Conformational Properties of Disulfide Bridges. Part 3.[1] An Estimation of Structural Flexibility from a Theoretical Study of Diethyl Disulfide

Carl Henrik Görbitz
Department of Chemistry, University of Oslo, N-0315 Oslo, Norway

The conformational flexibility of disulfide bridges is discussed on the basis of theoretical *ab initio* calculations with diethyl disulfide as a model molecule. The equilibrium structure of various disulfide bridge conformations can be stretched or compressed over a surprisingly wide range of $C^{\alpha} \cdots C^{\alpha}$ separations with a comparatively small energy penalty. This substantial flexibility is essential when disulfide bridges adapt to the surrounding peptide chains in the ternary structure of proteins. The total combined range for $C^{\alpha} \cdots C^{\alpha}$ separations in all disulfide conformations is *ca.* 4.5 Å (from 3.4 to 6.9 Å), which is superior to the normal ranges for other hypothetical covalent links between polypeptide chains.

Nature's choice of the disulfide bridge as the main covalent cross-link between polypeptide chains was dictated by two obvious demands: the ease of formation and breakage through the facile non-enzymatic oxidation and reduction of cysteine thiol groups, and conformational properties. Detailed knowledge of the various disulfide bridge structures is essential for studies of all molecules including this group, such as numerous extracellular proteins and biologically active oligopeptides.[2]

In order to understand and explain the conformational properties and preferences of disulfide bridges better, and provide results to be used in the refinement of force field parameters in molecular mechanics programs, a number of theoretical *ab initio* calculations has been carried out. The results are presented in a series of three papers. Paper 1[3] dealt with the C–S rotation in ethyl hydrodisulfide, while Paper 2[1] used diethyl disulfide, $CH_3CH_2SSCH_2CH_3$, as a model molecule.

The conformational freedom of diethyl disulfide can be described in terms of the three torsion angles C–C–S–S (χ_{CS}), C–S–S–C (χ_{SS}) and S–S–C–C (χ_{SC}). Each C–S (or S–C) rotation has *gauche*+ (*G*), *gauche*− (*G'*) and *trans* (*T*) minima, while the S–S rotation has *G* and *G'* '*gauche*' minima at ± *ca.* 90°. This gives a total of 18 minimum structures, but certain pairs involve the same structures rotated by 180° (*e.g. GGT–TGG*) and other pairs are mirror images, such as *GG'G–G'GG'*. Positive and negative disulfide chirality yield non-equivalent disulfide conformations in peptides and proteins, but for diethyl disulfide mirror images are structurally equivalent. This means that the total number of distinct minima is reduced to six: *GGG*, *TGT*, *G'GG'*, *GGT*, *GGG'* and *TGG'* (positive disulfide chirality). The three symmetric minima are shown in Fig. 1. Paper 2 presented results for rotational potentials in diethyl disulfide, and established the stability order for the six energy minima as shown in Fig. 2. There is an even spread of energies, but even the energy of the least stable minimum, *G'GG'*, is comparatively low with an *ab initio* estimate for $\Delta H_{298} = 6.87$ kJ mol^{-1}. The results are in excellent agreement with the frequency with which the corresponding disulfide bridge conformations are observed in peptide[4] and protein structures.[2] Thus, the all-*gauche* 'spiral' global energy minima for diethyl disulfide occur in the disulfide 'left-handed spiral' conformation, the most populated structural family in proteins.[2,5,6]

All the above results refer to the well-defined minimum structures of diethyl disulfide, but when the fragment is incorporated into a protein as a disulfide bridge, the geometry must always adjust to the surrounding ternary structure of the polypeptide chain. The specific modifications needed may involve stretching or compressing the $C^{\alpha} \cdots C^{\alpha}$ distance $r(C^{\alpha} \cdots C^{\alpha})$ {corresponding to r[C(1)···C(4)] in diethyl disulfide}, as well as twisting and bending the structure. The energies involved in the last two distortions are difficult to monitor and estimate, but the stretch/compress part is amenable to theoretical calculations. The present paper gives results from *ab initio* calculations on manipulated diethyl

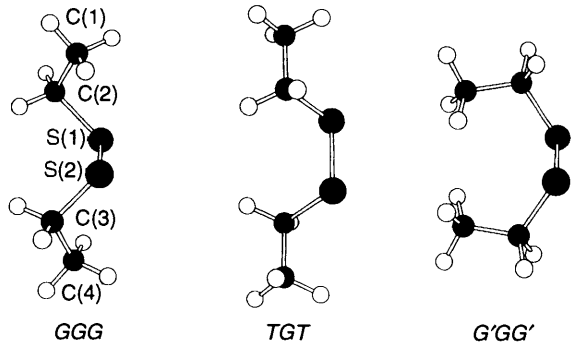

Fig. 1 Three symmetric minimum structures for diethyl disulfide with atomic numbering indicated. There is steric conflict between terminal methyl groups in the *G'GG'* conformation.

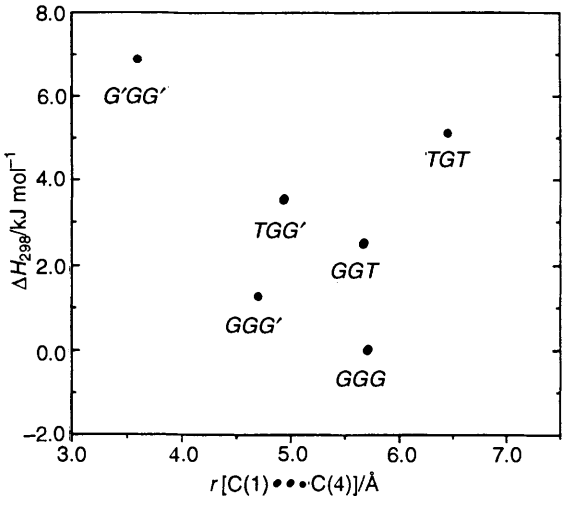

Fig. 2 Scattergram showing conformational energies (ΔH_{298}) for the six diethyl disulfide energy minima and equilibrium C(1) ··· C(4) distances

Table 1 Changes in energy (kJ mol^{-1}) and geometry on stretching and compressing symmetric diethyl disulfide minimum structures[a]

Conformation	r[C(1)···C(4)]	Δ[b]	E[c]	C–C	C–S	S–S	C–C–S	C–S–S	χ_{CS}	χ_{SS}
GGG	6.344	0.500	6.52	1.530	1.835	2.060	114.93	104.98	87.81	97.88
	6.094	0.250	1.62	1.527	1.829	2.057	114.74	104.18	77.79	94.88
	5.844	0.000	0.00	1.523	1.825	2.055	114.52	103.72	70.53	88.86
	5.594	−0.250	1.27	1.521	1.824	2.055	114.67	103.80	64.67	81.46
	5.344	−0.500	4.16	1.520	1.823	2.056	115.02	104.10	59.79	74.30
TGT	7.046	0.500	11.43	1.544	1.843	2.075	110.80	106.52	161.10	106.18
	6.796	0.250	2.21	1.533	1.832	2.059	109.86	104.69	170.46	96.94
	6.546	0.000	0.00	1.526	1.826	2.051	109.37	103.22	177.45	88.28
	6.296	−0.250	1.37	1.523	1.825	2.050	109.48	102.15	−172.81	84.87
	6.046	−0.500	4.54	1.522	1.826	2.050	109.75	101.68	−163.44	82.89
G'GG'	4.246	0.500	4.06	1.522	1.828	2.070	116.54	104.45	−82.96	121.88
	3.996	0.250	1.19	1.522	1.827	2.066	116.50	104.46	−78.17	117.23
	3.746	0.000	0.00	1.522	1.826	2.063	116.41	104.41	−73.45	112.47
	3.496	−0.250	1.95	1.522	1.825	2.059	116.30	104.24	−68.77	107.64
	3.246	−0.500	10.35	1.523	1.825	2.056	116.23	103.86	−63.59	102.95

[a] Distances in Å, angles in degrees. [b] r[C(1)···C(4)] distortion. [c] Energy relative to stationary-point structure.

disulfide conformations and thus illustrates the remarkable flexibility of the structurally equivalent disulfide bridges.

To appreciate better the results from such a study, supplementary calculations have been carried out for two additional molecules as models for alternative protein cross-links, *n*-hexane and *p*-xylene. *p*-Xylene is admittedly rather far-fetched, but is interesting as a rigid structure as distinct from the more flexible open chains. A brief discussion dealing with the properties of still other chemical groups is given at the end of this paper.

Method

All *ab initio* calculations were carried out with the GAUSSIAN90[7] and GAUSSIAN92[8] molecular orbital program systems, and were run on Convex, Cray and IBM computers. The geometry of the six energy minima for diethyl disulfide had previously been optimized at the HF/6-31G* level.[1] The object of the present paper, the study of molecular flexibility, was addressed by performing for each of the three symmetric structures *GGG*, *TGT* and *G'GG'* a set of four calculations in which r[C(1)···C(4)] was reduced or increased relative to the value observed for the respective minimum geometry. The distortions introduced in these calculations were Δ = −0.50, −0.25, 0.25 and 0.50 Å. Non-standard *Z*-matrixes which did not explicitly define the S–S covalent bond or the usual torsion angles were employed in this study. Instead, the C(2)–C(1)···C(4)–C(3) and S(1)–C(2)–C(1)···C(4) torsion angles were included as *Z*-matrix parameters, as well as r[C(1)···C(4)], which could then be fixed at any desired value.

The three C–C–C–C torsion angles in *n*-hexane are described in terms of *gauche*+, *gauche*− and *trans* minima, using the previously defined terminology (*G*, *G'* and *T*). Structures with adjacent *G* and *G'* torsion angles are affected by very unfavourable 1,5 methylene–methylene interactions which cause a substantial rise in relative energy (> 14 kJ mol^{-1}).[9] Leaving out these structures, *n*-hexane has seven distinct energy minima: *GGG*, *GGT*, *GTG*, *GTG'*, *GTT*, *TGT* and *TTT*. Fully relaxed HF/6-31G* molecular geometries were obtained for the symmetric *GGG*, *TGT* and *TTT* structures as well as for *p*-xylene. Each calculation was supplemented by two manipulated refinements with distortions Δ = −0.25 and 0.25 Å for r[C(1)···C(6)]. Additionally, refinements with distortions Δ = −0.50 and 0.50 Å were carried out for the *GGG n*-hexane minimum. The calculations on stretched and compressed structures followed procedures similar to those taken for diethyl disulfide.

Results and Discussion

Relative energies and molecular geometries for diethyl disulfide structures are given in Table 1, while similar data for *n*-hexane and *p*-xylene are given in Table 2.

When a structure corresponding to an energy minimum is either stretched or compressed, an increase in energy relative to the specific minimum will occur. For each of the diethyl disulfide minima this gives rise to an energy potential for r[C(1)···C(4)]. The potentials obtained for the three symmetric minima *GGG*, *TGT* and *G'GG'* are shown in Fig. 3. These curves allow direct assessment of the energy penalty associated with changes to r[C(1)···C(4)] relative to the equilibrium structure of the minimum being studied.

It is immediately clear that the *GGG* minimum structure responds most effortlessly to external pressure. The equilibrium r[C(1)···C(4)] (5.84 Å) can be reduced to 5.29 Å (Δ = −0.55 Å) or increased to 6.29 Å (Δ = 0.45 Å) while keeping the relative energy < 5.0 kJ mol^{-1} above the minimum (Fig. 3). Accordingly, the flexibility within a 5.0 kJ mol^{-1} penalty [hereafter referred to as 'range(5.0)'] is 1.00 Å. The correlation between r[C(1)···C(4)] and covalent bond lengths in Table 1 is evident, but the geometry modifications rendering such substantial flexibility possible occur for the torsion angles. Thus, χ_{CS} and χ_{SC} change from 59.79 to 87.81° and χ_{SS} from 74.30 to 97.88° when r[C(1)···C(4)] is increased from 5.34 to 6.34 Å. The large shifts reflect the fairly low barrier for C–S rotations[1,3] and the relatively flat bottom of the S–S potential well.[10] Incidentally, there is a linear correlation between r[C(1)···C(4)] and the C(2)–C(1)···C(4)–C(3) torsion angle, showing that compression implies unwinding the *GGG* 'spiral', while extending r[C(1)···C(4)] means winding the structure.

The higher-energy *TGT* structure is fairly easily compressed, but stretching the already extended structure is energetically more costly. A sharp energy increase is observed for r[C(1)···C(4)] above *ca.* 6.8 Å. The stretching is strongly reflected in all bond distances and bond angles as well as in the torsion angles. In contrast, only the torsion angles are significantly affected upon manipulation of r[C(1)···C(4)] in the *G'GG'* structure. The exception is the S–S bond length which as usual is correlated with χ_{SS}.[11] *G'GG'* is the only minimum affected by significant steric conflict (Fig. 1), and the close inter-

Table 2 Changes in energy (kJ mol^{-1}) and molecular geometry on stretching and compressing *p*-xylene and symmetric *n*-hexane minimum structures[a]

Conformation	*r*[C(1)···C(6)]	Δ[b]	*E*[c]	C(1)–C(2)	C(2)–C(3)	C(3)–C(4)	C(1)–C(2)–C(3)[d]	C(2)–C(3)–C(4)[e]	C(1)–C(2)–C(3)–C(4)	C(2)–C(3)–C(4)–C(5)
	p-xylene									
	6.095	0.250	25.11	1.580	1.400	1.402	123.20	123.20	180.00	0.00
	5.845	0.000	0.00	1.512	1.389	1.385	121.15	121.15	180.00	0.00
	5.595	−0.250	28.68	1.453	1.379	1.368	118.62	118.62	180.00	0.00
	n-hexane									
GGG	5.390	0.500	2.44	1.532	1.537	1.537	114.45	115.78	71.96	73.85
	5.140	0.250	0.74	1.531	1.536	1.536	114.53	116.23	65.97	65.85
	4.890	0.000	0.00	1.530	1.534	1.536	114.42	116.09	61.47	59.40
	4.640	−0.250	1.19	1.528	1.533	1.535	114.19	115.71	57.63	53.93
	4.390	−0.500	6.05	1.524	1.533	1.536	114.03	115.45	53.44	49.68
TGT	6.340	0.250	6.89	1.547	1.540	1.549	113.85	117.84	169.92	78.07
	6.090	0.000	0.00	1.529	1.531	1.533	112.62	114.88	176.09	66.49
	5.840	−0.250	4.46	1.521	1.531	1.529	112.63	112.85	−167.86	75.02
	6.686	0.250	13.47	1.562	1.554	1.560	117.87	114.54	180.00	180.00
TTT	6.436	0.000	0.00	1.528	1.530	1.530	113.04	113.36	180.00	180.00
	6.186	−0.250	15.13	1.499	1.509	1.499	107.81	112.29	180.00	180.00

[a] Distances in Å, angles in degrees. [b] *r*[C(1)···C(6)] distortion. [c] Energy relative to stationary-point structure. [d] C(1)–C(2)–C(31) and C(1)–C(2)–C(32) for *p*-xylene. [e] C(2)–C(31)–C(41) and C(2)–C(32)–C(42) for *p*-xylene.

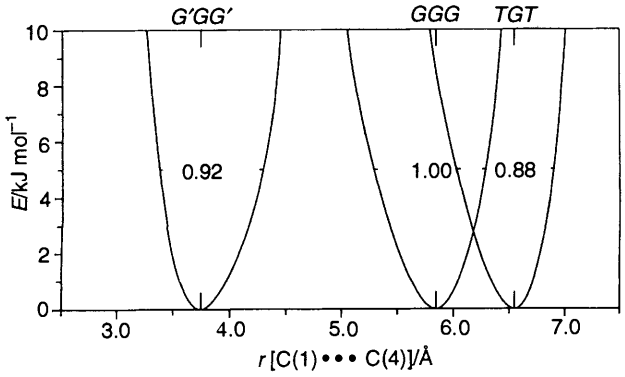

Fig. 3 *Ab initio* energy potentials for strectching and compressing the equilibrium C(1)···C(4) distances, indicated by vertical bars, for the three symmetric diethyl disulfide minima. The width of each potential (in Å) at 5.0 kJ mol^{-1} penalty [range(5.0)] has been indicated.

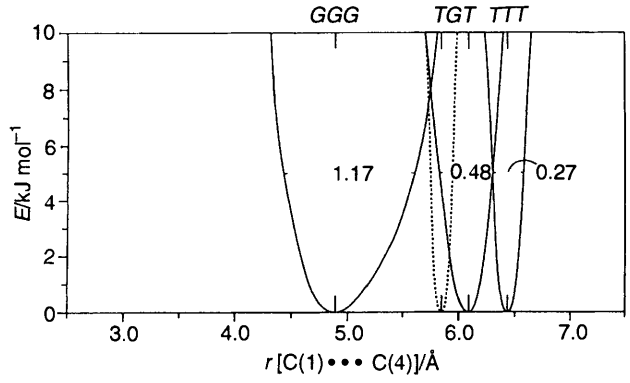

Fig. 4 *Ab initio* energy potentials for stretching and compressing the equilibrium C(1)···C(6) distances for three symmetric *n*-hexane minima (——) and *p*-xylene (– – –). Potential widths have been indicated as in Fig. 3, except for *p*-xylene, which has range(5.0) = 0.20 Å.

ethyl H···H contacts are further aggravated by a reduction of *r*[C(1)···C(4)]. Thus, compression is paralleled by a steep increase in relative energy. The shortest H···H distance in the structure with $\Delta = -0.500$ is only 2.114 Å. On the other hand, increasing *r*[C(1)···C(4)] alleviates bad steric interactions. In the structure with $\Delta = 0.500$ Å χ_{SS} is 121.88°, by far the largest value encountered in this study, but with a relative energy a modest 4.1 kJ mol^{-1} above that of the *G'GG'* minimum.

No calculations were carried out for the *GGT*, *GGG'* and *TGG'* structure, but the rather independent behaviour of the two halves of the diethyl disulfide molecule[1] suggests that the stretching potentials for these minima are approximate blends of the three potentials shown in Fig. 3, with range(5.0) between 0.88 and 1.00 Å.

n-Hexane differs from diethyl disulfide in that the all-*trans* *TTT* extended conformation is the global energy minimum. There is also an all-*gauche* *GGG* minimum, but with a significantly higher energy 8.00 kJ mol^{-1} above *TTT* at the MP3/6-31G*//HF/6-31G* level.[9] Fig. 4 gives potential-energy curves for stretching and compressing both these structures and also for the *TGT* minimum.

The *TTT* structure is very inflexible as any change to *r*[C(1)···C(6)] involves expensive modifications of bond lengths and bond angles. No strain can be absorbed by the torsion angles which are all fixed to 180°. The relative energies of structures with $\Delta = 0.25$ and -0.25 Å are 13.5 and 15.1 kJ mol^{-1}, respectively, *i.e.* about one order of magnitude higher than the values obtained for similar distortions of the diethyl disulfide *GGG* minimum, and comparable to the energies associated with equivalent distortions in *p*-xylene, 25.1 and 28.7 kJ mol^{-1}, respectively.

Not unexpectedly, the *n*-hexane *GGG* structure is considerably more flexible than *TTT*. Its 1.17 Å range(5.0) in fact supersedes those of the diethyl disulfide structures. The *n*-hexane *TGT* minimum, on the other hand, is significantly less flexible than its diethyl disulfide counterpart, with range(5.0) 0.48 and 0.88 Å, respectively. This finding illustrates an important difference between the two molecules: While high flexibility for *n*-hexane resides only with high-energy conformations (*GGG* and probably to some extent *GTG*, *GGT* and *GTG'*), diethyl disulfide not only has the global energy minimum as its most flexible structure, but also displays an overall high degree of flexibility regardless of the specific conformation.

In addition to obvious demands for flexibility, a hypothetical covalent link between peptide chains must also be able to adapt to a suitable range of $C^{\alpha}\cdots C^{\alpha}$ distances without excessive increases in relative conformational energies. The observed range for $r(C^{\alpha}\cdots C^{\alpha})$ in protein disulfide bridges is 3.8–6.8 Å,[1] while the calculated rage(5.0) for diethyl disulfide is *ca.* 3.4–6.9 Å. The importance of having a large 3.5 Å range should not be underestimated. An all-carbon link would have only a 2.1 Å range(5.0) (from 4.5 to 6.6 Å), Fig. 4.

Other Covalent Links.—In this paper only two alternatives to the disulfide bridge are considered. Which other links would be plausible based on chemical properties, and how can our knowledge of the conformational properties be used to eliminate some of these alternatives? Without embarking on a comprehensive discussion on this topic, a few ideas may be put forward.

In any link C^{α}–X–$(Y)_n$–X–C^{α} substituents other than H on the atoms X should be avoided. Substituents would not only interfere with rotation around the C^{α}–X bond, but also constrain the conformation of both peptide main chains at the link.[12] Either property is clearly undesirable for a versatile link of general use. The quest for flexibility would also make the use of multiple bonds or ring systems, as in *p*-xylene, unlikely.

We may then focus on the length of the bridging group. The three molecules described in this paper have chains of six atoms, giving a total length of five covalent bonds. A four-bond all-carbon link would have a range(5.0) from *ca.* 3.4 to 5.3 Å. The upper limit could be increased somewhat by the introduction of heteroatoms like S and P, but given that the experimental distribution for $r(C^{\alpha}\cdots C^{\alpha})$ in protein disulfide bridges has a peak around 5.7 Å,[2] it is obvious that four-bond links would restrict the relative positions of the two main-chain segments involved in the bridge much more than a five bond link. Very short links (≤four covalent bonds) have, like ring systems and multiple bonds, been introduced in model peptides in order to force β-turns.[13]

Hypothetical five-bond links with a central C–C bond, such as C^{α}–O–CH_2–CH_2–O–C^{α}, would have conformational properties rather similar to the all-carbon bridge, shown above to be inferior to those of the disulfide bridge. This is true also for a C^{α}–CH_2–CH(OH)–O–CH_2–C^{α} hemiacetal link. A C^{α}–CH_2–O–O–CH_2–C^{α} link, on the other hand, would have conformational properties rather similar to the disulfide bridge, but the high chemical reactivity of the peroxide group makes it less suitable for incorporation into a protein structure. Similarly, the C^{α}–CH_2–NH–NH–CH_2–C^{α} link is a poor candidate. Hydrazine and its derivatives are thermodynamically unstable, and at the same time the group possesses presumably unfavorable basic properties.

The number of possible six-bond or longer covalent links is essentially unlimited, but most alternatives can clearly be discarded as they lack general flexibility and/or suitable $C^{\alpha}\cdots C^{\alpha}$ range. It is noted here only that any alternative

based primarily on a carbon-atom skeleton would yield a mixture of conformations dominated by *trans* torsion angles and thus have limited flexibility and no possibility of attaining short $C^{\alpha} \cdots C^{\alpha}$ distances which may sometimes be needed in protein structures.

Other cross-links do occur in nature, however. The best known is lysinonorleucine (C^{α}–CH_2–CH_2–NH–CH_2–CH_2–C^{α}) which occurs in collagen and elastin.[14] More complex covalent links, such as desmosine and isodesmosine in elastin and histidino-dehydrohydroxy merodesmosine in collagen tie together no fewer than four side chains. These groups occur in functionally specialized proteins and cannot be regarded as general alternatives to the disulfide bridge.

Conclusions

The conformational properties of disulfide bridges are more advantageous than those of other hypothetical covalent links in four different respects: they have unique all-*gauche* global energy minima; all minimum structures have entensive flexibility; the comparatively small energy differences between minima mean that all conformations can be used in protein structures, and the flexible minimum structures together span an unparalleled large range of $C^{\alpha} \cdots C^{\alpha}$ distances.

The *GGG* and *G'G'G'* 'spiral' conformations are particularly attractive as they represent the global energy minima and at the same time offer superior flexibility. The very nature of these structures, like short molecular springs, discloses their unique applicability as links between peptide chains in a truly diverse array of protein structures.

Acknowledgements

The author thanks the Norwegian Research Council (NFR) for a generous allocation of Cray computer time.

References

1 Part 2. C. H. Görbitz, *J. Phys. Org. Chem.*, in the press.
2 N. Srinivasan, R. Sowdhamini, C. Ramakrishnan and P. Balaram, *Int. J. Peptide Protein Res.*, 1990, **36**, 147.
3 C. H. Görbitz, *J. Phys. Org. Chem.*, 1993, **6**, 615.
4 C. H. Görbitz, *Acta Chem. Scand.*, 1990, **44**, 584.
5 J. M. Thornton, *J. Mol. Biol.*, 1981, **151**, 261.
6 J. Richardson, *Adv. Protein Chem.*, 1981, **34**, 167.
7 GAUSSIAN90, M. J. Frisch, M. Head-Gordon, G. W. Trucks, J. B. Foresman, H. B. Schlegel, K. Raghavachari, M. Robb, J. S. Binkley, C. Gonzalez, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. L. Martin, L. R. Kahn, J. J. P. Stewart, S. Topiol and J. A. Pople, Gaussian Inc., Pittsburgh PA, 1990.
8 GAUSSIAN92, Revision C, M. J. Frisch, G. W. Trucks, M. Head-Gordon, P. M. W. Gill, M. W. Wong, J. B. Foresman, B. G. Johnsen, H. B. Schlegel, M. A. Robb, E. S. Reploge, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzalez, R. L. Martin, D. J. Fox, D. J. Defrees, J. Baker, J. J. P. Stewart and J. A. Pople, Gaussian Inc., Pittsburgh PA, 1992.
9 K. B. Wiberg and M. A. Murcko, *J. Am. Chem. Soc.*, 1988, **110**, 8029.
10 D. Jiao, M. Barfield, J. E. Combariza and V. J. Hruby, *J. Am. Chem. Soc.*, 1992, **114**, 3639.
11 M. Aida and C. Nagata, *Theor. Chim. Acta*, 1986, **70**, 73.
12 S. J. Leach, G. Némethy and H. A. Scheraga, *Biopolymers*, 1966, **4**, 369.
13 W. C. Ripka, G. V. De Lucca, A. C. Bach, R. S. Pottorf and J. M. Blaney, *Tetrahedron*, 1993, **49**, 3593.
14 D. Voet and J. G. Voet, *Biochemistry*, John Wiley, New York, 1990.

Paper 3/05511F
Received 14*th September* 1993
Accepted 18*th October* 1993