<sup>34</sup>  $S$ -propylthio-L-cysteine,<sup>34</sup> and  $S$ -methylthiogluthathione.<sup>34</sup> These results, in conjunction with Figure 8, require that  ${}_3P_{g^+,3} > {}_3P_{g^-,3}$ . The a priori probabilities presented in Figure 9 are in agreement with this inequality.

The observed optical activity of the longest wavelength electronic transition in  $S$ -methylthio-L-cysteine,  $S$ -ethylthio-L-cysteine, and  $S$ -propylthio-L-cysteine decreases substantially with increasing temperature.<sup>34</sup> Circular dichroism spectra show that the maximum ellipticity at  $RT = 0.7$  kcal/mol is only 67–71% of the value at  $RT = 0.6$  kcal/mol. The configuration partition function for  $N$ -acetyl- $S$ -methylthio-L-cysteine methylamide can be written as

$$Z = {}_1U_{n_1}({}_2U_1 \otimes {}_3U_1)({}_3U_2)({}_3U_3) \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad (14)$$

The  ${}_3P_{g^\pm,3}$  were calculated using the  $w_i$  and  $s_i$  for  $RT = 0.6$  and  $0.7$  kcal/mol. This procedure indicates that  ${}_3P_{g^+,3} - {}_3P_{g^-,3}$  at  $RT = 0.7$  kcal/mol would be 68% of the value at  $RT = 0.6$  kcal/mol, a result in excellent agreement with the observed temperature dependence of the circular dichroism.

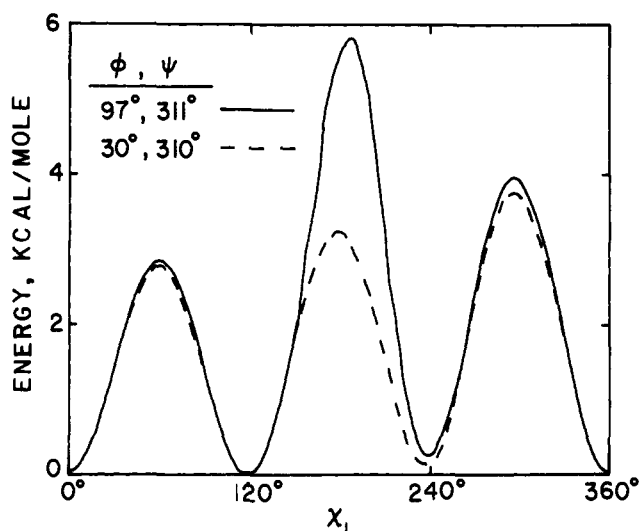
An estimate of  ${}_3P_{g^+,3}$  can be obtained from eq 15,

$$R_{\text{obsd}} = R[{}_3P_{g^+,3} - (1 - {}_3P_{g^+,3})] \quad (15)$$

where  $R_{\text{obsd}}$  is the observed rotatory strength and  $R$  is the rotatory strength predicted for  $\chi_3 = 100^\circ$ . Coleman and Blout<sup>33</sup> find that  $R_{\text{obsd}}$  is  $-1.641 \pm 0.235 \times 10^{-40}$  erg  $\text{cm}^3$  for the 262.3-nm band of  $N,N'$ -diacetyl-L-cystine bis(methylamide). Webb et al.<sup>32</sup> calculate that the net rotatory strength for the two closely situated, lowest energy electronic transitions should be  $-121 \times 10^{-40}$  and  $-193 \times 10^{-40}$  erg  $\text{cm}^3$  when  $\chi_3$  is 90 and  $105^\circ$ , respectively. Linear interpolation provides an estimate of  $-169 \times 10^{-40}$  erg  $\text{cm}^3$  for  $R$  when  $\chi_3$  is  $100^\circ$ . Use of these values for  $R_{\text{obsd}}$  and  $R$  in eq 15 provides a result of  $0.50485 \pm$

Table IV. States Used for  $\phi$ ,  $\psi$  in Cys(Pro)

Designation	Energy, kcal/mol	$\phi$ , deg	$\psi$ , deg.	$\langle\phi\rangle$ , deg	$\langle\psi\rangle$ , deg	Statistical wt
$\beta_1$	0.00	100	320	97	311	0.49 ( $w_1$ )
$\beta_2$	0.14	20	310-320	30	310	0.51 ( $w_2$ )
	2.66	240	260	241	276	

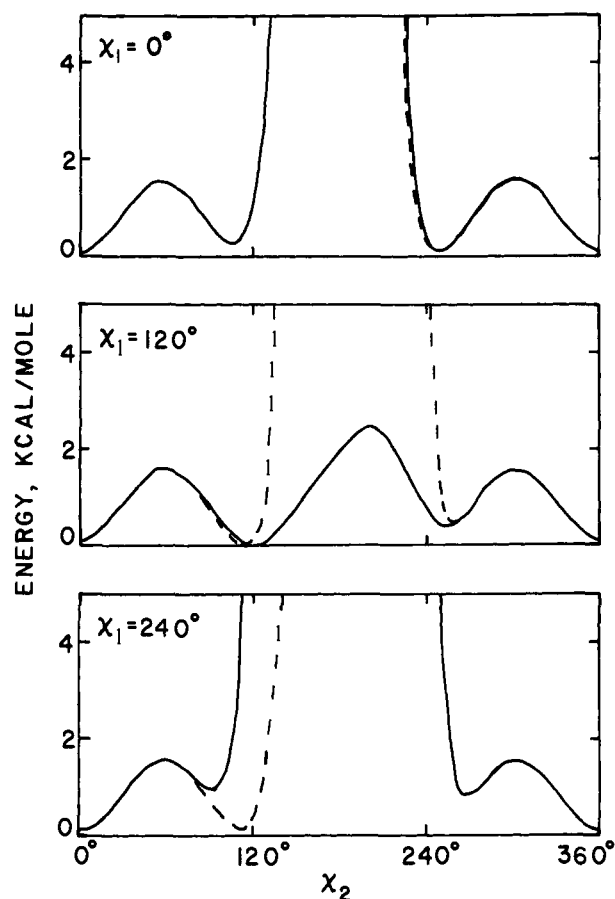
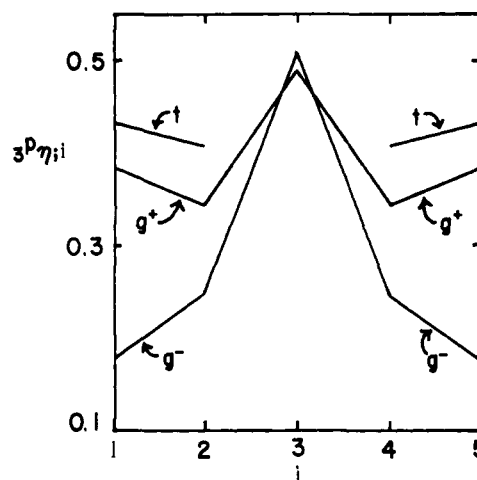
Figure 11. Conformational energy as a function of  $\chi_1$  for L-cystine followed by L-proline.

0.0007 for  $3p_{g^+;3}$ . The value calculated from the rotational isomeric state treatment (0.5068) is in surprisingly good agreement with the estimate from eq 15.

The effect on  $3p_{g^+;3}$  of doubling and halving the energies associated with the  $s_i$  was explored. A larger value of  $3p_{g^+;3}$  is obtained with an increase in the energy associated with  $s_3$ ,  $s_4$ ,  $s_5$ ,  $s_6$ ,  $s_8$ , or  $s_{11}$  or with a decrease in the energy associated with  $s_7$ ,  $s_9$ ,  $s_{10}$ ,  $s_{12}$ ,  $s_{13}$ , or  $s_{16}$ . Little effect is obtained upon variation of the energy associated with  $s_1$ ,  $s_2$ ,  $s_{14}$ , or  $s_{15}$ . A larger  $3p_{g^+;3}$  is also obtained by increasing  $w_1$  or  $w_3$  or by decreasing  $w_2$ . The result is insensitive to  $w_4$ . These variations in the  $w_i$  and the energies associated with the  $s_i$  do not significantly improve the ability to reproduce the observed spectral characteristics of small molecules related to L-cystine.

**Treatment when Proline Follows Cysteine.** The conformational energy map for an L-alanyl residue is altered when it is followed by an L-prolyl residue.<sup>18</sup> For this reason it can be anticipated that a different result will be obtained for the cystine configuration partition function when the amino acid sequence is -Cys-Pro-. The major cause of the change is the steric interactions generated by the  $\delta$  ( $\text{CH}_2$ ) of the prolyl residue. In order to estimate the changes required, the hydrogen atom bonded to  $2N_2$  was replaced by a methylene unit, using a  $\text{C}^\delta\text{-N}$  bond length of 1.47 Å and a  $\text{C}'\text{-N-C}^\delta$  angle of  $126^\circ$ .<sup>13</sup>

Reevaluation of the energies appropriate for consideration in  $1U_{n1}$  yields Figure 10. This figure bears a close resemblance to the conformational energy map computed by Schimmel and Flory<sup>18</sup> for an L-alanyl residue followed by an L-prolyl residue. The minimum energy in Figure 10 is at  $\phi$ ,  $\psi = 100^\circ$ ,  $320^\circ$ , which is close to the location ( $100^\circ$ ,  $300^\circ$ ) found for Ala-Pro.<sup>18</sup> This minimum, and two local minima within 3 kcal/mol of the minimum, is denoted by  $\times$  in Figure 10 and summarized in Table IV. The statistical weights,  $\langle\phi\rangle$ , and  $\langle\psi\rangle$  were obtained by the same procedure applied for  $1U_{n1}$  when the residue following cysteine is not L-prolyl. The  $\beta_1$  and  $\beta_2$  designations bear an obvious correspondence to two of the entries in Table II. No state is assigned to the region at  $\phi = 240^\circ$  because of its low

Figure 12. Conformational energy as a function of  $\chi_2$  for the three values of  $\chi_1$  and two values of  $\phi$ ,  $\psi$  used in Figure 11.Figure 13. A priori probabilities for the bonds in branch 3 when L-proline follows L-cysteine in both polypeptide chains.  $RT = 0.6$  kcal/mol.

statistical weight (0.002). The statistical weight matrix  $1U_{n1}$  is therefore given by eq 5 with  $w_1 = 0.49$ ,  $w_2 = 0.51$ ,  $w_3 = w_4 = 0$ .

The energies appropriate for consideration in  $3U_1$  are shown

Table V. Changes Required for Cys(Pro)

---

$w_1 = 0.49$
$w_2 = 0.51$
$w_3 = w_4 = 0$
$E_{s_5} = E_{s_7} = E_{s_8} = 0.0$ kcal/mol
$E_{s_{12}} = 1.5$ kcal/mol
$\langle \varphi \rangle, \langle \psi \rangle = 97^\circ, 311^\circ$ for $\beta_1$
$\langle \varphi \rangle, \langle \psi \rangle = 30^\circ, 310^\circ$ for $\beta_2$

---

in Figure 11. There is a close correspondence with the analogous two curves in Figure 5. Hence eq 7 can still be used for  $3U_1$ .

Figure 12 presents the energies which are reflected in  $3U_2$ . Comparison with Figure 7 reveals a general similarity, but there are differences in detail. The interaction between  $3S_3$  and  $2N_2$  when  $\chi_1, \chi_2 = 0^\circ, 110^\circ$  is relieved upon a reduction of  $\psi$  from  $328$  to  $311^\circ$ . For this reason the energy associated with  $s_5$  is estimated to be  $0.0$  kcal/mol in Cys(Pro). The energies at  $\chi_1, \chi_2 = 0^\circ, 250^\circ$  and for  $\beta_2 g^+ g^+$  are sufficiently close to that obtained when  $\chi_2 = 0^\circ$  to permit assignment of zero energy for the statistical weights denoted by  $s_7$  and  $s_8$ . Finally, the energy associated with the  $\beta_1 g^- g^-$  state must be changed due to the increased interaction between  $3S_3$  and the oxygen atom bonded to  $2C'_1$  when  $\psi$  decreases from  $328$  to  $311^\circ$ . The energy associated with  $s_{12}$  is increased from  $0.5$  to  $1.5$  kcal/mol. A summary of the changes required is presented in Table V. The energies associated with  $s_1, s_2, s_6, s_{14},$  and  $s_{15}$  are irrelevant since terms containing these statistical weights will be rejected. A priori probabilities for the bonds in branch 3 are shown in Figure 13. The value of  $3p_{g^+,3}$  is  $0.4920$ . It can be made to exceed  $0.5$  by minor ( $0.1$  kcal/mol) adjustment of the energies associated with the  $s_i$ . The adjustments are in the direction which serves to increase  $3p_{g^+,3}$  for the case where the amino acid residue following cysteine is not proline.

**Acknowledgment.** This research was supported by NSF Grant 72-02416 A01.

## References and Notes

- (1) V. M. Volkensteln, "Configurational Statistics of Polymeric Chains", S. N. Timasheff and M. J. Timasheff, Transl., Interscience, New York, N.Y., 1963.
- (2) T. M. Birshtein and O. B. Ptitsyn, "Conformations of Macromolecules", S. N. Timasheff and M. J. Timasheff, Transl., Interscience, New York, N.Y., 1966.
- (3) P. J. Flory, "Statistical Mechanics of Chain Molecules", Interscience, New York, N.Y., 1969.
- (4) P. J. Flory, *Macromolecules*, **7**, 381 (1974).
- (5) D. A. Brant and P. J. Flory, *J. Am. Chem. Soc.*, **87**, 663 (1965).
- (6) D. A. Brant and P. J. Flory, *J. Am. Chem. Soc.*, **87**, 2788 (1965).
- (7) D. A. Brant and P. J. Flory, *J. Am. Chem. Soc.*, **87**, 2791 (1965).
- (8) W. G. Miller and P. J. Flory, *J. Mol. Biol.*, **15**, 298 (1966).
- (9) D. A. Brant, W. G. Miller, and P. J. Flory, *J. Mol. Biol.*, **23**, 47 (1967).
- (10) A. E. Tonelli and F. A. Bovey, *Macromolecules*, **3**, 410 (1970).
- (11) W. L. Mattice and L. Mandelkern, *J. Am. Chem. Soc.*, **93**, 1769 (1971).
- (12) W. L. Mattice and J. T. Lo, *Macromolecules*, **5**, 734 (1972).
- (13) W. L. Mattice, K. Nishikawa, and T. Ooi, *Macromolecules*, **6**, 443 (1973).
- (14) T. Ooi, D. S. Clark, and W. L. Mattice, *Macromolecules*, **7**, 337 (1974).
- (15) S. Tanaka and H. A. Scheraga, *Macromolecules*, **8**, 623 (1975).
- (16) P. J. Flory and P. R. Schimmel, *J. Am. Chem. Soc.*, **89**, 6807 (1967).
- (17) W. G. Miller, D. A. Brant, and P. J. Flory, *J. Mol. Biol.*, **23**, 67 (1967).
- (18) P. R. Schimmel and P. J. Flory, *J. Mol. Biol.*, **34**, 105 (1968).
- (19) S. P. Rao, D. E. Carlstrom, and W. G. Miller, *Biochemistry*, **13**, 943 (1974).
- (20) W. L. Mattice and L. Mandelkern, *Biochemistry*, **10**, 1934 (1971).
- (21) W. G. Miller and C. V. Goebel, *Biochemistry*, **7**, 3925 (1968).
- (22) W. L. Mattice, J. M. Riser, and D. S. Clark, *Biochemistry*, **15**, 4264 (1976).
- (23) W. L. Mattice, *Macromolecules*, **8**, 644 (1975).
- (24) W. L. Mattice, *Macromolecules*, **9**, 48 (1976).
- (25) H. L. Yakei, Jr., and E. W. Hughes, *Acta Crystallogr.*, **7**, 291 (1954).
- (26) H. A. Scheraga, *Adv. Phys. Org. Chem.*, **6**, 103 (1968).
- (27) D. P. Stevenson and J. Y. Beach, *J. Am. Chem. Soc.*, **60**, 2872 (1938).
- (28) A. Abe, R. L. Jernigan, and P. J. Flory, *J. Am. Chem. Soc.*, **88**, 631 (1966).
- (29) H. D. Rudolph, D. Sutter, A. Jaeschke, and H. Dreizler, *Trans. Am. Crystallogr. Assoc.*, **2**, 197 (1966).
- (30) N. L. Allinger, M. J. Hickey, and J. Kao, *J. Am. Chem. Soc.*, **98**, 2741 (1976).
- (31) R. W. Woody, *Tetrahedron*, **29**, 1273 (1973).
- (32) J. Webb, R. W. Strickland, and F. S. Richardson, *J. Am. Chem. Soc.*, **95**, 4775 (1973).
- (33) D. L. Coleman and E. R. Blout, *J. Am. Chem. Soc.*, **90**, 2405 (1968).
- (34) M. Ottnad, C. Ottnad, P. Hartter, and G. Jung, *Tetrahedron*, **31**, 1155 (1975).