evitably lead to a large characteristic ratio.[15] In the present instance the effect is accentuated because only very extended conformations are available to this molecule on the (ϕ,ψ) energy surface.[4] The only other homopolypeptide that has been reported to have such extended chain dimensions is that of proline.[15,16] Since β-methylvaline would not be expected to undergo cis–trans isomerism across the peptide bond, a polymer composed of this residue would provide a simpler model for the hydrodynamic behavior of a randomly coiled rigid chain than does polyproline.

The severe restrictions on the number of backbone conformations available to β-methylvaline may render the random coil model inappropriate to a polymer of these residues. It is possible that poly-β-methylvaline is unable to adopt a random coil in any solvent and may exist only as a broken rod, an extended helix, or some other extended form which is not determined by the energy maps for a single residue. It should be possible to establish this by deriving experimentally the dimensions of this polymer based on a random coil model and comparing these values with the theoretical results obtained in this study.

Acknowledgments. We are indebted to Dr. A. W. Burgess for many helpful criticisms of this manuscript. We would also like to thank Professor H. A. Scheraga for making the program ECEPP available to us and to Dr. G. Némethy for many useful suggestions. This work was supported by a grant from the Australian Research Grants Committee to S.J.L. Y.P. is a recipient of a Melbourne University Postgraduate Research award.

Supplementary Material Available: Tables of the elements of the average transformation matrices $\langle \mathbf{T} \rangle$ from which the characteristic ratios shown in Tables I and II are derived (Tables III–VI) (4 pages). Ordering information can be found on any current masthead page.

References and Notes

(1) P. J. Flory, "Statistical Mechanics of Chain Molecules", Interscience, New York, N.Y., 1969.
(2) F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *J. Phys. Chem.*, **79**, 2361 (1975).
(3) β-Methylalanine is the trivial name of 2-aminobutanoic acid and β-methylvaline is the trivial name of 2-amino-3-dimethylbutanoic acid. All amino acids and their residues referred to in this paper are in the L configuration.
(4) Y. Paterson and S. J. Leach, *Macromolecules*, preceding paper in this issue.
(5) D. A. Brant and P. J. Flory, *J. Am. Chem. Soc.*, **87**, 2791 (1965).
(6) ECEPP was developed in the Department of Chemistry, Cornell University, and is available from the Quantum Chemistry Program Exchange, Chemistry Department, Room 204, Indiana University, Bloomington, Ind. 47401, as program number QCPE 286.
(7) M. J. D. Powell, *Comput. J.*, **7**, 155 (1964).
(8) W. G. Miller and C. V. Goebel, *Biochemistry*, **7**, 3925 (1968).
(9) D. A. Brant, W. G. Miller, and P. J. Flory, *J. Mol. Biol.*, **23**, 47 (1967).
(10) S. Tanaka and A. Nakajima, *Polym. J.* **2**, 717 (1971).
(11) A. W. Burgess, Y. Paterson, and S. J. Leach, *J. Polym. Sci., Polym. Symp.*, **No. 49**, 75 (1975).
(12) D. A. Brant and P. J. Flory, *J. Am. Chem. Soc.*, **87**, 2788 (1965).
(13) H. Fujita, A. Teramato, T. Yamashita, K. Okita, and S. Ikada, *Biopolymers*, **4**, 781 (1966).
(14) M. Terbojevitch, E. Peggion, A. Cosani, G. D'Este, and E. Scoffone, *Eur. Polym. J.*, **3**, 681 (1967).
(15) P. R. Schimmel and P. J. Flory, *Proc. Natl. Acad. Sci. U.S.A.*, **58**, 52 (1967).
(16) S. Tanaka and H. A. Scheraga, *Macromolecules*, **8**, 623 (1975).

The Random Coil Dimensions of Sequential Copolypeptides Containing N^5-(2-Hydroxyethyl)-L-glutamine. 1. Theoretical Studies

Yvonne Paterson and Sydney J. Leach*

Russell Grimwade School of Biochemistry, University of Melbourne, Parkville, Victoria, 3052, Australia. *Received August 22, 1977*

ABSTRACT: The statistically most probable backbone and side-chain conformations for the N-acetyl-N'-methylamide derivative of N^5-(2-hydroxyethyl)-glutamine (HEG) have been determined using a 12-state model for each of the backbone dihedral angles ϕ and ψ and a 3- or 6-state model for each of the side-chain torsion angles. The most probable backbone conformation is an extended one and the most probable side-chain conformation is one which facilitates the formation of a hydrogen bond between the side-chain hydroxyl group and the backbone carbonyl oxygen atom of the N-acetyl group. The characteristic ratios for the random coil forms of the homopolypeptide, $(\text{HEG})_n$, and its copolypeptides $(\text{HEG-X-X})_n$, where X = gly, ala, or β-methylala, have been derived using this conformational energy surface and are respectively 9.75, 2.27, 8.56, and 7.73.

The identification of conformations of minimum energy of single amino acid residues is finding increasing use in studies on the tertiary structure of peptides.[1-3] The validity of this type of calculation may be assessed by comparing the predicted structure with that derived experimentally, either by x-ray crystallographic[2] or solution[4,5] studies. For polypeptides, interactions between residues which are nearest neighbors are dominant in the computations. This simplification may invalidate the comparisons so that experimental observations fail to agree with the theoretically predicted structure of minimum energy. The theoretically predicted total conformational space available to a residue can also be tested for validity by comparing the characteristic ratio of a randomly coiled homopolypeptide (derived theoretically from the conformational energy surface of the monomer unit over the (ϕ,ψ) plane) with that deduced experimentally from the hydrodynamic behavior of the polymer. Excluding the case of amino acid residues containing a pyrollidine ring, there are few exceptions[6,7] to the close agreement between the theoretically predicted[8-10] and experimentally observed[11-13] unperturbed dimensions of polyamino acids. In such cases intraresidue interactions alone are sufficient to account for the behavior of randomly coiled unperturbed homopolypeptides. This is not the case with polymers containing the proline and hydroxyproline residues where some experimental estimates of the characteristic ratio[14,15] have been markedly lower than

0024-9297/78/2211-0418$01.00/0

the ratios predicted theoretically for poly-L-proline using noncooperative configurational statistics.[16] However, these amino acids exhibit the phenomenon of cis–trans isomerism across the peptide bond and even a small number of cis-peptide units can markedly lower the chain dimensions.[16] Theoretical estimates of residues with the peptide group in the cis conformation must not ignore interactions between neighboring residues and where ω, the dihedral angle about the peptide bond, is allowed to vary continuously, cooperative methods must be used to calculate the characteristic ratio.[17] Theoretical values obtained using such methods for poly-L-proline including a proportion of 5% cis residues in the chain[17] agree closely with experimental estimates[14] on this polymer.

Most polyamino acids have a characteristic ratio similar to that of poly-L-alanine, namely 8 to 9.[8–10] Polyglycine, however, which has no side chains and in which there is, therefore, a much larger conformational space available to the backbone, has a calculated characteristic ratio of about 2. It is impossible to verify this theoretical prediction due to the highly insoluble nature of polyglycine,[18] but glycine residues have been incorporated, in varying proportions up to 46%, into random copolymers containing more soluble amino acid residues. The measured dimensions of these products show varying degrees of agreement with the theoretically predicted copolymer dimensions.[19–21] The types of solubilizing residues used so far have either been soluble only in strong acids[20] or have been electrolytes[19] or have contained a pyrollidine ring.[21] Hydroxyalkylglutamines have been used extensively in copolymers to "solubilize" amino acid residues in order to measure their helix–coil stability constants in water.[22,23] The advantages of using polymers containing this type of residue are that they are (a) likely to be water soluble and thus provide a better model for the behavior of denatured proteins in aqueous solutions than would copolymers in nonaqueous solvents, (b) they are nonionic and charge effects need not, therefore, be considered, and (c) in the case of hydroxyethylglutamine, its polymer shows no tendency to form helical structures in water,[24] and its conformation in this solvent is considered to be typical for a truly statistical random coil.[25]

The present study sets out to compare the effect of side-chain length on the experimentally determined and theoretically predicted characteristic ratio of polypeptides in the unperturbed state. Copolymers were chosen containing hydroxyethylglutamine, together with residues carrying increasingly long side chains as the major component. Three sequential copolymers were selected, with the repeating sequence [hydroxyethylglutamine-X-X], where X = glycine, alanine, or β-methylalanine. The use of sequential rather than random copolymers eliminates the possibility of incorporating blocks of one type of residue. This may occur in random polymerization leading to problems with solubility, localized formation of secondary structures, and difficulties in the theoretical assessment of unperturbed dimensions of such conformationally inhomogeneous polymers.

This paper describes the theoretical prediction of the unperturbed dimensions of these copolymers. Experimentally derived dimensions are currently under investigation.

Methods

Nomenclature and Abbreviations. The amino acid 2-aminobutanoic acid is referred to as β-methylalanine in this paper. It has also been referred to as α-amino-n-butyric acid elsewhere in the literature. N^5-(2-Hydroxyethyl)-L-glutamine is abbreviated to HEG. All other nomenclature used is that recommended by the IUPAC-IUB Commission.[26] All amino acids are considered in the L configuration.

Conformational Energy Calculations. Conformational energy surfaces for the N-acetyl-N'-methylamides of each amino acid residue in the copolymers were calculated using ECEPP[27] (Empirical Conformational Energy Program for Peptides) using a Cyber 73 computer. ECEPP utilizes the empirical potential energy parameters and functions of Momany et al.[28] The methods used to generate the conformational energy surface of glycine, alanine, and β-methylalanine are reported elsewhere.[29]

Table I
Partial Charges, Bond Lengths, and Bond Angles of the Side Chain of Hydroxyethyl-L-glutamine

Side-chain atom	Partial charge, esu	Bond	length, Å	Bond	angle, deg
C^β	−0.030	C^α–C^β	1.53	$C'C^\alpha C^\beta$	111.0
$H(\beta)$	0.020	C^β–H	1.0	$N\,C^\alpha C^\beta$	111.0
C^γ	−0.110	C^β–C^γ	1.53	$C^\alpha C^\beta C^\gamma$	111.0
$H(\gamma)$	0.053	C^γ–H	1.0	$C^\beta C^\gamma C^\delta$	111.0
C^δ	0.465	C^γ–C^δ	1.53	$C^\gamma C^\delta O$	120.5
$O(\delta)$	−0.387	C^δ=O	1.23	$O\,C^\delta N^\epsilon$	124.5
N^ϵ	−0.344	C^δ–N^ϵ	1.325	$C^\gamma C^\delta N^\epsilon$	115.0
$H(\epsilon)$	0.164	N^ϵ–H	1.0	$C^\delta N^\epsilon H$	124.0
C^ζ	0.045	N^ϵ–C^ζ	1.453	$C^\delta N^\epsilon C^\zeta$	121.0
$H(\zeta)$	0.025	C^ζ–H	1.0	$H\,N^\epsilon C^\zeta$	115.0
C^η	0.130	C^ζ–C^η	1.53	$N^\epsilon C^\zeta C^\eta$	111.1
$H(\eta)$	0.02	C^η–H	1.0	$C^\zeta C^\eta O^\theta$	112.0
O^θ	−0.310	C^η–O^θ	1.425	$C^\eta O^\theta H$	110.0
$H(\theta)$	0.170	O^θ–H	1.0	$N\,C^\alpha C'$	110.3

The standard end group data supplied with ECEPP were used for the N'-methyl and N-acetyl end groups. Since direct crystallographic data were not available for the structure of hydroxyethylglutamine the geometry and partial electronic charges of this molecule were adopted from the data available on the most closely related amino acids[28] and model compounds.[30] The values of the partial charges, bond lengths, and bond angles used for the side chain of hydroxyethylglutamine are shown in Table I. The backbone geometry and partial electronic charges used were those described by Momany et al.[28] for all amino acid residues, and the overall charge on the residue was equal to zero.

The conformational space of the backbone of hydroxyethylglutamine was covered using 30° increments in ϕ and ψ within those regions energetically allowed for the alanine residue,[29] 31 grid points being taken in all. The conformations considered for side-chain bond rotations for those bond rotations with torsional energy minima, i.e., χ^1, χ^2, χ^6, and χ^7, were those generated by $\chi^i = \pm 60$ or 180°. The angles of rotation around the peptide bonds and the torsion angle of the side-chain amide bond, χ^4, were kept at 180° so that the substituent atoms to these bonds were kept in a planar trans configuration. A six-state rotational isomeric model was used for the side-chain torsion angles χ^3 and χ^5. The values of the angles were chosen to maximize the distance between the adjoined substitutents to these bonds and were ±30°, ±90°, and ±150°. Thus, the total number of side-chain conformations calculated for each backbone conformation was 2916.

Energy minimizations were performed on selected low-energy conformations of the hydroxyethylglutamine residue, using ECEPP in conjunction with a function minimizing subroutine.[31] Minimization was terminated when the conformational energy changed by less than 0.01 kcal/mol between successive calculations.

Conformational Probability Calculations. The partition function, Z, for the entire conformational space of the molecule is the sum of the Boltzmann distributions of the conformational energy value for each degree of freedom in the molecule.[32]

$$Z = \sum_{\phi}\sum_{\psi}\sum_{\chi^1}\cdots\sum_{\chi^7}\exp[-E(\phi,\psi,\chi^1,\ldots\chi^7)/RT] \qquad (1)$$

where R is the gas constant and T, the temperature, was taken to be 293 K throughout this study. The statistical weight $w(\phi,\psi)$ for each backbone conformation is:

$$w(\phi,\psi) = \sum_{\chi^1} \cdots \sum_{\chi^7} \exp[-E(\chi^1 \cdots \chi^7)/RT] \quad (2)$$

Thus, the conformational probability $P(\phi,\psi) = w(\phi,\psi)/Z$ is the probability for the occurrence of each backbone conformation of the molecule.

The statistical weight $w(\chi^1 \cdots \chi^7)$ for each side-chain conformation is given by:

$$w(\chi^1 \cdots \chi^7) = \sum_{\phi}\sum_{\psi} \exp[-E(\phi,\psi)/RT] \quad (3)$$

and

$$P(\chi^1 \cdots \chi^7) = w(\chi^1 \cdots \chi^7)/Z \quad (4)$$

is the probability for the occurrence of each side-chain conformation of the molecule. The probability for the occurrence of each rotational isomeric state for each variable side-chain torsion angle for the hydroxyethylglutamine residue $P(\chi_k{}^i) = w(\chi_k{}^i)/Z$, where $\chi_k{}^i$ refers to the kth rotational isomeric state of the ith side-chain torsion angles.

$$w(\chi_k{}^i) = \sum_{\phi}\sum_{\psi} \sum_{\substack{\chi^j (j\neq i) \\ n(n\neq k)}} \exp[-E(\phi,\psi,\chi_n{}^j)/RT] \quad (5)$$

Unperturbed Dimensions of Poly(hydroxyethylglutamine). The unperturbed dimensions of homopolypeptides are usually expressed as the characteristic ratio (C_∞)[8] of the mean-square end-to-end distance $\langle r^2\rangle_0$ to the chain length n_p by the square of the virtual bond distance l_p.

$$C_\infty = (\langle r^2\rangle_0/n_p l_p{}^2) = [(\mathbf{E} + \langle\mathbf{T}\rangle)(\mathbf{E} - \langle\mathbf{T}\rangle)^{-1} - (2/n_p)(\langle\mathbf{T}\rangle)(\mathbf{E} - \langle\mathbf{T}\rangle^{n_p})(\mathbf{E} - \langle\mathbf{T}\rangle)^{-2}]_{1,1} \quad (6)$$

where $\mathbf{E}$ is the identity matrix of order 3 and $\langle\mathbf{T}\rangle$ is the statistical average of the set of matrices $\{\mathbf{T}(\phi,\psi)\}$ that transforms the coordinate system of the $(i + 1)$th residue to that of the ith residue over the entire (ϕ,ψ) space available to the backbone.

Where multiple side-chain conformations are available to the molecule, each transformation matrix for each (ϕ,ψ) pair may be statistically weighted over all side-chain rotational isomeric states.[10]

$$\langle\mathbf{T}\rangle = Z^{-1} \sum_{\phi}\sum_{\psi} \mathbf{T}(\phi,\psi) \sum_{i-1}^{7} \sum_{k=1}^{n} \exp[-E(\phi,\psi,\chi_k{}^i)/RT] \quad (7)$$

where n is the number of rotational isomeric states chosen for the side-chain torsion angle χ^i.

The numerical integration interval used for the evaluation of the (ϕ,ψ) space available to this molecule was 30° within the space allowed energetically for alanine. Although a more refined interval would have been more accurate, recent evidence[10] suggests that this interval is adequate for homopolypeptides with "alanine-like" characteristic ratios of 8 ± 2. Since the characteristic ratio of poly(hydroxyethylglutamine) using this integration interval fell into this range, further refinements in the description of the backbone conformational space were not thought to be necessary.

Unperturbed Dimensions of Sequential Copolypeptide (HEG-X-X)$_n$. The characteristic ratio of a sequential copolypeptide $[M_aM_bM_c]_n$ where $n = n_p/3$, is given by[19]

$$\langle r^2\rangle_0/n_p l_p{}^2 = 1 + (2/n_p l_p{}^2)[1,0,0,0,0] \times [\mathbf{G}_{M_b}\mathbf{G}_{M_c}\mathbf{G}_{M_a}]^{1/3(n_p+1)-1} \times \mathbf{G}_{M_b} \begin{bmatrix} 0 \\ l_p \\ 1 \end{bmatrix} \quad (8)$$

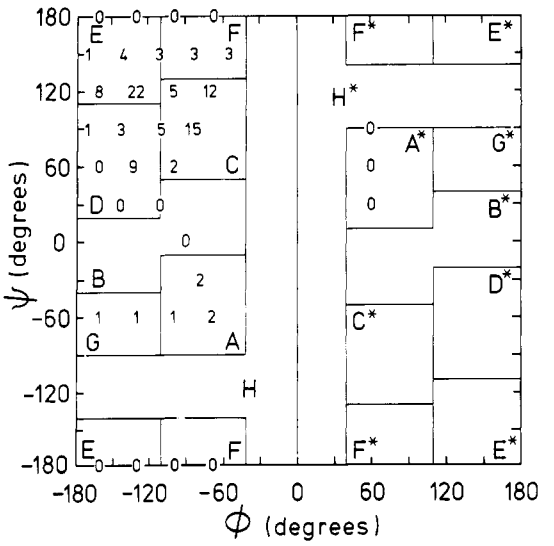

Figure 1. Conformational probability diagram for the N-acetyl-N'-methylamide derivative of hydroxyethylglutamine. The probabilities (%) over all side-chain conformations available to the molecule are plotted as values to the nearest integer. The regions defined by the conformational letter code[27] are marked.

where the vector $\mathbf{l}_p = (l_p,0,0)$ and $\mathbf{G}_{Mx}$ is a composite matrix of order 5 × 5 constructed from the average transformation matrix $\langle\mathbf{T}\rangle_x$ for each type of residue in the polymer, thus

$$\mathbf{G}_{M_x} = \begin{bmatrix} 1 & \mathbf{l}_p{}^T\langle\mathbf{T}\rangle & 0 \\ \mathbf{0} & \langle\mathbf{T}\rangle & \mathbf{l}_p \\ 0 & \mathbf{0} & 1 \end{bmatrix} \quad (9)$$

where $\mathbf{0}$ is the zero matrix of order 3 and $\mathbf{l}_p{}^T$ is the transpose of $\mathbf{l}_p$.

The average transformation matrix $\langle\mathbf{T}\rangle$ for hydroxyethylglutamine was evaluated from eq 7. The matrices for glycine, alanine, and β-methylalanine were evaluated in a previous study[10] at 10° intervals over ϕ and ψ and are shown in Table V. In the case of alanine and β-methylalanine, side-chain angles were allowed some freedom to move from the intrinsic torsional minima in order to minimize the overall conformational energy and the transformation matrix for β-methylalanine was averaged over all rotational isomeric states as described above.

The characteristic ratio for the copolypeptides [HEG-Gly-Gly]$_n$, [HEG-Ala-Ala]$_n$, and [HEG-β-methylAla-β-methylAla]$_n$ were calculated from eq 8 and 9 for increasing values of n until a limiting value was reached.

Results

Energy Calculations. The conformational space available to the backbone of the hydroxyethylglutamine residue is represented in Figure 1 as the probability (%) to the nearest integer over all side-chain conformations available to the molecule within the isomeric state model chosen for this residue. This (ϕ,ψ) map is divided into regions which are considered to contain similar backbone conformations[33] and are labeled according to the conformational letter code devised by Zimmerman et al.[33] The most probable (P = 22%) backbone conformation occurs in region E, which contains extended conformations, and the next most probable in region C, which contains the $C_7{}^{eq}$ ring conformations. The probability of the backbone conformation of HEG falling into either of the regions A and A* (which contain the right- and left-handed α helix) or the region G is less than 5%.

The conformation of lowest energy within each set of 2916

Table II
Lowest Side-Chain Energy Conformations for Each (ϕ,ψ) Pair of the Backbone Conformational Space of the Hydroxyethylglutamine Residue

Dihedral angles, deg[a]								
ϕ	ψ	χ^1	χ^2	χ^3	χ^5	χ^6	χ^7	ΔE[b]
−130	120	−60	−60	−30	90	−60	60	0.0[c]
−131.3	146.2	−67.3	−62.8	−31.9	100.5	−55.1	75.1	−3.02
−130	60	−60	−60	−30	90	−60	60	0.50
−115.4	90.6	−63.9	−62.0	−41.1	103.4	−60.2	81.6	−2.41
−110	150	−60	−60	−30	90	−60	60	1.35
−140	150	−60	−60	−30	90	60	−60	1.71
−160	120	180	60	−90	90	60	−60	1.87
−156.0	135.4	−169.0	56.0	−108.0	85.0	57.5	−84.5	−4.90[d]
−130	−60	−60	−60	−30	90	−60	60	2.02
−140	90	−60	−60	−30	90	−60	60	2.16
−70	120	180	60	−90	90	60	−60	2.23
−74.8	124.8	−175.1	59.3	−103.2	79.1	55.9	−63.9	1.42
−130	180	−60	−60	−30	90	−60	60	2.26
−80	90	−60	−60	−90	150	−60	60	2.45
−114.8	88.1	−67.0	−67.6	−72.2	158.9	−63.4	67.0	−3.20
−110	90	−60	−60	−30	90	60	−60	2.86
−80	150	−60	−60	−90	150	−60	60	2.87
−100	120	180	60	−90	90	60	−60	2.88

[a] ω and χ^4 are set at 180° for all conformations. [b] Relative to $E = -11.651$ for conformation ($\phi = -130$, $\psi = 120$). [c] A perspective drawing of this conformation is shown in Figure 2a. [d] A perspective drawing of this conformation is shown in Figure 2d.

Table III
Probability of Occurrence of Side-Chain Conformations for the Hydroxyethylglutamyl Residue

Dihedral angles, deg							Probability,
χ^1	χ^2	χ^3	χ^4	χ^5	χ^6	χ^7	%
−60	−60	−30	180	90	−60	60	31.8
−60	−60	−30	180	90	60	−60	3.3
180	60	−90	180	90	60	−60	1.8
180	60	−90	180	90	180	180	1.2
					All others		≤1.0

side-chain conformations, for those points on the (ϕ,ψ) plane with a conformational energy within 3 kcal/mol of the lowest point overall, is shown in Table II. Most of the side-chain conformations are not represented in that table and only four side-chain conformations have a probability of occurrence of ≥1%. These are listed in Table III. Gauche rotamers predominate for the side-chain angles which have three side-chain rotational isomeric states. The probability of occurrence of each state for the side-chain angles of the HEG residues is shown in Table IV.

The most probable backbone and side-chain conformation is visualized in Figure 2a. This is also the overall conformation of lowest energy, before minimization, shown in Table II.

Since the purpose of this study was to predict the characteristic ratio of copolymers containing HEG, a total description of the conformational space available to the hydroxyethylglutamine residue was required. Thus, no attempt was made to locate the global minimum energy conformation of the residue or all of the local minima. However, since the probabilities of occurrence of certain side-chain and backbone conformations were very high, we examined each backbone conformation which had a probability of occurrence greater than 5% (i.e., five (ϕ,ψ) pairs in all) and selected those side-chain conformations which gave the lowest energy for the overall conformation. All dihedral angles (excepting ω and χ^4 which were set at 180°) were also minimized. These minimized conformations are shown in italics in Table II below the conformations that were the starting point for each minimization. In fact, the conformation of lowest energy of these five has neither the side-chain nor the backbone conformations of highest probability and may not be the global minimum for this residue. This molecule is visualized in Figure 2b. Both the minimized and starting conformations were checked for the occurrence of hydrogen bonds using the criteria of Zimmerman et al.[33] for hydrogen bond formation. The HEG residue is theoretically capable of forming three types of hydrogen bond: (a) backbone–backbone; (b) side chain–side chain, and (c) side chain–backbone. Each pair of the conformations tested, except the pair of highest energy which did not show any hydrogen bond formation, had only one hydrogen bond, and this was in category (c). In each case the bond was between the side-chain hydroxyl group as hydrogen donor and the backbone carbonyl group of the N-acetyl group as acceptor (see Figures 2a and 2b). For the minimized conformations the H · · · O distance was 1.68 Å.

Unperturbed Chain Dimensions. The average transformation matrix $\langle \mathbf{T} \rangle$ for polyHEG is shown in Table V, together with the $\langle \mathbf{T} \rangle$ matrices for glycine, alanine, and β-methylalanine derived previously[10] and used in this study to evaluate the characteristic ratios of the copolypeptides. The characteristic ratio C_∞ of the randomly coiled homopolypeptide HEG calculated from this matrix by eq 6 is 9.75.

The characteristic ratios for the three copolypeptides (HEG-Gly-Gly)$_n$, (HEG-Ala-Ala)$_n$, and (HEG-β-methyl-Ala-β-methylAla)$_n$ were calculated using eq 8 and 9 for increasing values of n_p (where $n_p = 3n$) are shown in Figure 3. The limiting values C_∞ as $n_p \rightarrow \infty$ are marked on the figure.

Discussion

From the probability of occurrence of backbone conformations (Figure 1) the most likely backbone conformation for the hydroxyethylglutamine residue is an extended one involving a distorted C_5 ring or a distorted C_7^{eq} ring. This is consistent with empirical conformational energy calculations on the naturally occurring amino acids[33] whose lowest conformational energy minima tend to fall in regions C and E with experimental data[34–37] that have demonstrated the existence of the C_5 and C_7^{eq} structures for N-acetyl-N'-methylamide derivatives of amino acid residues. All but four of the side-chain conformations tested for the HEG residue had probabilities of occurrence of less than 1%, but one conformation was overwhelmingly preferred with a probability of occurrence of 32% (see Table III). The strong preference for this side-chain conformation is reflected in the probability of occurrence of each rotational isomeric state for the side-chain di-

Table IV
Probability of Occurrence of Each Rotational Isomeric State for the Side-Chain Dihedral Angles of the Hydroxyethylglutamine Residue

Side-chain angle	Rotational isomeric state, deg (probability, %)					
χ^1	+60 (2.5)	−60 (59.7)	180 (37.8)			
χ^2	+60 (19.9)	−60 (40.6)	180 (39.5)			
χ^3	+30 (8.2)	−30 (43.5)	+90 (15.3)	−90 (33.0)	+150 (0.0)	−150 (0.0)
χ^5	+30 (1.3)	−30 (1.1)	+90 (59.3)	−90 (19.1)	+150 (10.0)	−150 (9.2)
χ^6	+60 (27.5)	−60 (55.0)	180 (17.5)			
χ^7	+60 (52.8)	−60 (24.3)	180 (22.9)			

Table V
⟨T⟩ Elements for the Amino Acid Components of the Copolypeptides [HEG-X-X] Where X = GLY, ALA, and β-Methylalanine

Residue	a_{11}	a_{12}	a_{13}	a_{21}	a_{22}	a_{23}	a_{31}	a_{32}	a_{33}
HEG	0.393	−0.037	0.770	−0.091	−0.699	−0.096	0.698	−0.086	−0.299
GLY	0.368	0.074	0.000	−0.166	−0.451	0.000	0.000	0.000	−0.202
ALA	0.379	−0.023	0.752	−0.126	−0.652	−0.089	0.661	−0.124	−0.271
β-Methyl alanine	0.341	−0.144	0.782	−0.123	−0.591	−0.202	0.614	−0.077	−0.257

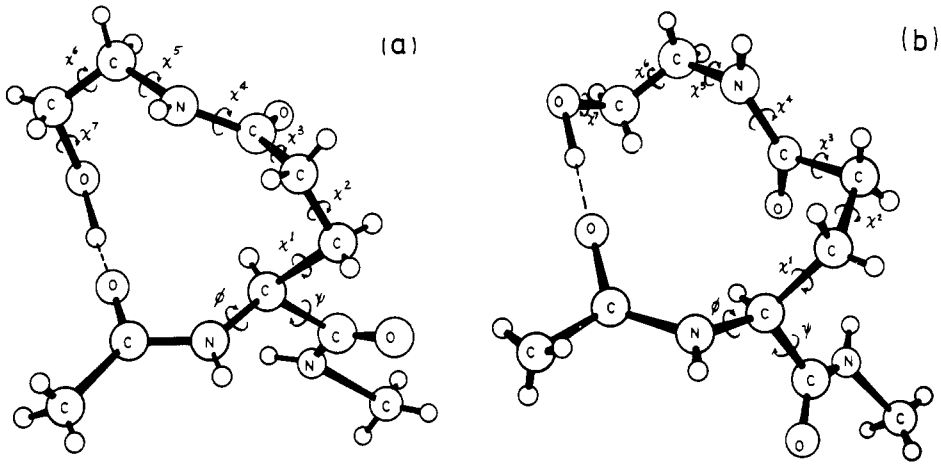

Figure 2. Perspective drawings of two conformations of the *N*-acetyl-*N'*-methylamide derivative of hydroxyethylglutamine: (a) $\phi = -130°$, $\psi = 120°$; (b) $\phi = 156.0°$, $\psi = 135.4°$ (see Table II for values of the other torsion angles). The dashed line represents a hydrogen bond.

hedral angles where, for the angles with torsional minima, the gauche$^-$ rotamers are most preferred for angles χ^1, χ^2, and χ^6 and the gauche$^+$ rotamer is preferred for χ^7, indicating that the side chain tends to bend back on itself toward the backbone. Computations of the interatomic distances between potential hydrogen bond donor and acceptor atoms for selected (ϕ,ψ) pairs with the most preferred side-chain conformation indicated that a strong hydrogen bond between the terminal hydroxyl hydrogen of the side chain and the carbonyl oxygen of the backbone is important in lowering the energy for these conformations. This forms an 11-membered ring. The high probability of this side-chain conformation is almost certainly because it orientates the side chain in a position where this hydrogen bond can be formed for a large number of backbone conformations. Side chain–backbone H bonds have also been shown[33] to play a prominent role in the minimum energy conformations of the *N*-acetyl-*N'*-methylamides of serine and threonine. In these two cases the H bonding produces a seven-membered ring. In proteins such as hemoglobin[38] larger H-bonded rings are observed involving the hydroxyl groups of serine and threonine interacting with main-chain carbonyl groups several residues removed. The seven-membered ring is probably allowed for HEG for certain backbone conformations. It does not appear, however, within side-chain conformations which occur significantly frequently taking into account *all* (ϕ,ψ) backbone values considered. The reason for the observed preference for an 11-membered H-bonded ring may be associated with additional favorable electrostatic and van der Waals energetic contributions.

Since solvent interactions are not incorporated in these calculations, such conformations involving hydrogen bonds may not be important for these residues in solution in good polar solvents where the backbone peptide groups may be solvated. However, these energy calculations have been applied in computing the characteristic ratio of polymers in the unperturbed state at the theta point, where interactions between neighboring chain units are dominant over long-range interactions between remote segments of the molecule and polymer–solvent interactions are negligible.[39] At the theta point conformations with intraresidue H bonds will contribute significantly in the overall spectrum of conformations available to the randomly coiled polymer. It may be that water is not the ideal solvent for deriving the unperturbed dimensions of the copolymers. However, our theoretically predicted value of the characteristic ratio of polyHEG of 9.75 is in excellent agreement with the experimentally derived random-coil dimensions (10 ± 1) of this polymer in water at 30 °C.[40] The agreement between these values indicates that the inclusion

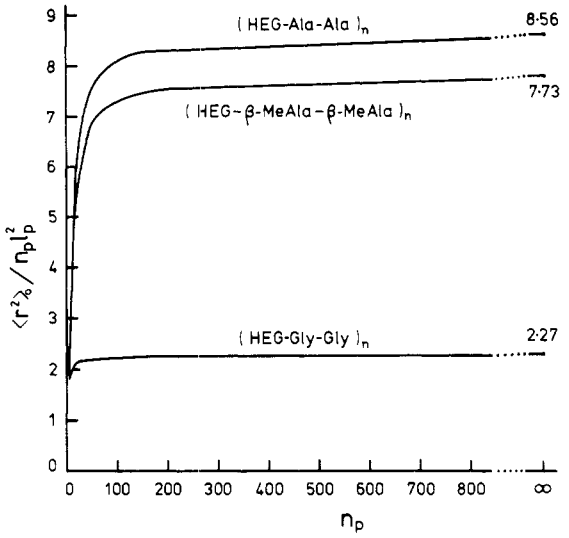

Figure 3. Characteristic ratios as functions of n_p for copolymers containing the hydroxyethylglutamine (HEG) residue. The limiting values of C_∞ for $n_p = \infty$ are marked.

of an intraresidue hydrogen-bonding potential in the overall energy term for the estimation of chain dimensions in the unperturbed state is not inappropriate. The findings of Joubert et al.[25] that the conformation of polyHEG in water is that of a truly statistical random coil also suggest that in this solvent interactions of polyHEG with water are negligible.

A characteristic ratio for polyHEG of 9.75 accords with the hypothesis of Brant and Flory[8] that all amino acid residues with side chains longer than alanine have similar dimensions. Various values of the characteristic ratio for alanine and other amino acid residues have been reported[8,9] but the most recent estimates[10] showed that the characteristic ratios of alanine, β-methylalanine, and valine were very similar (8.06, 7.02, and 9.21, respectively). The value reported here for polyHEG is somewhat higher than these values, but, due to the large number of side-chain conformations for this residue, less rigorous techniques for the energetic weighting were used in this study and less side-chain freedom was allowed for the molecule. The previous work[10] using the same techniques as in this study showed a slightly higher characteristic ratio for alanine of 8.77, thus, a more accurate estimation of the dimensions for polyHEG could well be 10% lower. Since this is within the reported accuracy for experimental estimates of chain dimensions of homopolypeptides,[11–13,40] the calculations on polyHEG should be sufficiently accurate to compare with experimental results.

The chain dimensions of the copolymers containing HEG are only slightly higher than the dimensions of the major components. The characteristic ratio of polyglycine, polyalanine, and poly-β-methylalanine calculated from the ⟨**T**⟩ matrices shown in Table V and reported previously[10] are 2.12, 8.06, and 7.02, respectively; those of the sequential copolypeptides containing two of each of these residues to one of the hydroxyethylglutamine are 2.27, 8.56, and 7.73. Although the β-methylalanine residue has a bulkier side chain and its conformational energy space is thus slightly more restricted[29] than that of alanine, poly-β-methylalanine has a slightly lower characteristic ratio than polyalanine which is reflected in the chain dimensions of the copolymers containing these residues. However, chain dimensions of poly(amino acids) are influenced not so much by the total space available to the residue as by the energetic distributions of conformations within that space. Differences in the methods and parameters used to define the conformational energy space and the methods of weighting the chain dimensions with this space can give markedly different values to the characteristic ratio.[10] The slightly lower dimensions of the β-methylalanine-containing polymers, therefore, may be the result of the approximations involved in using a three-state model for side-chain conformations of this molecule. However, the computed chain dimensions of the homopolymer of β-methylalanine are as close to those of polyalanine as the experimental limit in estimating chain dimensions. The published results for theoretical estimates of the characteristic ratios of sequential copolypeptides have so far been concerned with the effect of *N*-methylamino acids[41] or proline[21,42] on chain dimensions. Our findings are consistent with those of Miller et al.[19] who showed that a small percentage of randomly distributed glycine has a large effect on the chain dimensions of randomly coiled polyalanine. A random copolymer containing 66% glycine would be expected from their results to have a characteristic ratio of 2.4.

Conclusions

The conformational space of the *N*-acetyl-*N'*-methylamide derivative of N^5-(2-hydroxyethyl)-L-glutamine has been described in terms of the statistical percentage probabilities for 30° increments in ϕ and ψ within the areas of the (ϕ,ψ) map allowed for the alanine residue and averaged over all side-chain conformations within the rotational isomeric state model chosen for this molecule. The most probable backbone conformation is an extended one and the most probable side-chain conformation is one in which those gauche rotamers predominate, which facilitate the formation of a hydrogen bond between the terminal side-chain hydroxyl hydrogen atom and the carbonyl oxygen of the *N*-acetyl group.

The unperturbed dimensions of the water soluble nonionic homopolypeptide, polyHEG, and of the sequential copolymers $[HEG\text{-}X\text{-}X]_n$, where X = glycine, alanine, or β-methylalanine, are expressed as the characteristic ratio. The values obtained are polyHEG = 9.75, poly(HEG-Gly-Gly) = 2.27, poly(HEG-Ala-Ala) = 8.56, and poly(HEG-β-methylAla-β-methyl-Ala) = 7.73. These results will be compared with values obtained experimentally in future work.

Acknowledgments. We would like to thank Dr. A. W. Burgess for many stimulating discussions and for suggesting the hydroxyethylglutamine residue as a solubilizing monomer. We must also thank Dr. G. Némethy for helpful advice and Professor H. A. Scheraga for making the program ECEPP available to us. This work was supported by a grant from the Australian Research Grant Committee to S.J.L. Y.P. is a recipient of a Melbourne University Postgraduate Research Award.

References and Notes

(1) A. W. Burgess, F. A. Momany, and H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.*, **70**, 1456 (1973).
(2) S. S. Zimmerman, L. L. Shipman, and H. A. Scheraga, *J. Phys. Chem.*, **81**, 614 (1977).
(3) Y. Isogai, G. Némethy, and H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 414 (1977).
(4) B. P. Roques, C. Garoby-Jaureguiberry, R. Oberlin, M. Anteunis, and A. K. Lala, *Nature (London)*, **262**, 778 (1976).
(5) C. R. Jones, W. A. Gibbons, and V. Garsky, *Nature (London)*, **262**, 779 (1976).
(6) A. W. Burgess, Y. Paterson, and S. J. Leach, *J. Polym. Sci., Polym. Symp.*, **No. 49**, 75 (1975).
(7) R. B. Hawkins and A. Holzer, *Macromolecules*, **5**, 294 (1972).
(8) D. A. Brant and P. J. Flory, *J. Am. Chem. Soc.*, **87**, 2791 (1965).
(9) W. G. Miller and C. V. Goebel, *Biochemistry*, **7**, 3925 (1968).
(10) Y. Paterson and S. J. Leach, *Macromolecules*, preceding paper in this issue.
(11) D. A. Brant and P. J. Flory, *J. Am. Chem. Soc.*, **87**, 2788 (1965).
(12) H. Fujita, A. Teramato, T. Yamashita, K. Okita, and S. Ikada, *Biopolymers*, **4**, 781 (1966).
(13) M. Terbojevitch, E. Peggion, A. Cosani, G. D'este, and E. Scoffone, *Eur. Polym. J.*, **3**, 681 (1967).

(14) W. L. Mattice and L. Mandelkern, *J. Am. Chem. Soc.*, **93**, 1769 (1971).
(15) D. S. Clark and W. L. Mattice, *Macromolecules*, **10**, 369 (1977).
(16) P. R. Schimmel and P. J. Flory, *Proc. Natl. Acad. Sci. U.S.A.*, **58**, 52 (1967).
(17) S. Tanaka and H. A. Scheraga, *Macromolecules*, **8**, 623 (1975).
(18) D. A. Brant, W. G. Miller, and P. J. Flory, *J. Mol. Biol.*, **23**, 47 (1967).
(19) W. G. Miller, D. A. Brant, and P. J. Flory, *J. Mol. Biol.*, **23**, 67 (1967).
(20) M. V. R. Rao, M. Atreyi, S. C. Sinha, and A. R. Ray, *J. Macromol. Sci., Chem.*, **9**, 441 (1975).
(21) W. L. Mattice and L. Mandelkern, *Biochemistry*, **10**, 1934 (1971).
(22) P. H. von Dreele, N. Lotan, V. A. Ananthanarayanan, R. H. Andreatta, D. Poland, and H. A. Scheraga, *Macromolecules*, **4**, 408 (1971).
(23) M. K. Dygert, G. T. Taylor, F. Cardinaux, and H. A. Scheraga, *Macromolecules*, **9**, 794 (1976).
(24) N. Lotan, A. Yaron, and A. Berger, *Biopolymers*, **4**, 365 (1966).
(25) F. J. Joubert, N. Lotan, and H. A. Scheraga, *Biochemistry*, **9**, 2197 (1970).
(26) IUPAC-IUB Commission of Biochemical Nomenclature, *Biochemistry*, **9**, 3471 (1970).
(27) ECEPP was developed in the Department of Chemistry, Cornell University, and is available from the Quantum Chemistry Program Exchange, Chemistry Department, Room 204, Indiana University, Bloomington, Ind. 47401, U.S.A., as program number QCPE 286.
(28) F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *J. Phys. Chem.*, **79**, 2361 (1975).
(29) Y. Paterson and S. J. Leach, *Macromolecules*, preceding paper in this issue.
(30) F. A. Momany, L. M. Curruthers, R. F. McGuire, and H. A. Scheraga, *J. Phys. Chem.*, **78**, 1595 (1974).
(31) M. J. D. Powell, *Comput. J.*, **7**, 155 (1964).
(32) B. Pullman and A. Pullman, *Adv. Protein Chem.*, **28**, 347 (1974).
(33) S. S. Zimmerman, M. S. Pottle, G. Némethy, and H. A. Scheraga, *Macromolecules*, **10**, 1 (1977).
(34) M. Avignon and P. V. Huong, *Biopolymers*, **9**, 427 (1970).
(35) M. T. Cung, M. Marroud, and J. Néel, *Jerusalem Symp. Quantum Chem. Biochem.*, **5**, 69 (1973).
(36) Y. Grenie, M. Avignon, and G. Garrigou-Lagrange, *J. Mol. Struct.*, **24**, 293 (1975).
(37) G. M. Crippen and J. T. Yang, *J. Phys. Chem.*, **78**, 1127 (1974).
(38) M. F. Perutz, The Harvey Lectures (1967–68), Series 63, Academic Press, New York, N.Y., 1969, p 213.
(39) P. J. Flory, "Statistical Mechanics of Chain Molecules", Interscience, New York, N.Y., 1969.
(40) W. L. Mattice and J.-T. Lo, *Macromolecules*, **5**, 734 (1972).
(41) S. Tanaka and A. Nakajima, *Polym. J.*, **2**, 725 (1971).
(42) P. Schimmel and P. J. Flory, *J. Mol. Biol.*, **34**, 105 (1968).

Thermodynamics of Polymerization Involving a Gaseous Monomer-Condensed Polymer Equilibrium

J. C. Favier and J. Leonard*

Département de Chimie, Université Laval, Québec, Canada G1K 7P4.
Received March 7, 1977

ABSTRACT: The thermodynamics of polymerization involving an equilibrium between gaseous monomer and active polymer in the "solid" state is reconsidered taking into account the fact that under equilibrium conditions the polymer is not a pure component. A sizable amount of liquid monomer which forms a nonideal mixture with the polymer is then present and its activity is expressed through the Flory–Huggins treatment. An expression for $\Delta G°_{gc}$, the free enthalpy change upon the conversion of 1 mol of gaseous monomer at 1 atm to 1 base mol of pure amorphous polymer, is written in terms of the equilibrium vapor pressure of the monomer and its volume fraction in the condensed phase. This expression is applied to the published data for the polymerizations of dioxolane and tetrahydrofuran. The results are compared with the value of $\Delta G°_{gc}$, $\Delta H°_{gc}$, and $\Delta S°_{gc}$ computed from the free enthalpy of vaporization of the pure monomers and from ΔG_{lc}, the free enthalpy change upon the conversion of 1 mol of pure liquid monomer into 1 base mol of pure amorphous polymer. A good agreement is found in the case of dioxolane whereas the corrections to be applied to the tetrahydrofuran system are discussed.

In a certain number of equilibrium polymerizations it is found that an equilibrium is established between the monomer in the gaseous state and the active polymer in the condensed state. In the classical thermodynamic treatment of such polymerization systems, the gaseous monomer is considered to be in equilibrium with the pure amorphous polymer, which is simply termed "solid" polymer. The activity of the "solid" polymer is set equal to the activity of the gaseous monomer and this leads to the simple relationship:

$$\Delta G°_{gc} = RT \ln p \tag{1}$$

where $\Delta G°_{gc}$ is the free enthalpy change upon the conversion of 1 mol of gaseous monomer at a pressure of 1 atm (101.325 kPa) into 1 base mol of pure amorphous polymer, p is the vapor pressure of the monomer under equilibrium conditions, and R and T are the gas constant and the temperature, respectively.

However, in many cases the gaseous monomer is in fact in equilibrium with a mixture of monomer and active polymer. Under these conditions the term "solid" polymer becomes inappropriate and the more general term "condensed" polymer or polymer in the condensed state, which applies to polymer in a liquid mixture as well as to pure amorphous polymer, is used in this paper. Since the proportion of monomer in the condensed phase can be large, the thermodynamics of such equilibria ought to be reexamined accordingly.

Theory

In an equilibrium between gaseous monomer and active polymer in the condensed state, the transformation of 1 mol of gaseous monomer at a pressure p into 1 base mol of polymer in a solution made of monomer and polymer takes place. This process may be considered as the sum of the three following steps:

monomer (g, p) → monomer (g, 1 atm)
→ polymer (amorphous) → polymer (solution)

A free enthalpy term can be associated with each of these steps and the sum of these is equal to zero under equilibrium conditions:

$$\Delta G_g + \Delta G°_{gc} + \overline{\Delta G}_p/n = 0 \tag{2}$$

ΔG_g is the free enthalpy change when the pressure of 1 mol of gaseous monomer is changed from p to 1 atm and, assuming an ideal behavior, is expressed by

$$\Delta G_g = -RT \ln p(\text{atm}) \tag{3}$$

0024-9297/78/2211-0424$01.00/0