# Stereochemical Studies on Cyclic Peptides. IX. Conformational Studies on Cyclic Tetrapeptides Containing Alternating *cis* and *trans* Peptide Units*

G. MANJULA and C. RAMAKRISHNAN, *Molecular Biophysics Unit, Indian Institute of Science, Bangalore—560 012, India*

## Synopsis

Conformational analyses of cyclic tetrapeptides consisting of alternating *cis* and *trans* peptide units have been made using contact criteria and energy calculations. This study has been restricted to those structures having a symmetry element in the backbone ring, such as a twofold axis ($d$) or a center of inversion ($i$). There are five main results. (1) There are two *distinct* types of conformations, which are stereochemically favorable corresponding to each of twofold and inversion-symmetrical structures, designated as $d_1$, $d_2$ (for twofold symmetrical) and $i_1$, $i_2$ (for inversion-symmetrical). Among these, the $i_1$ type has the lowest energy when glycyl residues occur at all four $\alpha$-carbon atoms. (2) With the glycyl residue at all four $\alpha$-carbon atoms, methyl substitution at the *cis* peptide nitrogen atoms is possible in all the four types, whereas the substitution at *trans* peptide nitrogen atoms is possible only for the $i_1$ type. Thus only in the $i_1$ type can all the nitrogen atoms be methylated simultaneously. The conformation of the molecule in the crystal structure of cyclotetrasarcosyl belongs to the $i_1$ type. (3) When alanyl residues occur at all four $\alpha$-carbon atoms, the possible symmetrical type is dependent on the enantiomorphic form and the actual sequence of the alanyl residues. (4) The methyl substitution at peptide nitrogen atoms for cyclic tetrapeptides having alanyl residues causes more stereochemical restriction in the allowed conformations than with glycyl residues. (5) The prolyl residue can be incorporated favorably at the *cis-trans* junction of both $d$ and $i$ types of structures. The results of the present study are compared with the data on cyclic tetrapeptides available from the crystal structure and nmr studies. The results show an overall agreement both regarding the type of symmetry and the conformational parameters.

## INTRODUCTION

In an earlier paper of this series,[1] the conformational studies of cyclic tetrapeptides with all the peptide units in the *trans* conformation have been reported. At that time, the syntheses of only cyclotetraglycyl and cyclotetraalanyl were known,[2,3] and the nature of the peptide units in them was not clearly established. Since then, crystallographic and solution studies (by nmr) of a number of cyclic tetrapeptides[4–12] and tetradepsipeptides[13–16] have been reported. The majority of these contain *N*-methyl amino acid

* Contribution No. 116 from Molecular Biophysics Unit, Indian Institute of Science, Bangalore. Part VIII of this series is by Manjula et al. [(1977) *Proc. Ind. Acad. Sci.* **86A**, 443–454].

Biopolymers, Vol. 18, 591–607 (1979)
© 1979 John Wiley & Sons, Inc.
0006-3525/79/0018-0591$01.00

TABLE I
Symmetry of the Peptide Ring for the Various Cyclic Tetrapeptides and Tetradepsipeptides Studied by Experimental Methods

| Compound | Symmetry | Type of Study | Ref. |
|---|---|---|---|
| Peptide Sequence as *–cis-trans-cis-trans–* | | | |
| 1. Cyclo(Sar)$_4$ | Center of inversion | x-ray, nmr | 4,7 |
| 2. Dihydrotentoxin: cyclo(L-Leu-*N*-Me-D-Phe-Gly-*N*-Me-L-Ala) | Center of inversion | x-ray | 5 |
| 3. Cyclo(Sar-Sar-Sar-Gly) | Center of inversion | x-ray, nmr | 6,8 |
| 4. Cyclo(Sar-Gly-Sar-Gly) | Center of inversion | x-ray, nmr | 6,8 |
| 5. Cyclo(Sar-Sar-Sar-Ala) | Center of inversion | x-ray, nmr | 6,8 |
| 6. Cyclo(L-*N*-Me-Ilue-D-HyIV-L-*N*-Me-Leu-D-HyIV) | Twofold rotation | x-ray | 13 |
| 7. Cyclo(L-*N*-Me-Val-D-HyIV-L-*N*-Me-Val-D-HyIV) | Twofold rotation | x-ray | 14 |
| 8. Cyclo(D-*N*-Me-Val-L-HyIV-L-*N*-Me-Val-D-HyIV) | Center of inversion | x-ray | 15 |
| 9. Cyclo(D-*N*-Me-Val-D-HyIV-D-*N*-Me-Val-L-HyIV) | No symmetry | x-ray | 16 |
| 10. Cyclo(L-Pro-Gly)$_2$ | No symmetry | nmr | 10 |
| 11. Tentoxin: cyclo(*N*-Me-Ala-L-Leu-*N*-Me-(*Z*)-de Phe-Gly) | Twofold rotation | nmr | 12 |
| Peptide Sequence as *–trans-trans-trans-trans–* | | | |
| 1. Cyclo(Gly)$_4$ | Fourfold rotation, reflection | nmr | 9 |
| 2. Dihydrochlamydocin: Cyclo(Iabu-L-Phe-D-Pro-LX) | Twofold rotation | x-ray | 11 |

residues and are made up of both *cis* and *trans* peptide units, with the peptide sequence as *–cis-trans-cis-trans–*. The *cis* peptide units that do not normally occur in protein chains, except occasionally with prolyl residues, are essential in the formation of cyclic dipeptides[17,18] and cyclic tripeptides.[19–21] It has been established[1,22] that cyclic tetrapeptides cannot be formed with standard planar *trans* peptide units. Acceptable conformations can be obtained if the peptide units deviate from planarity by about 10°–15°. On the other hand, the *cis* peptide units, in combination with *trans* peptide units, can bring about reasonably good structures without invoking much nonplanarity. In this paper, these types of cyclic tetrapeptide structures are examined from a stereochemical point of view, using "contact criteria" and potential energy considerations.[23] This study has been confined to structures in which the peptide ring has a symmetry element, such as a twofold rotation axis or a center of inversion. While a complete study should include the nonsymmetrical conformations as well, this is not attempted here because of the extensive computer time required. Furthermore, in many of the cyclic tetrapeptides that have been studied experimentally, the peptide ring is found to have at least an approximate symmetry. The symmetry of the ring in the various cyclic tetrapeptides

and tetradepsipeptides studied by crystal structure analysis and by nmr methods is given in Table I.

### Combination of *cis* and *trans* Peptide Units

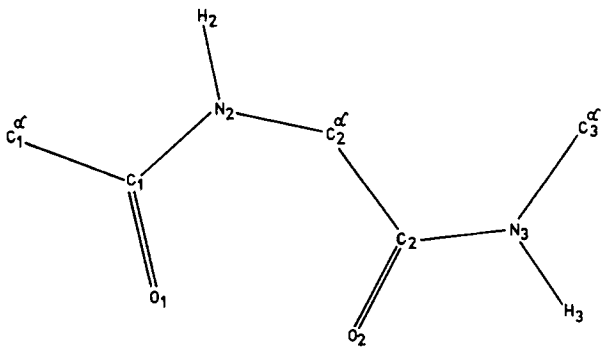
Formulation of cyclic tetrapeptides with both *cis* and *trans* peptide units can give rise to four possible combinations. (In this paper, the *trans* peptide unit will be referred to by the symbol *t* and the *cis* peptide unit by the symbol *c*.) The possibilities are (1) all *c*'s, (2) three *c*'s and one *t*, (3) two *c*'s and two *t*'s, and (4) one *c* and three *t*'s.

Preliminary model building shows that all these four types of structures can be formed geometrically, with a certain amount of nonplanarity for the peptide units. It has been shown by Sarathy[24] that an all-*cis* cyclic tetrapeptide structure is not stereochemically favorable. Possibilities (2) and (4) cannot have any symmetry for the ring and hence are not considered here. Combination (3) has two *c*'s and two *t*'s, which can be arranged in two ways, namely, –*c*-*c*-*t*-*t*– and –*c*-*t*-*c*-*t*–. The former sequence cannot have any symmetry in its ring, and this type of sequence has not been observed. Hence only the peptide sequence –*c*-*t*-*c*-*t*– has been considered in this paper.

### Method

The method used to form the cyclic tetrapeptides is as follows: two pairs of identical peptide units, each pair consisting of linked *t* and *c* peptide units, were joined together under appropriate conditions. The cyclic tetrapeptides having twofold symmetry will be denoted by the symbol *d* and those having inversion symmetry by *i*. The method used for the generation of *d*-symmetrical structures is the same as that given in Ref. 1. For generating the *i*-symmetrical structures, the center of inversion was taken to be at the midpoint of the line joining the terminal $C^{\alpha}$ atoms of the pair of linked peptide units. Thus the two halves of the cyclic tetrapeptide are automatically obtained and linked. However, the bond angle $\tau(NC^{\alpha}C)$ at the two $C^{\alpha}$ atoms may not always be around the tetrahedral value, and therefore, only certain sets of linked peptide units will be found suitable to form geometrically satisfactory structures. For both types of symmetrical cases, a pair of peptide units, with a given set of parameters ($\tau$, $\phi$, $\psi$),[25] is first generated and then the symmetry condition is applied. The method is shown schematically in Fig. 1.

In the present study, the initial pair of peptide units has been chosen as a system of linked *trans*-*cis* peptide units as shown in Scheme I:




Scheme I

Consequently, the third peptide unit will be $t$ and the fourth $c$. The atoms $C_2^\alpha$, $C_4^\alpha$ will then be situated at $t$-$c$ junctions and $C_1^\alpha$, $C_3^\alpha$ at $c$-$t$ junctions. Further, it is sufficient to specify the parameters at two consecutive $\alpha$-carbon atoms (say $C_1^\alpha$ and $C_2^\alpha$), as those at the other two $\alpha$-carbon atoms are automatically specified due to symmetry by the general relationships:

$$\tau_{n+2} = \tau_n, \quad \phi_{n+2} = \phi_n, \quad \text{and} \quad \psi_{n+2} = \psi_n \text{ for } d \text{ types}$$

$$\tau_{n+2} = \tau_n, \quad \phi_{n+2} = -\phi_n, \quad \text{and} \quad \psi_{n+2} = -\psi_n \text{ for } i \text{ types}$$

The geometrically possible structures were first examined using stereochemical criteria for the backbone atoms only. This was done by superposing the $(\phi - \psi)$ regions, which can give rise to geometrically feasible structures, over the $(\phi - \psi)$ contact map. Since the structure involves two different types of junctions ($c$-$t$ and $t$-$c$), it is necessary to do the superposition with two different $(\phi - \psi)$ maps. By this method conformations having short contacts between atoms of adjacent units are eliminated, and those left out are further examined for possible steric hindrance between atoms of opposite (symmetrically related) peptide units. Those conformations that do not have severe short contacts are further studied through energy considerations. In the present study, only the simple side groups, like glycine and alanine, are considered.

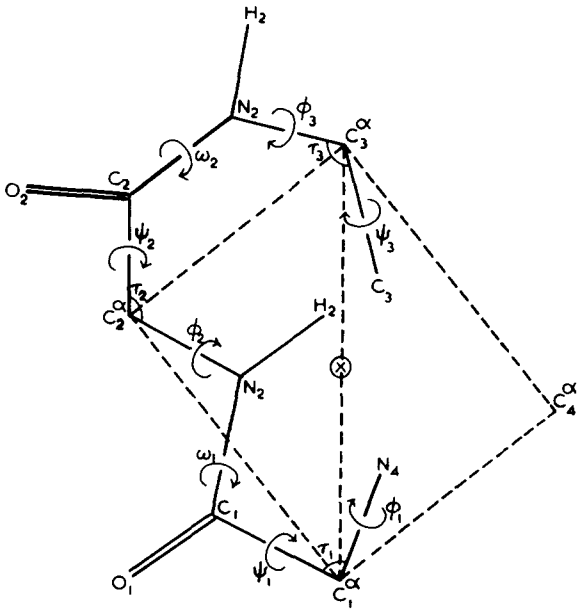


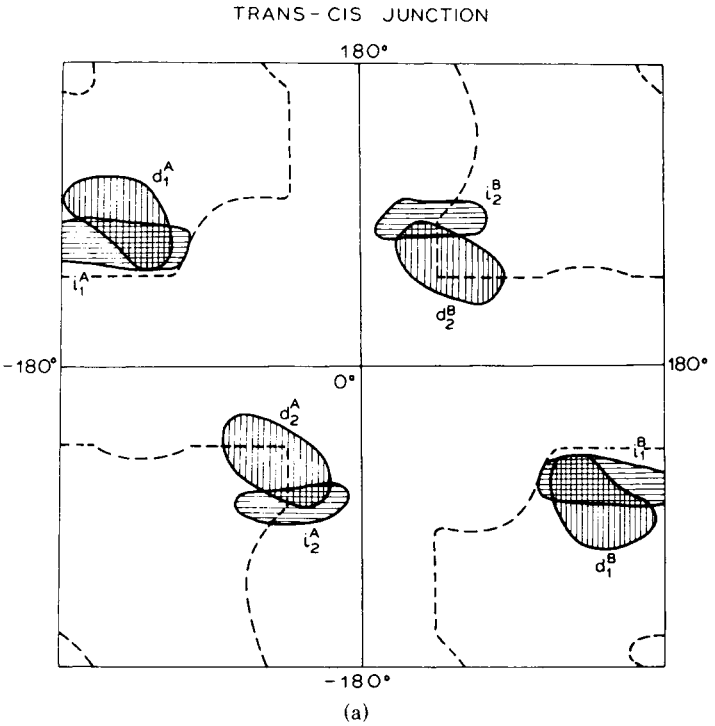

Fig. 1. Schematic diagram of a *trans-cis* dipeptide used in the generation of the symmetrical structures. The ⊗ corresponds to a twofold rotation axis for the $d$-symmetrical type and to a center of inversion for $i$-symmetrical type of structure.

## RESULTS AND DISCUSSION

### Two-Fold and Inversion Symmetrical Structures

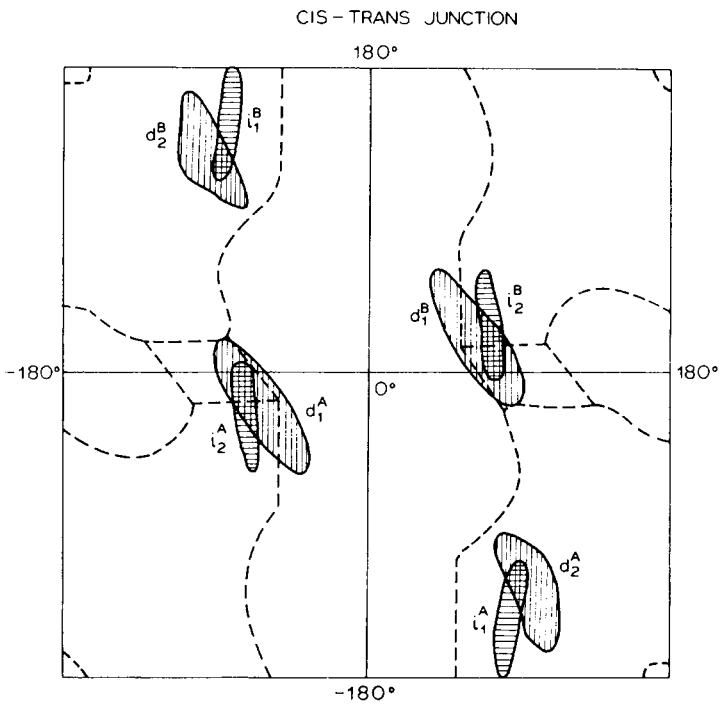
It was found that the major portion of geometrically allowed regions (a broad region in the case of $d$ symmetry and two narrow strips with $i$ symmetry)[26,27] does not overlap with the stereochemically allowed regions of the contact maps, even for glycyl residues The combination of geometry and stereochemistry results in four distinct regions, two each for $d$ and $i$ symmetries. This is shown in Fig. 2, marked as $d_1^A$, $d_2^A$, $i_1^A$, and $i_2^A$, in the two maps. (For details, see Figs. 1 and 2 of Ref. 26.) Since the backbone atoms alone are considered, four other symmetrically related regions denoted by the superscript B can be obtained. Thus, there are four types ($d_1$, $d_2$, $i_1$, and $i_2$) of stereochemically favorable symmetrical structures of cyclic tetrapeptide. The projection drawings of these conformations are shown in Fig. 3.

It can be seen from Fig. 3 that in all the four types of conformations, the C=O groups are oriented nearly perpendicular to the approximate plane formed by the four $\alpha$-carbon atoms. The actual directions (up or down) of the C=O bonds with respect to the ring are given in Table II.




(a)

Fig. 2. The stereochemically permitted regions belonging to the four different types are shown on the contact maps with glycyl residue, corresponding to the $t$-$c$ junction (a) and the $c$-$t$ junction (b).




(b)

Fig. 2. (*continued from previous page*)

The potential energy calculations were carried out for the different types with glycyl residues at all four $\alpha$-carbon atoms by varying the values of ($\phi$, $\psi$) in their corresponding regions at 5° intervals. The forms and the constants of the energy function used for the calculations are those given by Ramachandran and Sasisekharan.[23] The conformational parameters and energy values in the increasing order of their energy up to 0.6 kcal/mol/residue from their respective minima are listed in Table III for the four types. Comparing the minimum energy values for the different types, it can be seen that the $i_1$ type has the lowest minimum, and the minima corresponding to the types $d_1$, $d_2$, and $i_2$ are higher by more than 0.8 kcal/

TABLE II
Orientation of C=O Groups with Respect to the Approximate Plane of the Ring in *d* and *i* Type Cyclic Tetrapeptide Structures

| Symmetry and Type | Peptide Units | | | |
|---|---|---|---|---|
| | *trans*[a] | *cis* | *trans* | *cis* |
| $d_1$ | Up | Down | Up | Down |
| $d_2$ | Up | Up | Up | Up |
| $i_1$ | Up | Down | Down | Up |
| $i_2$ | Up | Up | Down | Down |

[a] The orientation in this unit is arbitrarily taken as up.

TABLE III
Conformational Parameters and Energy Values (Arranged According to the Increasing Order of Energy) for the Four Types of Cyclotetraglycyl Conformations

| Conformational Parameters (deg) | | | | | | Energy, $V_{tot}$ |
|---|---|---|---|---|---|---|
| $\tau_1$ | $\phi_1$ | $\psi_1$ | $\tau_2$ | $\phi_2$ | $\psi_2$ | (kcal/mol/residue) |
| | | | $d_1$ Type[a] | | | |
| 114 | −70 | −20 | 110 | −135 | 85 | 0.07 |
| 115 | −65 | −35 | 110 | −135 | 90 | 0.19 |
| 115 | −70 | −15 | 110 | −140 | 85 | 0.28 |
| 115 | −55 | −40 | 110 | −140 | 45 | 0.32 |
| 111 | −60 | −30 | 110 | −140 | 95 | 0.49 |
| 114 | −70 | −25 | 110 | −130 | 85 | 0.49 |
| 111 | −60 | −30 | 110 | −145 | 95 | 0.53 |
| 112 | −60 | −25 | 110 | −145 | 95 | 0.75 |
| | | | $d_2$ Type[b] | | | |
| 111 | 100 | −125 | 110 | −50 | −60 | 0.27 |
| 110 | 100 | −130 | 110 | −45 | −60 | 0.35 |
| 107 | 100 | −125 | 110 | −50 | −65 | 0.36 |
| 106 | 100 | −130 | 110 | −45 | −65 | 0.44 |
| 106 | 100 | −125 | 110 | −60 | −60 | 0.46 |
| 108 | 105 | −135 | 110 | −55 | −55 | 0.53 |
| 112 | 100 | −120 | 110 | −55 | −60 | 0.54 |
| 107 | 105 | −140 | 110 | −50 | −55 | 0.64 |
| 109 | 105 | −130 | 110 | −60 | −55 | 0.67 |
| 112 | 95 | −120 | 110 | −45 | −65 | 0.70 |
| 111 | 95 | −125 | 110 | −40 | −65 | 0.71 |
| 109 | 100 | −135 | 110 | −40 | −60 | 0.72 |
| 115 | 100 | −130 | 110 | −50 | −55 | 0.76 |
| 108 | 100 | −120 | 110 | −55 | −65 | 0.77 |
| 114 | 100 | −135 | 110 | −45 | −55 | 0.86 |
| 107 | 95 | −120 | 110 | −40 | −70 | 0.94 |
| | | | $i_1$ Type[c] | | | |
| 112 | 85 | −155 | 110 | −135 | 65 | −1.27 |
| 112 | 85 | −160 | 110 | −130 | 65 | −1.25 |
| 113 | 85 | −150 | 110 | −140 | 65 | −1.17 |
| 108 | 85 | −160 | 110 | −135 | 70 | −1.11 |
| 108 | 85 | −150 | 110 | −140 | 70 | −1.09 |
| 109 | 85 | −150 | 110 | −145 | 70 | −1.00 |
| 113 | 85 | −150 | 110 | −145 | 65 | −1.00 |
| 112 | 85 | −165 | 110 | −125 | 65 | −0.97 |
| 109 | 85 | −140 | 110 | −150 | 70 | −0.89 |
| 114 | 90 | −140 | 110 | −150 | 65 | −0.81 |
| 110 | 85 | −140 | 110 | −155 | 70 | −0.75 |
| 107 | 80 | −165 | 110 | −125 | 70 | −0.64 |
| | | | $i_2$ Type[d] | | | |
| 115 | −70 | −35 | 110 | −35 | −80 | 1.16 |
| 112 | −70 | −20 | 110 | −50 | −85 | 1.55 |
| 111 | −70 | −25 | 110 | −45 | −85 | 1.59 |
| 114 | −70 | −40 | 110 | −30 | −80 | 1.63 |
| 111 | −70 | −30 | 110 | −40 | −85 | 1.78 |

[a] These correspond to regions $d_1^A$ in Fig. 2.
[b] These correspond to regions $d_2^A$ in Fig. 2.
[c] These correspond to regions $i_1^A$ in Fig. 2.
[d] These correspond to regions $i_2^A$ in Fig. 2.

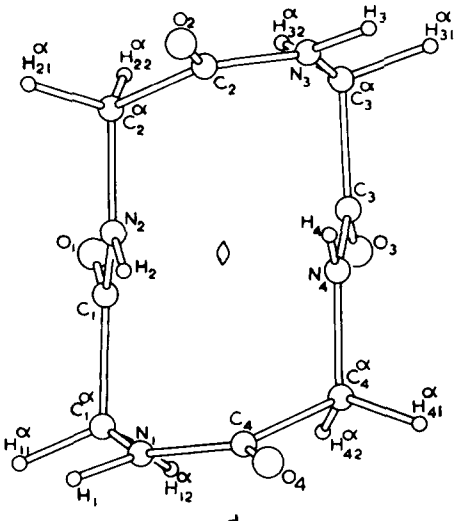



Fig. 3. Projection diagrams of types (a) $d_1$ (b) $d_2$, (c) $i_1$, and (d) $i_2$ of cyclotetraglycyl molecule.

mol/residue. The conformation of cyclo($Gly_4$) as determined by nmr studies[9] indicates an all-*trans* sequence having nonplanar peptide units and thus cannot be compared with the results of the present study. The relative stabilities of cyclo($Cly_4$) under different symmetrical conditions have been studied, and the results will be published elsewhere.

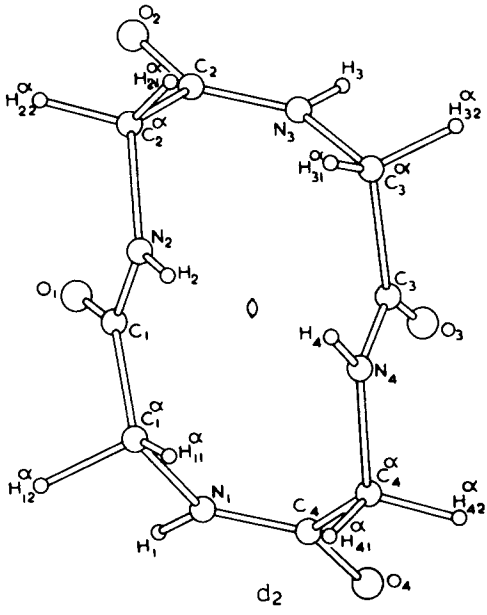
## *N*-Methyl Substituted Cyclic Tetrapeptide with Glycyl Residues

It can be seen from Table I that all the cyclic tetrapeptides except cyclo($Gly_4$) and dihydrochlamydocin have at least one methylated nitrogen atom and have the peptide sequence $\underline{-c\text{-}t\text{-}c\text{-}t-}$. The effect of methyl substitution on the low-energy conformations obtained with normal peptide units was studied using stereochemical "contact criteria" corresponding to the following situations:

1. The two *cis* peptide units containing *N*-methyl groups and the two *trans* units containing the usual NH groups;
2. The two *trans* peptide units containing *N*-methyl groups and the two *cis* units containing NH groups; and
3. All four peptide units containing *N*-methyl groups.

The principal results are summarized in Table IV, in which the minimum energy values when there are glycyl residues at the four $\alpha$-carbon atoms are given in parentheses. The following conclusions were drawn:

1. It is possible to substitute the methyl groups at the two *cis* peptide units without hindrance for all types of structures.
2. Simultaneous substitution of methyl groups at the two *trans* peptide units is possible only for the $i_1$ type.




(b)

Fig. 3. (*continued from previous page*)

3. For types $d_1$ and $i_2$, the methyl group cannot be substituted even at one *trans* unit, because the N—$CH_3$ bonds in these types point towards the inside of the peptide ring, and hence methyl groups come very close to the atoms of the opposite *trans* peptide units.
4. In the case of the $d_2$ type, the methyl groups at the two *trans* units give rise to very short contacts between them, and hence substitution at only one of the *trans* units is possible for this type.

TABLE IV

Effect of Methyl Substitution at Nitrogen Atoms on the Different Cyclic Tetrapeptide Structures Having a Glycyl Side Group at All the $\alpha$-Carbon Atoms[a]

| Type | Methyl Substitution | |
|---|---|---|
| | At *cis* peptide units | At *trans* peptide units |
| $d_1$ | 0.23[b] | c |
| $d_2$ | 0.36[b] | d |
| $i_1$ | −1.11[b] | 0.87[b,e] |
| $i_2$ | 0.52[b] | c |

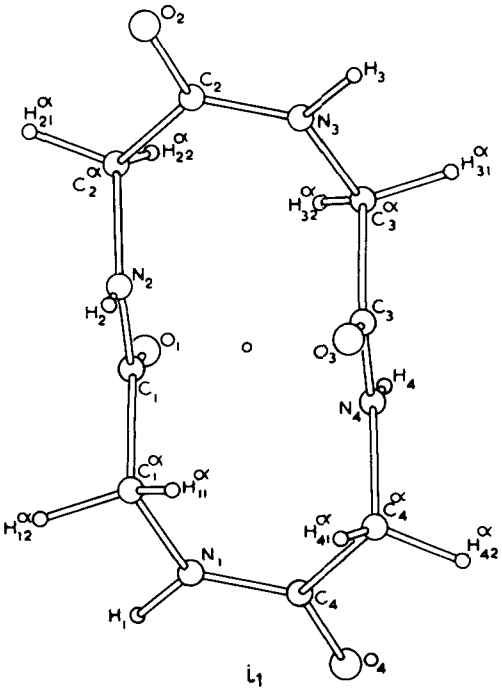
[a] The minimum energy values (in kcal/mol/residue) are given wherever methyl substitution is possible.

[b] Methyl groups *can* be substituted at both the opposite peptide units simultaneously.

[c] Methyl groups *cannot* be substituted at both the opposite peptide units simultaneously.

[d] Substitution is possible only at one *trans* peptide unit and not at both.

[e] The value of minimum energy when all the four units have an *N*-methyl group is −0.4 kcal/mol/residue.



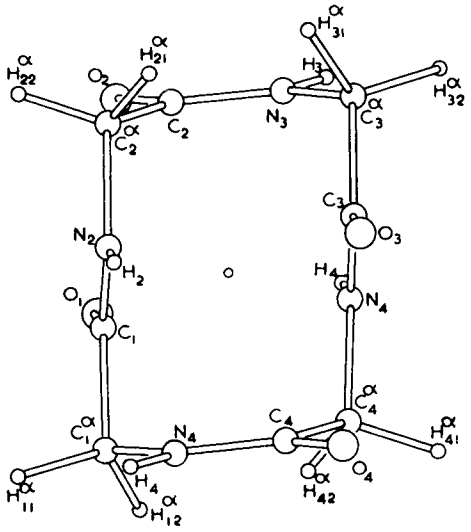

(c)

Fig. 3. (*continued from previous page*)

5. All the peptide units can be methylated only for the $i_1$ type.

Thus the compound cyclotetrasarcosyl, in which all four peptide units are methylated, can take up only the $i_1$ type of structure. This conclusion, which results from a systematic study, is in complete agreement with observations.[4,7]

The crystal structure data on three compounds containing glycyl and sarcosyl residues are available in the literature. In the crystal structure of cyclotetrasarcosyl,[4] the molecule is of the $i_1$ type and is the only type possible when all four peptide units are methylated. In the crystal structure of cyclo(Gly-Sar-Sar-Sar),[6] two of the three *N*-methyl groups (of sarcosyl residues) occur at the opposite *cis* units and the remaining one at one of the *trans* units. According to Table IV, this is possible for types $d_2$ and $i_1$. Among these two types, the conformation $i_1$ has a lower minimum than $d_2$ (by about 3 kcal/mol/residue). The observed conformation of this molecule is also of $i_1$ type. In cyclo(Gly-Sar-Gly-Sar),[6] *N*-methyl groups of the two sarcosyl residues occur at the two opposite *cis* units. According to Table IV, it is possible for all four types of cyclic tetrapeptide structures to accommodate methyl groups at the *cis* peptide units. However, the $i_1$ type has the lowest minimum, and the observed conformation of the molecule is also of the $i_1$ type.




(d)

Fig. 3. (*continued from previous page*)

In all the three examples mentioned above, the conformation of the molecule is of the $i_1$ type. The conformational parameters of the molecules (reported in Refs. 4 and 6) are compared with those calculated in the present study. Since in the observed molecules both *cis* and *trans* peptide units are nonplanar, a deviation of ±10° from planarity was introduced to obtain low-energy conformations so that effective comparison could be made. The results for these compounds are given in Table V. The conformation that fits best with observations is also given in Table V. Energetically, these are very close to the minimum energy conformations.

While the application of "contact criteria" is fully valid, comparison of the minimum energy conformations in the present study with those observed in the experimental study, such as x-ray or nmr, is open to question, since neither the intermolecular crystal packing interactions in the solid state nor the solvent interactions in solution has been taken into account. The authors feel that in the case of cyclic peptides, especially with a small number of residues, the molecule is fairly rigid, and the intramolecular interactions play a major role in stabilizing the conformation. The other effects can, however, introduce slight distortions, and hence the observed conformation need not exactly correspond to the minimum energy conformation. This is the reason why the conformations that fit best with the observation have also been listed in Table V. It can be seen that the deviations are small and reasonable.

TABLE V
Conformational Parameters of the Observed and Calculated Cyclic Tetrapeptide Conformations

| Compound | Conformation | Parameters (deg) | | | | | | | | Energy (kcal/mol/residue) |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $\omega_1$ | $\tau_1$ | $\phi_1$ | $\psi_1$ | $\omega_2$ | $\tau_2$ | $\phi_2$ | $\psi_2$ | |
| Cyclo($Sar_4$) | Observed | 170 | 112 | 94 | −170 | 5 | 111 | −121 | 66 | |
| | Minimum energy | 170 | 113 | 95 | −170 | 0 | 110 | −120 | 65 | −0.43 |
| | Best fit | 170 | 109 | 95 | −170 | 5 | 110 | −120 | 65 | −0.31 |
| Cyclo(Gly-Sar)$_2$ | Observed | 173 | 113 | 85 | −166 | −7 | 111 | −125 | 71 | |
| | Minimum energy | 175 | 114 | 90 | −160 | 0 | 110 | −130 | 65 | −1.11 |
| | Best fit | 175 | 113 | 85 | −165 | −5 | 110 | −125 | 70 | −1.00 |
| Cyclo(Gly-Sar-Sar-Sar)[a] | Observed | 172 | 114 | 92 | −168 | 5 | 110 | −122 | 63 | |
| | Minimum energy | 170 | 113 | 90 | −170 | −5 | 110 | −120 | 70 | −0.57 |
| | Best fit | 170 | 110 | 95 | −165 | 5 | 110 | −125 | 65 | −0.30 |

[a] The observed conformation has only an approximate center of inversion for the peptide ring. The values given here for the various parameters of the observed conformation are mean values.

## Incorporation of L- and D-Alanine Side Groups

The stereochemically feasible cyclic tetrapeptide conformations having L- and D-alanyl residues were deduced by examining the different types of cyclic tetrapeptide structures against a background of the corresponding contact maps for the alanine dipeptide (Fig. 4). The results are summarized in Table VI, in which the possible sequences and the minimum value of the total energy for that sequence are listed. The following conclusions can be drawn:

1. Cyclic tetrapeptides having L- (or D-) residues at all the $\alpha$-carbon atoms can take up only *d*-symmetrical, and *not* the *i*-symmetrical, structures. (It should be noted that the symmetry considered here applies only to the backbone.)
2. When the structure contains both L- and D-residues, depending on the actual sequence, either the *d* or *i* type becomes allowed. For example, the sequence –LLDD– is not possible for any of the types. It may appear that for the *i* type, since –DLLD– is possible, –LLDD– must also be possible because of the cyclic nature. But this is not so, since there are two types of junctions, *c-t* and *t-c*. This makes the two sequences conformationally different.

The information contained in Table VI can be effectively used to pick out the symmetrical type of conformation that any given isomeric sequence of nonglycyl residues can assume.

TABLE VI
Possibility of Incorporation of Alanyl Side Groups at the Four $\alpha$-Carbon Atoms in the Cyclic Tetrapeptide Structures[a]

| Type | Preferred L- or D-Ala at $C_1^\alpha$ ($c$-$t$) Junction | $C_2^\alpha$ ($t$-$c$) Junction | $C_3^\alpha$ ($c$-$t$) Junction | $C_4^\alpha$ ($t$-$c$) Junction | Possible Sequences | Energy (kcal/mol/residue) |
|---|---|---|---|---|---|---|
| $d_1$ | L,D | L | L,D | L | L L L L | 1.2 |
| | | | | | D L D L | 4.5 |
| | | | | | D L L L<br>L L D L | 3.2 |
| $d_2$ | D | L,D | D | L,D | D D D D | 2.0 |
| | | | | | D L D L | 3.0 |
| | | | | | D D D L<br>D L D D | 3.1 |
| $i_1$ | D | L | L | D | D L L D | 0.3 |
| $i_2$ | L | D | D | L | L D D L | 2.8 |

[a] The possible residues and the sequences mentioned here correspond to the region denoted by the superscript A in Fig. 4. The enantiomeric residues and sequences are equally possible stereochemically for the region denoted by the superscript B in the same figure.

## Effect of *N*-Methyl Groups with Alanyl Residues

The interaction of the alanyl side group with the *N*-methyl group will naturally cause more stereochemical restrictions than the glycyl side group. The types of conformations that are stereochemically possible under these conditions are summarized in Table VII. It is clear that the methyl substitution eliminates many sequences that are otherwise possible with normal peptide units. This is caused mainly by the unfavorable close contact between the alanyl side group and the methyl group attached to the nitrogen atom of the preceding peptide unit.

Our results can be compared with the available crystal structure data on cyclic tetrapeptides containing nonglycyl amino acid residues. In dihydrotentoxin,[5] alternate amino acid residues are *N*-methylated, and the sequence of the residues is –LDGL– or –GLLD–, where G stands for glycyl. Since the glycyl residue can easily be accommodated in place of either the L- or D-residue, the sequences to be tested are

$$\underline{\text{–LDLL–}},\ \underline{\text{–LDDL–}},\ \underline{\text{–LLLD–}}$$

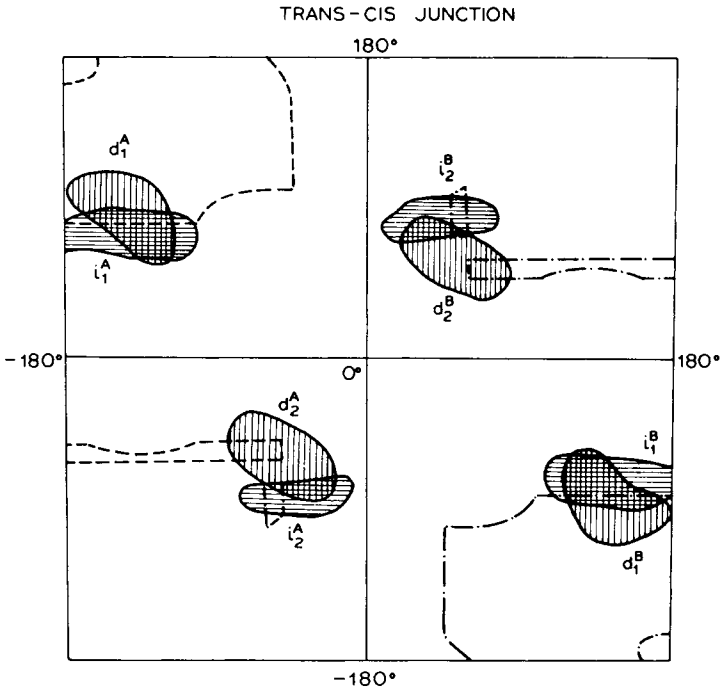
and

$$\underline{\text{–DLLD–}}$$

It can be seen from Table VII that only the sequence

$$\underline{\text{–DLLD–}}$$

can be accommodated and only by type $i_1$, with the two *trans* units being *N*-methylated. The observed conformation is also of $i_1$ type, and the *N*-methyl groups occur at *trans* units. This causes the ordinary amino acids (the secondary amide bonds) to be in the *cis* conformation. In all the




(a)

Fig. 4. The stereochemically permitted regions belonging to the four different types are shown on the contact maps with alanyl residue, corresponding to the *t-c* junction (a) and the *c-t* junction (b). The areas enclosed by the broken line show the regions allowed by the "extreme limit" contact criteria with L-alanyl residue and those enclosed by the dots and dashes show the corresponding regions with D-alanyl residue.

other cyclic tetrapeptides[4,6] and tetradepsipeptides[13–16] whose crystal structures are known, it is found that *N*-methyl groups occupy *cis* units in preference to *trans* units. If *N*-methyl groups were to occur in *cis* units for the sequence

$$\underline{\text{–GLLD–}}$$

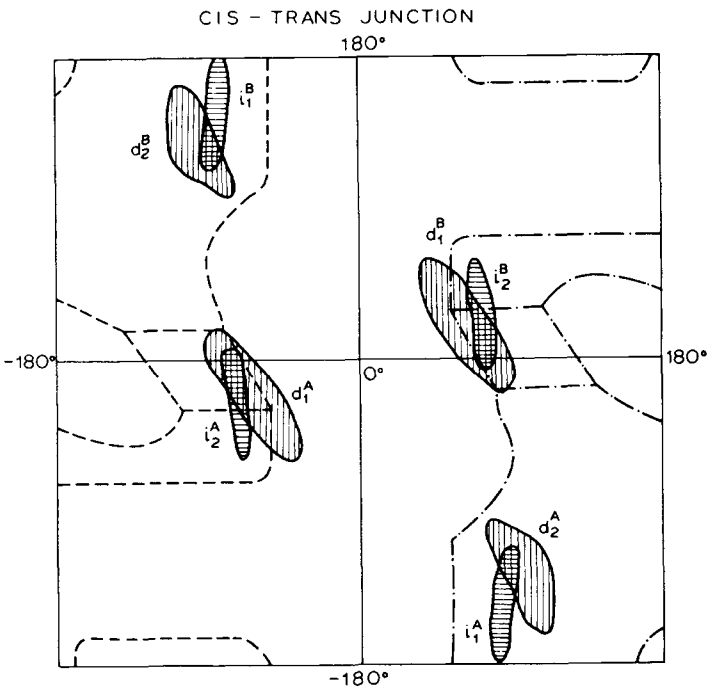
of dihydrotentoxin, none of the types of conformation is stereochemically

TABLE VII

Stereochemically Possible Sequences of L- and D-Alanyl Residues with *N*-Methylation in Different Types of Cyclic Tetrapeptide Conformations

| Type | N-$CH_3$ in | L or D at $C_1^\alpha$ (*c-t*) | L or D at $C_2^\alpha$ (*t-c*) | Possible Sequence |
|---|---|---|---|---|
| $d_1$ | *cis* | D | L | DLDL |
| $d_2$ | *trans* | D | D[a] | DDDD[a] |
| $i_1$ | *trans* | D | L | DLLD |

[a] Possible with methyl groups in only one of the *trans* peptide units.




(b)

Fig. 4. (*continued from previous page*)

favorable. This is perhaps the reason why the N-$CH_3$ groups occur in *trans* units in this compound.

In cyclo(D,L)(Ala-Sar-Sar-Sar),[6] there is only one nonglycyl residue, and the sequence is

$$\underline{\text{-GGGL-}} \text{ or } \underline{\text{-GGGD-}}$$

Types $d_1$ and $i_2$ are not favored, since the three consecutive amino acid residues are *N*-methylated. Among the other two types, $d_1$ and $i_2$, the minimum energy value for the $i_1$ (with the *cis* peptide units and one of the *trans* units *N*-methylated) is lower than that of $d_2$ by about 4.5 kcal/mol/residue. The observed conformation of the molecule is also the $i_1$ type, and the *N*-methyl groups occur in two *cis* units and one *trans* unit.

## Accommodation of Prolyl Residues

It is well known that the prolyl residue can assume both *cis* and *trans* forms. The feasibility of accommodating prolyl residues has been investigated here in a preliminary way by examining the range of $\phi$ values at the $\alpha$-carbon atoms. Subsequently, by fixing the proline group, the conformations were checked for possible short contacts. A summary of the results is given in Table VIII, where the conformation of the prolyl residue, as

TABLE VIII
Possibility of Incorporation of Prolyl Side Group at Four $\alpha$-Carbon Atoms in the Cyclic Tetrapeptide Structures

| Type | Incorporation at[a] $C_1^\alpha$ ($c$-$t$) | $C_2^\alpha$ ($t$-$c$) |
|---|---|---|
| $d_1$ | L-Pro ($cis'$) | c |
| $d_2$ | D-Pro ($trans'$) | L-Pro[b] ($cis'$) |
| $i_1$ | D-Pro ($trans'$) | c |
| $i_2$ | L-Pro ($cis'$) | c |

[a] The prolyl conformation is indicated in parentheses. The notations *cis'* and *trans'* correspond to regions with $\psi$ angle around $-60°$ and $120°$, respectively, for L-Pro.

[b] Simultaneous occurrence of proline residues at the two $t$-$c$ junctions not stereochemically permitted.

[c] Not possible.

determined by the value of $\psi$ at the prolyl $C^\alpha$ is indicated in parentheses. Table VIII shows the following:

1. For the $d_1$, $i_1$, and $i_2$ types of structures, the prolyl residue can be accommodated only at the $c$-$t$ junctions (i.e., the prolyl residue has to be in *cis* form). When the prolyl residues occur at the two $c$-$t$ junctions, they have to be in the same enantiomorphic form for the $d_1$ type and in the opposite forms for the $i$ types.
2. The $d_2$ type can accommodate the prolyl residue at all the $\alpha$-carbon atoms. However, due to steric hindrance, simultaneous occurrence of prolyl residues at the two $t$-$c$ junctions is forbidden.

Deber et al.[10] studied the conformation of cyclo(L-Pro-Gly)$_2$ by nmr. Their proposed conformation does not seem to possess any symmetry, but contains the

$$\underline{\text{-}cis\text{-}trans\text{-}cis\text{-}trans\text{-}}$$

sequence for the backbone, with the two prolyl residues having *cis'* and *trans'* conformations. It appears that the results of the nmr studies can still be satisfied if it is assumed that both $d_1$ and $d_2$ types of structures are simultaneously present in solution and are interchanging with one another. In their proposed conformation, one-half of the molecule corresponds to $d_1^A$ and the other half to $d_2^B$ and, as such, can be designated as $(\frac{1}{2}d_1^A + \frac{1}{2}d_2^B)$ in our notation. Thus it is an intermediate between $d_1^A$ and $d_2^B$. From Table II it can be seen that for type $d_1$, the adjacent oxygen atoms are pointing in the opposite directions; whereas for the type $d_2$, all four oxygen atoms are pointing in the same direction. In the observed conformation, the three consecutive oxygen atoms are on one side of the ring, with the remaining one pointing in the opposite direction. Thus, if one of the *trans* peptide units in the $d_1$ or $d_2$ type is flipped by 180°, the resulting conformation will correspond to their model. The detailed energy calculation

on cyclo(Pro-Gly)$_2$ to find out the relative stabilities of symmetrical and asymmetrical conformations is underway. We believe that like the symmetrical conformations of cyclo(L-Pro-Gly)$_3$ (Ref. 28) and cyclo(L-Pro-Gly)$_4$ (Ref. 29) in solution, this molecule should also exhibit a twofold symmetrical conformation in a nonpolar solvent.

The authors would like to thank Professor G. N. Ramachandran for his interest and Dr. P. Balaram, with whom they had many useful discussions. This work has been partly supported by a grant from Department of Science and Technology, India, and by PL-480 USPHS Grant 01-126-N.

## References

1. Ramakrishnan, C. & Sarathy, K. P. (1968) *Biochem. Biophys. Acta* **168,** 402–410.
2. Schwyzer, R., Iselin, B., Rittel, W. & Steber, P. (1956) *Helv. Chim. Acta* **339,** 872–883.
3. Balasubramanian, D. & Wetlaufer, D. B. (1967) in *Conformation of Biopolymers,* Vol. 1, Ramachandran, G. N., Ed., Academic, London, pp. 147–156.
4. Groth, P. (1970) *Acta Chem. Scand.* **24,** 780–790.
5. Meyer, W. L., Kuyper, L. F., Phelps, D. W. & Cordes, A. W. (1974) *Chem. Commun.,* 339–340.
6. Declercq, J. P., Germain, G., Van Meersche, M., Dabaerdemaeker, T., Dale, J. & Titlestad, K. (1975) *Bull. Soc. Chim. Belg.* **84,** 275–287.
7. Dale, J. & Titlestad, K. (1969) *Chem. Commun.,* 656–659.
8. Dale, J. & Titlestad, K. (1970) *Chem. Commun.,* 1403–1404.
9. Grathwohl, C., Tun-kyi, A., Bundi, A., Schwyzer, R. & Wüthrich, K. (1975) *Helv. Chim. Acta* **58,** 415–423.
10. Deber, C. M., Fossel, E. T. & Blout, E. R. (1974) *J. Am. Chem. Soc.* **96,** 4015–4017.
11. Flippen, J. L. & Karle, I. L. (1976) *Biopolymers* **15,** 1081–1092.
12. Rich, D. H. & Bhatnagar, P. (1977) in *Peptides: Proceedings of the Fifth American Peptide Symposium,* Goodman, M. & Meinhofer, J., Eds., Wiley, New York, pp. 340–342.
13. Konnert, J. & Karle, I. L. (1969) *J. Am. Chem. Soc.* **91,** 4888–4892.
14. Tishchenko, G. N., Karimov, Z. & Nazimova, N. V. (1975) *Bioorg. Khim.* **1,** 375–377.
15. Tishchenko, G. N., Karimov, Z. & Borisov, V. V. (1975) *Bioorg. Khim.* **1,** 378–385.
16. Tishchenko, G. N., Nazimova, N. V., Andrianov, V. I. & Karimov, Z. (1975) *Bioorg. Khim.* **1,** 386–392.
17. Degeilh, R. & Marsh, R. E. (1959) *Acta Crystallogr.* **12,** 1007–1014.
18. Sletten, E. (1970) *J. Am. Chem. Soc.* **92,** 172–177.
19. Venkatachalam, C. M. (1968) *Biochim. Biophys. Acta* **168,** 397–401.
20. Kartha, G., Ambady, G. & Shankar, P. V. (1974) *Nature* **247,** 204–205.
21. Deber, C. M., Torchia, D. A. & Blout, E. R. (1971) *J. Am. Chem. Soc.* **93,** 4893–4897.
22. Go, N. & Scheraga, H. A. (1970) *Macromolecules* **3,** 178–187.
23. Ramachandran, G. N. & Sasisekharan, V. (1968) *Adv. Protein Chem.* **23,** 283–437.
24. Sarathy, K. P. (1970) Ph.D. thesis, Madras University, Madras, India.
25. IUPAC-IUB Commission on Biochemical Nomenclature (1970) *Biochemistry* **9,** 3471–3479.
26. Manjula, G. & Ramakrishnan, C. (1977) in *Peptides: Proceedings of the Fifth American Peptide Symposium,* Goodman, M. & Meinhofer, J., Eds., Wiley, New York, pp. 296–299.
27. Manjula, G. (1976) Ph.D. thesis, Indian Institute of Science, Bangalore, India.
28. Deber, C. M., Torchia, D. A., Wong, S. C. K. & Blout, E. R. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1825–1829.
29. Madison, V., Deber, C. M. & Blout, E. R. (1977) *J. Am. Chem. Soc.* **99,** 4788–4798.

Received January 12, 1978
Returned for Revision April 21, 1978
Accepted June 30, 1978