set of four hydrogen atoms which are involved in bifurcated hydrogen bonds in which the oxygen atoms are in common. Conformations of type B (HG_8, HG_{11}, HG_{12}, HG_{14}, HG_{18}, and HG_{23}) also have a pair of intramolecular hydrogen bonds. It does not seem possible to decide from the present experimental data[24] which one of these MEC's or a mixture of them is realized in solution.

The calculations carried out in this paper should be regarded as the first step of a process to determine the most stable conformation of an isolated molecule in solution.[4,25,26] The second and third steps of the process require the determination of the conformational entropy and the introduction of the flexibility of bond lengths and bond angles in the molecule, as was done in the conformational analysis of *cyclo*-(Gly_3Pro_2).[27,28] As was found in ref 28, the relative conformational energy may change when flexibility is introduced. The effect of the introduction of flexibility would not be as strong as observed for *cyclo*-(Gly_3Pro_2) because *cyclo*-hexaglycyl is a bit less crowded than *cyclo*-(Gly_3Pro_2). However, the possibility of interchange of the relative order of stability by introduction of flexibility cannot be excluded. Therefore, the order of stabilities of the MEC's obtained in this paper should be regarded as tentative. However, the conformations themselves should not change too much when flexibility is introduced.

(24) (a) K. D. Kopple, M. Ohnishi, and A. Gō, *J. Amer. Chem. Soc.*, **91**, 4264 (1969); (b) K. D. Kopple, M. Ohnishi, and A. Gō, *Biochemistry*, **8**, 4087 (1969); (c) D. A. Torchia, A. di Corato, S. C. K. Wong, C. M. Deber, and E. R. Blout, *J. Amer. Chem. Soc.*, **94**, 609 (1972); (d) D. A. Torchia, S. C. K. Wong, C. M. Deber, and E. R. Blout, *ibid.*, **94**, 616 (1972); (e) V. T. Ivanov, V. V. Shilin, G. A. Kogan, E. N. Meshcheryakova, L. B. Senyavina, E. S. Efremov, and Yu. A. Ovchinnikov, *Tetrahedron Lett.*, **30**, 2841 (1971); (f) S. L. Portnova, V. V. Shilin, T. A. Balashova, J. Biernat, V. F. Bystrov, V. T. Ivanov, and Yu. A. Ovchinnikov, *ibid.*, **33**, 3085 (1971); (g) G. A. Kogan, V. M. Tulchinskii, V. V. Shilin, and V. T. Ivanov, *Chem. Nat. Prod. (Khim. Prir. Soed.)*, **3**, 361 (1972); (h) A. E. Tonelli and A. I. Brewster, *J. Amer. Chem. Soc.*, **94**, 2851 (1972).

(25) N. Gō and H. A. Scheraga, *J. Chem. Phys.*, **51**, 4751 (1969).
(26) The importance of intermolecular interactions in the crystal of this molecule has already been discussed in this section.
(27) N. Gō, P. N. Lewis, and H. A. Scheraga, *Macromolecules*, **3**, 628 (1970).
(28) G. C.-C. Niu, N. Gō, and H. A. Scheraga, *ibid.*, **6**, 91 (1973).

Investigation of the Cis and Trans Isomers of Sarcosylsarcosine by Nuclear Magnetic Resonance Spectroscopy and Conformational Energy Calculations[1]

J. C. Howard,[2a] F. A. Momany,[2b] R. H. Andreatta, and H. A. Scheraga*

Department of Chemistry, Cornell University, Ithaca, New York 14850.
Received January 24, 1973

ABSTRACT: The proton nuclear magnetic resonance spectrum of the dipeptide sarcosylsarcosine in D_2O shows the presence of both the cis and trans isomeric forms of the peptide group; *i.e.*, the rate of interconversion between the two forms is slow on the nmr time scale so that the resonance peaks of the α-CH_2 and N-CH_3 protons each exhibit different chemical shifts in the two forms. The peaks for the N-terminal N-CH_3 protons are the most easily resolved, and the ratio of the areas for these protons in the cis and trans forms represents the equilibrium constant for the interconversion. The temperature dependence of this equilibrium constant leads to an enthalpy change of 610 ± 90 cal/mol, with the trans isomer having the lower enthalpy. Conformational energy calculations are carried out, and the low-energy conformations of both the cis and trans peptide are found. These calculations account for the observed occurrence of only the cis conformation in the crystalline state and also identify the forces responsible for the occurrence of *both* cis and trans conformations in aqueous solution.

In this paper, we consider some of the forces which affect the preference for the cis or trans conformation of the dipeptide sarcosylsarcosine in the crystalline state and in aqueous solution.

In the absence of constraints such as those imposed by the presence of small rings (as in diketopiperazines) the peptide group -CONH- in polypeptides and proteins appears to exist in the planar trans conformation.[3,4] The planarity of the peptide group, as well as the high barrier to rotational interconversion between the cis and trans forms,[5-7] can be directly related to the partial double bond character of the C′-N bond.[3,4,8] However, it is difficult to understand by theoretical methods[8] the origin of the preference for the cis or trans forms, in terms of electronic effects. When the amide nitrogen is methylated, as in proline or sarcosine, the cis ⇄ trans equilibrium is shifted sufficiently toward the cis form, as, *e.g.*, in poly(L-proline)[7] and polysarcosine,[9] so that it is possible to detect the cis form experimentally. On the other hand, according to molecular orbital calculations,[8] the barriers to rotation about the C′-N bond in *N*′-methyl- and *N*,*N*-dimethylacetamide are nearly equivalent, but striking ste-

(1) This work was supported by a research grant (AM-13743) from the National Institute of Arthritis and Metabolic Diseases of the National Institutes of Health, U. S. Public Health Service, and by a Research grant (GB-28469X2) from the National Science Foundation.
(2) (a) NSF predoctoral trainee, 1966-1970; NIH predoctoral trainee 1970-1973. (b) Special Fellow of the National Institute of General Medical Sciences, National Institutes of Health, 1968-1969.
(3) L. Pauling, R. B. Corey, and H. R. Branson, *Proc. Nat. Acad. Sci. U. S.*, **37**, 205 (1951).
(4) J. Donohue, *Proc. Nat. Acad. Sci. U. S.*, **39**, 470 (1953).
(5) The rotational barrier ranges from 13 to 32 kcal per mol for unsubstituted amides,[6] and is about 20 kcal/mol for the prolyl residue in poly(L-proline).[7]
(6) W. E. Stewart and T. H. Siddall, *Chem. Rev.*, **70**, 517 (1970).
(7) I. Z. Steinberg, W. F. Harrington, A. Berger, M. Sela, and E. Katchalski, *J. Amer. Chem. Soc.*, **82**, 5263 (1960).
(8) J. F. Yan, F. A. Momany, R. Hoffmann, and H. A. Scheraga, *J. Phys. Chem.*, **74**, 420 (1970).
(9) F. A. Bovey, J. J. Ryan, and F. P. Hood, *Macromolecules*, **1**, 305 (1968).

ric effects for rotation of the methyl groups are exhibited when two of the methyl groups are cis to one another across the peptide bond. In order to assess the contribution of the various forces which determine the preference for the cis or trans forms, it was considered desirable to obtain an experimental measure (by means of the nmr technique) of the enthalpy difference between the cis and trans forms of a suitable model compound. Conformational energy calculations would then be carried out on the same compound to explain the observed cis:trans ratio both in the crystal and in aqueous solution. A dipeptide, *viz.*, sarcosylsarcosine, which contains a methylated amide group, was selected to ensure the presence of reasonable amounts of both isomeric forms, so that their interconversion could be examined experimentally. In order to maintain water solubility, and also to avoid complications from additional peptide bonds, protecting groups were not coupled to the charged end groups.

The energy difference between the cis and trans forms had been estimated earlier from spectroscopic evidence to be greater than 2.0 kcal/mol for *N*-methylformamide in carbon tetrachloride[10] and 1.6 kcal/mol for *N*-methylacetamide in Nujol mull,[11] with the trans isomer having the lower energy. The enthalpy difference was determined from nmr measurements to be ~0.78 kcal/mol for *N*-acetylsarcosine in aqueous solution (at pH ~1.6) with the trans isomer having the lower enthalpy.[12] In contrast to the above compounds which exhibit a preference for the trans conformation, it was found from nmr measurements on *N*-acetyl-L-proline-*N'*,*N'*-diisopropylamide in dioxane[13] that the enthalpy favored the cis conformation by 2.2 kcal/mol; these authors[13] concluded that the preference for cis arose from steric effects of the large isopropyl groups.

Our experimental results for the *dipeptide* sarcosylsarcosine in aqueous solution at neutral pH indicate a preference for the trans conformation, with an enthalpy difference similar to that of *N*-acetylsarcosine. In addition, our conformational energy calculations account for the preferred conformations of sarcosylsarcosine in both the crystalline state and in aqueous solution.

Experimental Section

Materials. Sarcosine (Aldrich Chemical Co., lot no. S140) and *N*-acetylsarcosine (Cyclo Chemical Corp., lot no. F-1729) each exhibited only one spot on thin-layer chromatography in the solvent system *n*-butyl alcohol (60%)–acetic acid (20%)–water (20%), and in the same solvents in the ratio 40%:30%:30%, and were used without further purification. D_2O from International Chemical and Nuclear Corp. had an isotopic purity of 99.8%. Diaprep Inc. 20% DCl in D_2O was diluted with D_2O to give 1 *N* DCl.

Synthesis. Sarcosine anhydride (cyclosarcosylsarcosyl) was prepared according to the second procedure of Sigmund and Liedl.[14] Sarcosylsarcosine was prepared from sarcosine anhydride by treatment with aqueous $Ba(OH)_2$ at 30°, also according to the procedure of Sigmund and Liedl.[14] After three recrystallizations from H_2O–absolute ethanol, the yield was 65% and the mp 189–190°. The amount of sarcosine anhydride contamination was estimated to be 8.5% from nmr measurements and 8.0% by titration.

Nmr Spectra. The nmr spectra were obtained with a Varian Associates A-60A nmr spectrometer with a variable-temperature probe. The temperature was controlled to within ±2° with a Varian V6040 temperature controller. Precision bore Wilmad Co. nmr sample tubes were used, and chemical shifts were measured with respect to the internal standard sodium 2,2-dimethyl-2-silapentane-5-sulfonate at ~1.3% w/v (~0.06 *M*) concentration. The radiofrequency field was adjusted to avoid saturation effects in all spectra, and the spectra were resolved, assuming Lorentzian line shape, with the aid of a Du Pont Model 310 curve resolver. The sample concentrations ranged from 0.2 to 10%, w/v (0.01–0.50 *M*), and the apparent pD was measured with a Radiometer Model PHM 4c pH meter using a Model GK2302B combined glass and saturated calomel electrode.

Since prolonged heating (~5 hr) of sarcosylsarcosine in water led to precipitation, the nmr experiments were carried out with freshly prepared samples, allowing only enough time to attain thermal equilibrium and obtain the spectrum at each temperature. It will be shown that, under these conditions, the cis ⇌ trans interconversion appears to be reversible (with no precipitation occurring during the nmr experiments).

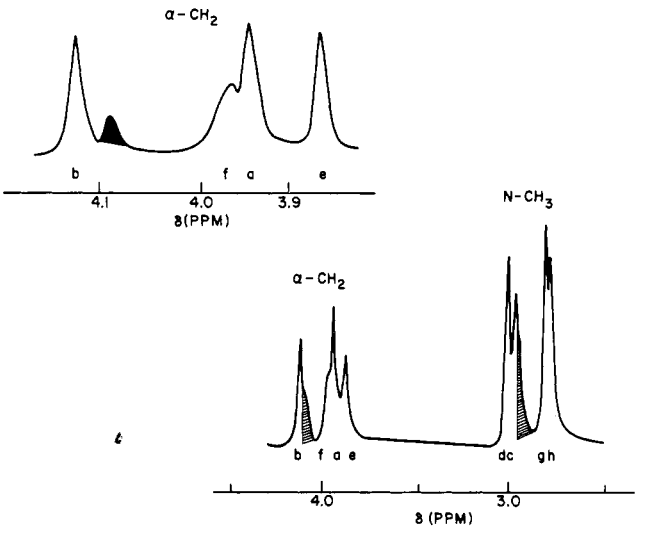

Figure 1. Nmr spectrum of sarcosylsarcosine at 60 MHz and 37°, for a 5% (w/v), ~0.25 *M*, solution of apparent pD of 6.6 (measured at 25°). Chemical shifts were measured with respect to the internal standard (sodium 2,2-dimethyl-2-silapentanesulfonate). The shaded regions arise from a *cyclo*-sarcosylsarcosyl impurity. The letters a, b, c, etc., correspond to the sarcosylsarcosine protons discussed in Table I and Figure 2. The α-CH_2 region is expanded (by running at lower sweep width) in the upper left-hand corner.

Experimental Results

The nmr spectrum of sarcosylsarcosine at 37° (shown in Figure 1) exhibits two regions of interest, those of the *N*-methyl and α-CH_2 protons with chemical shifts δ of 2.7–3.2 and 3.8–4.2 ppm, respectively. If a mixture of two conformations did not exist, *i.e.*, if the sample consisted of only one form, then the spectrum would be expected to consist of four separate peaks, one for each of the two α-CH_2 groups (each peak corresponding to two protons), and one for each of the two N-CH_3 groups (each peak corresponding to three protons). In fact, as seen in Figure 1 (see especially the insert in the upper left-hand corner), the spectrum consists of eight peaks.

The presence of the eight peaks is attributed to the existence of both cis and trans forms in equilibrium. The rate of interconversion is slow on the nmr time scale at 37°, so that the spectrum of each isomer consists of a discrete set of sharp peaks.

The peaks of Figure 1, designated by the letters a, b, c, etc., correspond to the protons indicated in Figure 2. The assignments were made with the aid of nmr measurements on the related compounds listed in Table I. From the areas in the spectra of sarcosine (A) and *cyclo*-sarcosylsarcosyl (B), it is clear that the α and β protons arise from the α-CH_2 and N-CH_3 groups, respectively. Both sets of protons in the cyclic compound are deshielded (relative to those of sarcosine) because they are bonded to an

(10) R. M. Badger and H. Rubalcava, *Proc. Nat. Acad. Sci. U. S.*, **40**, 12 (1954).
(11) T. Miyazawa, *J. Mol. Spectrosc.*, **4**, 155 (1960).
(12) J. T. Gerig, *Biopolymers*, **10**, 2435 (1971).
(13) V. Madison and J. Schellman, *Biopolymers*, **9**, 511 (1970).
(14) F. Sigmund and E. Liedl, *Hoppe-Seyler's Z. Physiol. Chem.*, **202**, 268 (1931).

Table I
Chemical Shifts of Sarcosylsarcosine and Some Related Compounds[a]

Compound	Solvent	Chemical Shift (ppm from DDS[b])									
(A) Sarcosine	D_2O (pD 5.1)					2.73 (β)				3.60 (α)	
	1 *N* DCl					2.82 (β)				4.00 (α)	
(B) *cyclo*-Sarcosyl-sarcosyl	D_2O (pD 4.0)					2.97 (β)				4.11 (α)	
	1 *N* DCl					2.98 (β)				4.12 (α)	
(C) *N*-Acetyl-sarcosine	D_2O (pD 4.9)	2.04 (γ)	2.14 (γ)			2.92 (β)	3.08 (β)			3.92 (α)	3.98 (α)
	D_2O (pD 2.1)	2.07 (γ)	2.17 (γ)			2.95 (β)	3.13 (β)			4.15 (α)	4.27 (α)
	1 *N* DCl	2.11 (γ)	2.19 (γ)			2.98 (β)	3.15 (β)			4.18 (α)	4.32 (α)
(D & E) Sarcosyl-sarcosine	D_2O (pD 6.6)			2.78 (h)	2.80 (g)	2.96 (c)	3.00 (d)	3.88 (e)	3.98 (f)	3.95 (a)	4.12 (b)

[a] The letters a, b, c, etc., and α, β, γ correspond to the protons indicated in Figure 2. [b] Sodium 2,2-dimethyl-2-silapentanesulfonate internal reference.

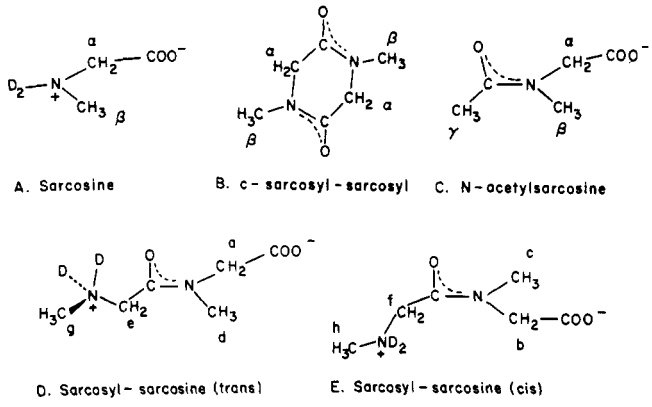

Figure 2. Chemical structures of compounds investigated here. The letters a, b, c, etc., for *cis*- and *trans*-sarcosylsarcosine correspond to the peaks in Figure 1; the letters α, β, and γ correspond to the proton designations of the other compounds listed in Table I.

amide group, and appear at lower field in the spectrum of the cyclic structure. The effect of pH on the chemical shifts of sarcosine is evident in Table I.

From a comparison of *N*-acetylsarcosine (C) at pD 2.1 and *cyclo*-sarcosylsarcosyl (B), the γ protons (from the acetyl CH_3 group) of the former are assigned to the peaks at 2.07 and 2.17 ppm since those at 2.95, 3.13 and 4.15, 4.27 correspond to the N-CH_3 and α-CH_2 groups, respectively, which are bonded to an amide group in both compounds. Again, the effect of pH on the chemical shifts of *N*-acetylsarcosine can be seen from the data of Table I. The doubling of the peaks in *N*-acetylsarcosine presumably arises from the presence of both cis and trans forms, in contrast to only cis in the cyclic compound. Since the α and β resonances in *N*-acetylsarcosine have similar chemical shifts to those of *cyclo*-sarcosylsarcosyl, and these peaks in both these compounds are downfield from the corresponding one of sarcosine, the deshielding effect in going from sarcosine to either *N*-acetylsarcosine or *cyclo*-sarcosylsarcosyl must arise from the loss of the free N-terminal group (rather than the loss of the free C-terminal carboxyl group) in forming the cyclic compound.

In sarcosylsarcosine (D and E), the a, b, e, f protons are assigned to the α-CH_2 groups and the c, d, g, h protons to the N-CH_3 groups on the basis of peak areas, and similarities of the chemical shifts to those of sarcosine, *cyclo*-sarcosylsarcosyl, and *N*-acetylsarcosine. Since the protons are deshielded in going from a compound with a free N-terminal amino group to a peptide amide group (as discussed above in the comparison of the spectra of sarcosine and *cyclo*-sarcosylsarcosyl), the lower field peaks a, b, and c, d are assigned to the C-terminal residue and the higher field peaks e, f and g, h are assigned to the N-terminal residue (see Table I and Figure 2).

The remaining assignment, that to the cis or trans forms in sarcosylsarcosine, is made as follows. Using a nuclear Overhauser effect, Anet and Bourn[15] found that the resonances of the protons of the methyl group, which is cis to the carbonyl oxygen in *N,N*-dimethylformamide, lie at higher fields that those of the other methyl group. In sarcosylsarcosine, the N-CH_3 group of the C-terminal residue is cis to the carbonyl oxygen in the cis isomer (see Figure 2E). Therefore, the higher field peak (c) for the C-terminal N-CH_3 pair is assigned to the cis isomer, and the lower field peak (d) to the trans isomer. It was possible to complete the assignments of the peaks a, b, e, f, g, and h, as shown in Table I and Figure 2 because the ratio of the areas of the peaks corresponding to the C-terminal N-CH_3 protons must be the same as those of the N-terminal N-CH_3 protons, the N-terminal α-CH_2 protons, and the C-terminal α-CH_2 protons, since these ratios represent the ratio of the cis and trans forms. These assignments are in agreement with those of other investigators[9,12,16] on *N*-acetylsarcosine methyl ester,[9] *N*-acetylsarcosine,[12] and proline oligomers.[16]

The small peaks at 4.09 and 2.96 ppm in Figure 1 arise from a small amount of *cyclo*-sarcosylsarcosyl impurity, as stated in the Synthesis section. This assignment is based on a correlation of the temperature dependences of the chemical shifts of these "impurity" peaks with those of the peaks of *cyclo*-sarcosylsarcosyl; also addition of pure *cyclo*-sarcosylsarcosyl enhanced the areas of these peaks.

It is conceivable that the doubling of the peaks in Figure 1 might be due to other conformations, rather than to cis and trans isomers, and this point will be considered in the Discussion section.

The nmr spectra of sarcosylsarcosine were obtained at a series of temperatures and at several concentrations. While all peaks were observed to alter in chemical shift with a change in temperature, only those corresponding to the N-terminal N-CH_3 protons (peaks g and h) were used to obtain trans:cis ratios; however, all areas were mea-

(15) F. A. L. Anet and A. J. R. Bourn, *J. Amer. Chem. Soc.*, 87, 5250 (1965).
(16) C. M. Deber, F. A. Bovey, J. P. Carver, and E. R. Blout, *J. Amer. Chem. Soc.*, 92, 6191 (1970).

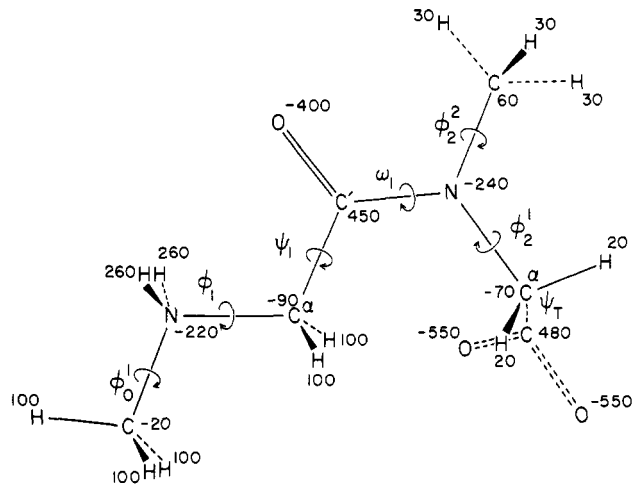

Figure 3. Partial charges and conformation of the "almost"-extended cis conformation (A of Table II) of charged sarcosylsarcosine. The charges must be divided by 1000 to obtain electronic units.

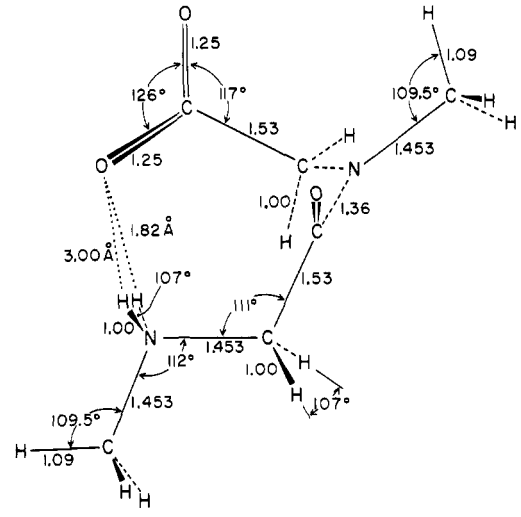

Figure 4. Bond lengths, bond angles, and conformation of the cyclic cis conformation of charged sarcosylsarcosine (B of Table II). The bond lengths and bond angles not given are the same as those shown in Figure 5, except for the angles around the peptide nitrogen, which are discussed in the text. The hydrogens on the charged nitrogen lie in the perpendicular plane bisecting the angle defined by the atoms $C(CH_3)$, N, and C^{α} and equally above and below the plane defined by these atoms.

sured at one temperature (37°) in order to make the peak assignments. The other peaks were not used at other temperatures because of the overlap of the resonances from the N-terminal α-CH_2 cis protons (f) with those of the C-terminal α-CH_2 trans protons (a), and the overlap of the resonances from the cyclic contaminant with those of the C-terminal α-CH_2 cis (b) and C-terminal N-CH_3 cis (c) protons (see Figure 1). The trans:cis ratio and chemical shifts of the N-terminal N-CH_3 protons of Figure 1 were independent of concentration in the range of 0.2–10%, indicating that only the monomeric species is present.

Conformational Energy Calculations

The nomenclature and conventions used here are those adopted by an IUPAC-IUB Commission.[17] The partial charges, bond angles, and bond lengths are shown in Figures 3, 4, and 5. The geometry was deduced from the structures of *cyclo*-tetrasarcosyl,[18] which contains both cis and trans N-methylated peptides, and tris(sarcosine)calcium chloride[19] which contains the charged ($CH_3NH_2^+CH_2$–) group (the same as the end group in Figure 4). The bond angles around the nitrogen of the peptide bond found for N-substituted amides by other workers[20,21] were also taken into consideration. In every case,[18,20,21] a shift in orientation of the substituent groups attached to the peptide nitrogen was observed. Thus, in our calculations, the angles 119 and 124° around the peptide nitrogen (as shown in the trans form of Figure 5) are interchanged when the cis forms are studied. The bond lengths and bond angles adopted here were later verified by the X-ray diffraction results of Stezowski and Hughes[22] on sarcosylsarcosine. The torsional potentials, and nonbonded and hydrogen-bond energies for both the charged and uncharged molecule, and the charges for the uncharged molecule are given elsewhere;[23] the geometry

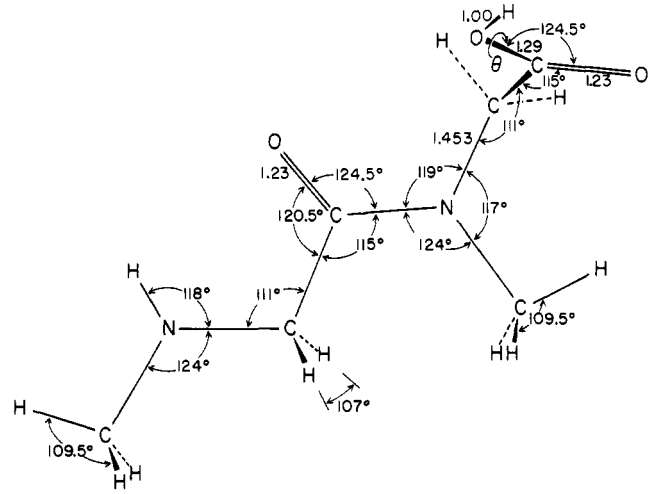

Figure 5. Bond lengths, bond angles and conformation of the trans uncharged "almost"-extended conformation of sarcosylsarcosine (A′ of Table II).

of the uncharged molecule is given in Figure 5 and elsewhere.[23] All bond angles and bond lengths were maintained fixed, except for τ(C′-N-C$^{\alpha}$) and τ(C′-N-C(CH_3)) which are reversed when going from the trans to the cis peptide conformation, as mentioned above. All dihedral angles for rotation about single bonds were allowed to vary except θ [the dihedral angle of the carboxyl end group (Figure 5)], which was maintained at 180° since initial calculations indicated that this angle did not vary significantly during energy minimization.

The total energy of all forms of sarcosylsarcosine was taken as a sum of the contributions from all intramolecular torsional, nonbonded, hydrogen-bonding, and electrostatic energy terms. An effective dielectric constant[24] of 2 was included in the electrostatic term. The conformational energy was minimized with respect to all dihedral angles except θ by a variation[25] of the Fletcher-Powell gra-

(17) IUPAC-IUB Commission on Biochemical Nomenclature, *Biochemistry*, **9**, 3471 (1970).
(18) P. Groth, *Acta Chem. Scand.*, **24**, 780 (1970).
(19) T. Ashida, S. Bando, and M. Kakudo, *Acta Crystallogr., Sect. B*, **28**, 1560 (1972).
(20) P. Ganis, G. Avitabile, E. Benedetti, C. Pedone, and M. Goodman, *Proc. Nat. Acad. Sci. U. S.*, **67**, 426 (1970).
(21) W. R. Krigbaum, R. J. Roe, and J. D. Woods, *Acta Crystallogr., Sect. B.*, **24**, 1304 (1968).
(22) J. J. Stezowski and R. E. Hughes, private communication; submitted for publication.
(23) F. A. Momany, R. F. McGuire, and H. A. Scheraga, *J. Phys. Chem.*, to be submitted.
(24) R. F. McGuire, F. A. Momany, and H. A. Scheraga, *J. Phys. Chem.*, **76**, 375 (1972).
(25) A. W. Burgess, private communication.

Table II
Low-Energy[a] Conformations of Sarcosylsarcosine

Conformation		ϕ_0^1	ϕ_1	ψ_1	ω_1	ϕ_2^1	ϕ_2^2	ψ_T	θ	E_{el}	E_{nb+HB}	E_{tor}	E_{tot}	ΔE
		Degrees								kcal/mol				
Charged														
Cis[b]	A	61.0	180.0	−169.3	−6.2	82.7	86.3	−177.2		19.29	2.38	0.24	21.91	0
Cis cyclic[c]	B	67.4	−168.4	−101.5	37.9	68.8	6.6	91.6		16.23	0.79	7.56	24.58	2.67
Trans cyclic	C	67.9	−160.7	73.9	−139.1	88.9	35.6	166.4		17.06	−0.73	8.58	24.91	3.00
Cis opp[d]	D	67.0	−153.7	93.7	−15.3	86.9	84.7	−140.6		20.02	3.81	1.32	25.15	3.24
Trans	E	60.2	−179.7	172.8	171.5	−96.6	65.6	−158.5		24.73	1.35	0.44	26.52	4.61
Trans opp[d]	F	60.8	−169.5	100.2	165.3	−105.1	76.9	−156.7		26.58	3.57	1.28	31.43	9.53
Uncharged														
Trans[e]	A′	61.2	179.0	179.1	−179.2	−105.0	58.8	76.8	180.0	16.51	−0.21	0.00	16.30	0
Cis	B′	54.9	171.0	−169.3	−8.3	90.9	79.6	−114.4	180.0	16.53	0.70	0.41	17.64	1.34
Trans opp[d]	C′	67.4	−159.7	79.1	−177.8	−106.9	63.8	77.8	180.0	16.53	1.30	0.02	17.85	1.55
Cis opp[d]	D′	63.6	−156.8	87.1	−12.7	92.5	79.1	−119.4	180.0	16.56	1.60	0.98	19.14	2.84
Trans cyclic	E′	61.9	−161.2	82.0	177.8	99.6	68.3	84.0	180.0	18.46	1.22	0.03	19.71	3.41
Cis cyclic	F′	82.9	−53.5	−88.9	14.5	153.5	18.6	106.0	180.0	16.16	3.16	1.14	20.46	4.16
Pseudocharged														
Trans[f]	A″	59.9	−179.5	179.6	−179.5	104.0	60.3	−65.6	180.0	15.98	0.39	0.00	16.37	0.60
Cis[f]	B″	61.9	−176.9	−169.1	−8.6	90.4	80.4	−112.5	180.0	14.16	1.23	0.48	15.77	0

[a] The energies are: E_{el}, electrostatic; E_{nb+HB}, nonbonded plus hydrogen bonding; E_{tor}, torsional; E_{tot}, total; ΔE, the total energy normalized to zero at the lowest energy conformation in each group. E_{tot} of the charged and uncharged molecules cannot be compared. [b] Lowest energy conformation of the *charged* molecule (shown in Figure 3), called "almost"-extended cis in this paper, is equivalent to that found by X-ray diffraction.[22] [c] Conformation shown in Figure 4. [d] Ends of molecule are twisted in opposite directions. [e] Lowest-energy conformation of the *uncharged* molecule (shown in Figure 5), called "almost"-extended trans in this paper. [f] Similar to conformations A and E (charged) with the change in charge discussed in the text.

dient procedure.[26,27] The minimization was carried out until the energy changed by less than 0.01 kcal/cycle, and the first derivatives were very small (*i.e.*, $\sim 10^{-4}$ kcal/radian).

The selection of starting conformations was simplified by the near mirror plane of symmetry and the limitation to two conformations (cis and trans) of the peptide CO–N group. Starting conformations were chosen from low-energy regions obtained by incrementing both ϕ_1 and ψ_1 in 10° intervals (holding all other dihedral angles fixed). This mapping procedure led to a relatively small set of conformations, which could be used as starting points for energy minimization.

Theoretical Results

Starting with low-energy conformations obtained by mapping, the total energy was minimized with respect to all dihedral angles except θ, and the results are shown in Table II. Since every conformation has an energetically equivalent symmetry-related conformational isomer, only half of the minimum-energy conformations are given in Table II.

In the *charged* state, without solvent interactions, the conformation of lowest energy is that with a cis peptide bond, the N-terminal end nearly extended, and the C-terminal carboxyl group nearly perpendicular to the plane of the peptide group (called "almost" extended throughout this paper). This conformation, shown in Figure 3, is nearly equivalent in every dihedral angle to the conformation found by Stezowski and Hughes[22] in an X-ray diffraction study of this molecule. The next lowest energy (*i.e.*, 2.7 kcal/mol) charged conformation is that of the "almost"-cyclic cis form shown in Figure 4. This conformation, which is ~0.3 kcal/mol lower in energy than the "almost"-cyclic trans conformation, is probably a precursor for the formation of *cyclo*-sarcosylsarcosine. A cis conformation with the charged groups on opposite sides of the peptide group is next, and the trans ("almost"-extended) form is found to be ~4.6 kcal/mol higher in energy than the cis ("almost"-extended) form.

The uncharged conformation of lowest energy is found to be the trans ("almost"-extended) conformation, followed by the cis ("almost"-extended) form ~1.3 kcal/mol higher in energy. We see immediately that the trans form of the *neutral* species is energetically preferred, and will be shown to be even more so when the conformational entropy contribution to the total entropy is included. This preference for the trans form would be in accord with expectation if it could be shown that the molecule behaves as if it were nearly neutral in aqueous solution—*i.e.*, if the solvent shields the partial atomic charges either by a dielectric medium effect or by forming a stable hydrate with the polar atoms. Since the effect of solvent was not included in these calculations, a very crude test of the effect of charge on the cis–trans preference was carried out by removing the 0.55 unit of electronic charge from *one* of the oxygens of the C-terminal carboxyl group, and replacing this oxygen by an –OH group with its original partial charges; the partial charges on all remaining atoms (including the N-terminal amino group) were left intact as given in Figure 3. This device (leading to a molecule designated as pseudocharged) would simulate, to some extent, the shielding effect of the charge on the COO^- group by the solvent. Such a molecule still has a preference for the "almost"-extended cis over the "almost"-extended trans conformation (see Table II, conformations B″ and A″), but now by only ~0.6 kcal/mol compared to ~2.7 kcal/mol when the COO^- group retains its full charge. Thus, we see that the progressive removal of the charge on the C-terminal COO^- (simulating a solvent shielding effect) from the fully charged to the pseudocharged to the uncharged molecule leads to a shift from a preference for cis to one for trans.

Although the cis conformation is predicted to be the one of lowest energy in the charged state, in agreement with

(26) M. J. D. Powell, *Computer J.*, 7, 155 (1964).
(27) W. I. Zangwill, *Computer J.*, 10, 293 (1965).

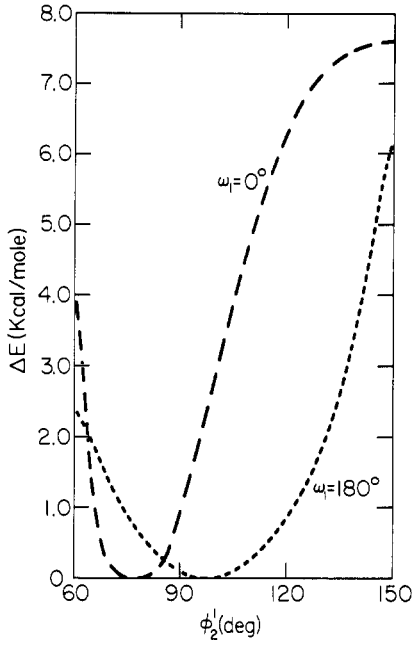

Figure 6. Energy *vs.* ϕ_2^1 for the charged cis (ω_1 = 0°) (A) and trans (ω_1 = 180°) (E) conformations of sarcosylsarcosine. The energy was minimized with respect to ϕ_2^2 and ψ_T at 10° intervals of ϕ_2^1, with $\phi_0^1 = \phi_1 = \psi_1 = 180°$ for $\omega_1 = 0°$ and 180°.

the X-ray observations, the nmr data reported here for sarcosylsarcosine in water indicate a slight preference for the trans conformation. In order to resolve this apparent discrepancy, we first include a contribution to the free energy which arises from the internal rotational (conformational) entropy. We assume that this rotational freedom does not exist in the crystal, but only in solution. The energies of the cis (ω_1 = 0°) and trans (ω_1 = 180°) isomers [forms A and E of Table II, chosen because the N-terminal ends are the same in both (extended) conformations] were computed at 10° increments in ϕ_2^1 by minimizing the total energy with respect to ϕ_2^2 and ψ_T (keeping ϕ_2^1 and ω_1 at the values indicated, and ϕ_0^1, ϕ_1 and ψ_1 at 180°), and the results are shown in Figure 6. It can be seen that ϕ_2^1 is much more restricted in the cis form compared to the trans form. By fitting the curves of Figure 6 with parabolas, and using the expressions for the entropy part of the partition function given by Gō *et al.*[28,29] the entropy contribution to the free energy which arises from the difference in the shapes of the two parabolas was calculated. The result (doubled because there are two equivalent symmetry-related conformations for each cis and trans isomer) showed a preference for the "almost"-extended trans form of 1.2 kcal/mol (at 300°K) over the "almost"-extended cis form. Thus, we conclude that the conformational entropy, as well as the reduction of the COO^- charge, lead to a preference for the trans form in water, which agrees with the nmr results. Although we have computed the solvent effect only very crudely in the computations presented here, it is worth noting that this system is a useful one for testing various models for calculating the effect of solvent interactions; this work will be presented elsewhere.[30] It should be pointed out here that the above contribution to the free energy obtained from the conformational entropy cannot be compared with the *total* entropy of the system because solvent contributions

(28) N. Gō, M. Gō, and H. A. Scheraga, *Proc. Nat. Acad. Sci. U. S.*, **59**, 1030 (1968).
(29) N. Gō and H. A. Scheraga, *J. Chem. Phys.*, **51**, 4751 (1969).
(30) F. A. Momany and H. A. Scheraga, in preparation.

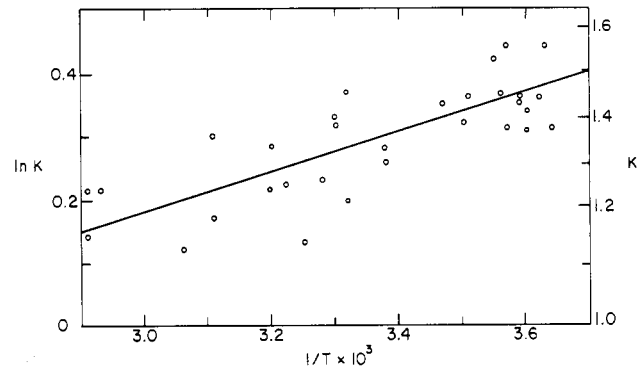

Figure 7. Plot of K and ln K against $1/T$ for the cis–trans interconversion of sarcosylsarcosine. The right-hand scale (for K) is logarithmic.

to the total entropy, which are probably significant, have not been included in these calculations.

Discussion

In the Experimental Results section we noted that it was conceivable that conformations other than cis and trans might cause the doubling of the peaks observed in Figure 1. However, from the conformational energy calculations, we do not find a reasonable probability of occurrence of any other pairs of conformations which would exhibit this doubling; thus, we attribute the presence of eight peaks in the spectrum of Figure 1 to the coexistence of cis and trans peptide forms.

As the temperature was raised, the various sets of peaks, corresponding to the different protons in the cis and trans forms, respectively, approached each other (since the rate of interconversion increases with increasing temperature). However, the various sets of peaks coalesced at different temperatures. This behavior can be understood in terms of the analysis of coalescence and exchange given by Gutowsky and Holm.[31] Below the coalescence temperature of peaks g and h (see Figure 1), the areas under these peaks were determined. The ratios of the areas g:h of the trans to the cis forms, which represent the equilibrium constant K for the reaction

$$\text{cis} \rightleftarrows \text{trans} \qquad (1)$$

were plotted (together with their natural logarithms) against the reciprocal of the absolute temperature T in Figure 7. Since data obtained during heating and cooling fall on the same curve, the cis–trans interconversion appears to be reversible. From the relationship

$$\Delta H^\circ = -R\mathrm{d}(\ln K)/\mathrm{d}(1/T) \qquad (2)$$

a value of $\Delta H^\circ = -610 \pm 90$ cal/mol was determined from the slope of the least-squares straight line in Figure 7.

Having shown that the doubling of the peaks in the nmr spectrum arises from the cis–trans interconversion, we can use the results of the conformational energy calculations to attempt to understand the factors responsible for this equilibrium. In the fully charged state, the calculations show that the charged carboxyl group prefers to be as far from the peptide carbonyl as possible under vacuum. This conformation is the "almost"-extended cis conformation, and is nearly identical to the conformation found in the X-ray structure of Stezowski and Hughes.[22] However, two factors destabilize the cis form and make the trans form

(31) H. S. Gutowsky and C. H. Holm, *J. Amer. Chem. Soc.*, **25**, 1228 (1956).

slightly more stable in water. The first is the shielding of the charge by solvent molecules which surround the carboxyl group, and the second is the conformational entropy.

It is of interest to apply this argument to other N-substituted molecules whose cis–trans equilibrium has been observed. For example, consider the difference between *N*-methyl-*N*-ethylformamide (40% with ethyl cis[6] to carbonyl oxygen) and *N*-methyl-*N*-ethylacetamide (51% with ethyl cis[6] to carbonyl oxygen). Since the attachment of a methyl group on the carbonyl carbon (in going from the formamide to the acetamide compound) would not be expected to change the barrier to rotation about the peptide bond[8] sufficiently to account for this difference, we attribute the difference to a similar entropic contribution, which arises from a steric effect as follows. Assuming that the barrier to rotation about the *N*-(ethyl CH_2) bond is small,[8] and that a CH_2 group is similar to a CH_3 group in steric behavior, the low-energy conformation (when the ethyl is trans to the carbonyl oxygen) would be one in which the ethyl CH_3 is pointed *away* from the acetyl CH_3 group of the acetamide compound to avoid steric interference. This restriction is not as severe when the ethyl is cis to the carbonyl oxygen. Hence, there is a smaller range available to the *N*-(ethyl CH_2) dihedral angle (hence, a smaller entropy) when the ethyl is trans to the carbonyl oxygen. In the formamide compound the restriction on the *N*-(ethyl CH_2) dihedral angle is less severe (than in the acetamide compound) when the ethyl is trans to the carbonyl oxygen, since formamide has an H atom in place of a methyl group; thus, a higher per cent of this trans conformation can occur in the formamide compound. We believe that this entropy effect accounts for the difference in behavior between these formamide and acetamide compounds.

Since the conformational entropy appears to be important for the relative stabilities of cis and trans peptides, we may now extend these arguments to explain the observed cis or trans preference for N-substituted peptides in naturally occurring polypeptides. The (small) enthalpy difference observed here for the cis $\rightleftharpoons$ trans isomerization suggests that a cis peptide bond *could* exist in a polypeptide or protein at the L-prolyl or sarcosyl residue if the trans-preferred enthalpy and conformational entropy were overcome by other energetic and entropic factors. At present, it appears that only those naturally occurring molecules in which a covalent cyclic structure exists (*e.g.*, actinomycin D[32] which has two cis peptide bonds per ring) have sufficient constraints so that the above factors which favor trans may be overcome.

Acknowledgment. We are indebted to Mrs. Jane Derbenwick for help with the synthetic work.

(32) H. M. Sobell, S. C. Jain, T. D. Sakore, and C. E. Nordman, *Nature (London), New Biol.*, 231, 200 (1971).

Conformational Energy Calculations. Thermodynamic Parameters of the Helix–Coil Transition for Poly(L-lysine) in Aqueous Salt Solution[1]

F. Theodore Hesselink,[2a] Tatsuo Ooi,[2b] and Harold A. Scheraga*

Department of Chemistry, Cornell University, Ithaca, New York 14850.
Received March 2, 1973

ABSTRACT: The properties of the helix–coil transition of poly(L-lysine) at various degrees of ionization in aqueous salt solution are computed with the aid of empirical energy functions. The free-energy change in the process is considered to be a sum of a contribution (ΔG_0) from the neutral polymer in the aqueous medium and one (ΔG_e) from the electrostatic interactions between the charged side chains. ΔG_0 is obtained from an earlier calculation for poly(L-alanine) in water, and the effect of the lysine side chain is incorporated. ΔG_e is obtained by computing the partition function for the helix and coil at various degrees of ionization, the coil conformations being generated by a Monte Carlo procedure which yields a characteristic ratio of 8.69 for a neutral poly(L-lysine) chain of 40 residues in a Θ solvent. A Debye–Hückel screening potential was used in the computation of ΔG_e. The expansion factor, α, of the poly(L-lysine) coil due to electrostatic interactions between the side chains was computed and found to obey the relation $\alpha^5 - \alpha^3 \sim 1/C_s$ for salt concentrations $C_s \geq 0.1\ M$; however, the calculated numerical values of α were somewhat lower than experimental ones. The computed value of $(\Delta G_0 + \Delta G_e)$ is zero at a degree of ionization of ~50%, compared to an experimental value of ~35%; the discrepancy may be due to the underestimate of the expansion of the coil. The stability of the poly(L-lysine) helix in 95% methanol at acid pH is attributed to ion-pair formation to the extent of about 35%.

It is well known that the conformation of a homopoly(amino acid) molecule with ionizable side chains depends on the pH of the solution. At high pH a poly(L-lysine) chain appears to be in an α-helical conformation, and at low pH adopts the conformation of a polylectrolyte random coil.[3] Zimm and Rice[4] have presented a statistical mechanical theory for this phenomenon, assuming a conformational transition from an α helix to a locally extended form. Several investigators[5-7] have applied conformational energy calculations to this problem, but these

(1) This work was supported by research grants from the National Science Foundation (No. GB-28469X2), and from the National Institute of General Medical Sciences of the National Institutes of Health, U. S. Public Health Service (No. GM-14312).

(2) (a) On leave from the Van't Hoff Laboratorium der Rijksuniversiteit, Utrecht, Netherlands, 1971–1972. (b) On leave from the Institute for Chemical Research, Kyoto University, Uji Kyoto-Fu, Japan, 1972.

(3) J. Applequist and P. Doty in "Polyamino Acids, Polypeptides and Proteins," M. A. Stahmann, Ed., Univ. Wisconsin Press, Madison, Wis., 1962, p 161.

(4) B. H. Zimm and S. A. Rice, *Mol. Phys.*, 3, 391 (1960).

(5) S. Krimm and J. E. Mark, *Proc. Nat. Acad. Sci. U. S.*, 60, 1122 (1968).

(6) W. A. Hiltner, A. J. Hopfinger, and A. G. Walton, *J. Amer. Chem. Soc.*, 94, 4324 (1972).

(7) A. J. Hopfinger, *Macromolecules*, 4, 731 (1971).