*MINIMIZATION OF POLYPEPTIDE ENERGY, II. PRELIMINARY STRUCTURES OF OXYTOCIN, VASOPRESSIN, AND AN OCTAPEPTIDE FROM RIBONUCLEASE**

By Kenneth D. Gibson and Harold A. Scheraga†

DEPARTMENT OF CHEMISTRY, CORNELL UNIVERSITY, ITHACA, NEW YORK

Communicated August 3, 1967

In the first paper in this series[1] we presented a somewhat simplified formulation for the energy of a polypeptide in aqueous solution. We also briefly discussed methods for finding local minima and applied one of these to the S-peptide from bovine pancreatic ribonuclease. Starting from five different regular conformations, five different structures were obtained, each of which was essentially a somewhat irregular version of the starting structure, with most of the dihedral angles still quite close to their initial values.

We now extend our expression for the energy to take account of the presence of disulfide bridges, and apply the minimization technique to obtain apparent local energy minima for oxytocin and vasopressin. We have also applied the method to the cyclic octapeptide consisting of residues 65 to 72 from bovine pancreatic ribonuclease, some of whose allowed conformations were computed earlier.[2] The work reported here does not in any sense represent an exhaustive examination of any of these peptides; consequently the results must be regarded as preliminary.

Disulfide Bridge Energy and Minimization.—Six terms were added to the expression for the energy given in reference 1 to account for the presence of disulfide bridges. Two of these terms were torsional potentials for rotation around the C-S bonds; these potentials included nonbonded interactions and were taken to be the same as for bond 3 of the methionine side chain (ref. 1, Table 2). The third term was a torsional potential for rotation about the S-S bond; this was assumed to be twofold, with minima at $\chi = 90°$ and $\chi = 270°$ and a barrier height[3] of 12.0 kcal/mole. The remaining energy terms consisted of one to close the bond gap between the sulfur atoms and two terms to bring the C-S-S bond angles to their correct value. To close the gap we use the simple harmonic function

$$U_{SS} = (1/2)K_{SS}(r_{SS} - r_0)^2, \tag{1}$$

where r_{SS} is the S-S distance and r_0 its equilibrium value[3, 4] (2.1 Å). The force constant K_{SS} had to be large enough to close the gap but not so large as to make this term dominate all others. A value of 1000 kcal/mole was satisfactory. For the C-S-S angle terms the formula

$$U_{CSS} = (1/2)H_{CSS}\{1 - \cos(\theta - \theta_0)\} \tag{2}$$

was used, where θ is the C-S-S bond angle with equilibrium value[4] $\theta_0 = 104°$. A similar argument led to a value of 100 kcal/mole for H_{CSS}.

Only one of these six terms was computed if the distance between the two sulfur atoms was greater than 5.7 Å; this was the term for closing the S-S bond gap (eq. (1)). The remaining five terms were set equal to zero in this case. To preserve continuity, these five terms were multiplied by the factor $g(r_{SS})$ when r_{SS} was less

than 5.7 Å, where $g(r)$ is the function in reference 1, equation (3) (this function is almost exactly equal to 1 when r_{SS} is close to its equilibrium value).

The search for local energy minima was carried out by Davidon's method,[5] as before.[1]

Results.—Oxytocin and vasopressin: Two starting points were used for minimizing the energy of oxytocin and vasopressin. One was the approximate α-helix used before,[1] with $\phi = 120°$, $\psi = 130°$. The other was an approximate β-conformation with $\phi = 60°$, $\psi = 300°$ for all residues except proline, for which $\phi = 120°$ and $\psi = 300°$. All ionizable groups except the side chain of tyrosine were in their charged forms. Minimization was continued until all components of the gradient were less than 0.05 kcal/mole/radian.

Two apparent minima were obtained with both oxytocin and vasopressin, one for each starting point; also, the two structures have quite different dihedral angles. However, there is a striking similarity between the dihedral angles of the backbone of oxytocin in a given minimum and those of the backbone of vasopressin in the corresponding minimum; this can be seen in Table 1 (*upper half*) and Figure 1 for the conformation reached from the β-conformation and in Table 1 (*lower half*) for the conformation obtained when the starting point was an α-helix. In the latter case, there were significant differences in the conformations of the last two residues, but the structure of the ring was virtually the same.

These results suggest that the main factor determining the conformation of the ring in these two peptides is the need to close it, and that other factors, such as the specific nature of the side chains, affect only the details of the structure. To test

TABLE 1

MINIMA OF OXYTOCIN AND VASOPRESSIN[a]

Peptide: / Final Structure:[b] / Residue	Vasopressin ϕ	Vasopressin ψ	Oxytocin ϕ	Oxytocin ψ	Oxytocin (no solvent; $D = 3.0$) ϕ	Oxytocin (no solvent; $D = 3.0$) ψ	Oxytocin (no solvent; $D = 1.0$) ϕ	Oxytocin (no solvent; $D = 1.0$) ψ
1. ½-Cystine	—	327.2	—	315.4	—	306.9	—	288.8
2. Tyrosine	110.0	7.5	120.8	334.4	121.9	339.4	122.6	352.5
3. Isoleucine[c]	106.6	266.1	132.7	276.8	135.2	276.9	132.7	279.0
4. Glutamine	31.8	6.6	22.1	7.5	25.5	6.0	21.3	355.8
5. Asparagine	86.1	239.5	86.0	238.5	84.5	234.4	91.8	266.2
6. ½-Cystine	91.6	325.1	95.2	324.9	93.5	324.8	37.2	324.6
7. Proline	120.1	289.2	120.3	298.9	120.2	297.9	120.5	299.6
8. Leucine[d]	76.9	258.8	87.8	326.4	90.5	321.5	77.9	328.8
9. Glycine[e]	105.0	292.9	119.8	295.6	90.2	327.6	124.6	295.3
Final energy (kcal/mole)	−6.17		10.86		−45.32		−166.55	
1. ½-Cystine	—	214.2	—	286.9	—	237.2	—	259.6
2. Tyrosine	85.8	141.4	43.1	121.6	93.3	168.1	44.2	112.7
3. Isoleucine[c]	137.6	109.5	95.2	104.9	97.5	101.2	90.6	145.6
4. Glutamine	351.1	243.7	18.8	242.0	6.4	231.7	56.0	23.1
5. Asparagine	62.7	112.5	83.8	144.2	82.7	124.3	248.6	97.8
6. ½-Cystine	111.2	126.9	111.3	126.3	110.6	125.8	122.9	120.1
7. Proline	120.4	137.5	120.3	146.2	119.9	154.8	119.3	115.4
8. Leucine[d]	127.2	118.4	258.0	344.8	254.5	96.8	229.1	32.4
9. Glycine[e]	109.1	84.0	125.3	142.0	90.0	70.8	244.2	52.5
Final energy (kcal/mole)	30.96		66.14		−16.35		−150.61	

[a] Starting conformation, upper group: $\phi = 60°$, $\psi = 300°$, except for proline, for which $\phi = 120°$, $\psi = 300°$; ower group, $\phi = 120°$, $\psi = 130°$.
[b] For conventions defining dihedral angles, see ref. 12.
[c] Phenylalanine in vasopressin.
[d] Lysine in vasopressin.
[e] C-terminal amide in both peptides.

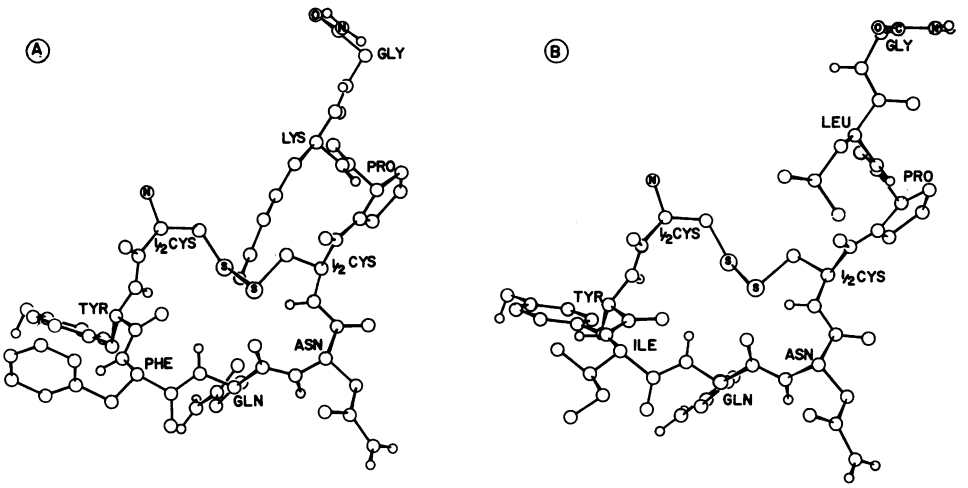

FIG. 1.—Computed structures of vasopressin (*A*) and oxytocin (*B*), obtained by energy minimization, starting from the β-conformation. These are the first two conformations listed in Table 1. These drawings and those in Fig. 2 were obtained from a computer program.[13]

this, the energy was modified by omitting the solvation free energy, and minimizations were carried out with oxytocin from the same starting points as before (Table 1). The final conformations were nearly the same as before, the only significant difference being in the last two residues of the chain. In this calculation, the dielectric constant was equal to 3.0, as it was in the previous calculation which included the solvent energy. There has been some question in the literature about the correct value to take for the dielectric constant when computing the structures of nonsolvated peptides, and values ranging from 1.0 to 4.5 have been used by various authors.[6-10] The partial charges used in the present work were essentially those of Poland and Scheraga,[11] which were computed from theoretical considerations. The correct value of the dielectric constant in calculations with their charge distributions is actually 1.0 if there is no solvent present. The calculations with oxytocin were, therefore, repeated with omission of the solvent energy *and* alteration of the dielectric constant to 1.0. This is presumably a more realistic approximation to the correct energy expression for a nonsolvated peptide; however, we have included the results obtained with $D = 3.0$ partly because there may still be some question about the correct value for D, and partly because they illustrate a point which will be made in the *Discussion*.

The conformations obtained by minimizing the energy of oxytocin in the absence of solvent and with $D = 1.0$ are shown in the last column of Table 1. Starting from the β-conformation led to almost exactly the same final conformation as before even though the total energy at the minimum was now lower by more than 120 kcal/mole. When minimization was started from the α-helix with a dielectric constant of 1.0, a somewhat greater change was observed in the final minimum; this occurred mainly in the region of residue 5 (asparagine), and its effect was to make this residue approach more closely to residue 1 on the other side of the ring. Apart from this, the structure strongly resembles the others in the lower half of Table 1.

Ribonuclease loop: Seventeen starting points were chosen for minimization of the energy of this peptide, namely, the approximate α-helix ($\phi = 120°, \psi = 130°$), the β-conformation ($\phi = 60°, \psi = 300°$), and the 15 sterically allowed conformations found by Nemethy and Scheraga.[2] In the latter, ϕ and ψ were originally constrained to be multiples of 60°. The expression for the energy included solvation in all cases, and the dielectric constant was 3.0. All ionizable side chains were taken to be charged; however, the N- and C-termini were not. Instead, an N-acetyl group was placed at the N-terminal end of the chain and a primary amide group at the C-terminal end, to take account of the fact that the octapeptide is actually in the middle of a long chain.

Each of the 17 starting points led to a different stationary point. In each of these structures, at least two backbone dihedral angles differed by more than 60° from the corresponding angles in other conformations. In most cases, no angle changed by more than 60° in going from the starting conformation to the final stationary point; the exceptions were the minimizations started from the α-helix, from the β-conformation, and from conformations 5, 8, and 12 of reference 2, Table VII. Many of the final structures showed some degree of similarity in the shape of the loop; this is also evident in the backbone angles shown in Table 2. However, the two conformations with lowest energy do not resemble each other particularly (Fig. 2). On the basis of the limited crystallographic data available at present,[14] none of the structures we have computed bears a very close resemblance to the conformation of the octapeptide in crystalline ribonuclease.

Discussion.—The results reported here and in our earlier paper[1] suggest strongly that there is a large number of minimum energy conformations available to a small peptide. This is true both for open-chain and cyclic peptides. Also, the minima are not necessarily all confined to a small set of dihedral angles, but are spread fairly widely over the whole possible range of conformations, with minima whose energies are nearly the same often being quite dissimilar in conformation. Presumably, a protein will show the same behavior, the difference being that one or a few minima will probably have much lower energies than the rest and will thus become the native conformation. Unfortunately, at the present stage of development of computers, it might turn out to be impossible to determine all the energy minima of a protein within a reasonable period of time; hence, it will probably be necessary to find some other method for locating the global minimum of the energy before the problem can be solved. We have made some progress along these lines which will be reported later.

There is one other inference that can be drawn from the results presented here. This concerns the relative importance of the individual terms contributing to the total energy. Omitting the solvent contribution from the expression for the energy of oxytocin (but retaining a value of 3.0 for the dielectric constant) affected the conformation obtained by minimization only slightly. Also, increasing the relative contribution of the electrostatic energy by a factor of 3 (by setting the dielectric constant equal to 1.0 instead of 3.0) gave rise to structures that were not grossly altered, although in this case the change was more marked. These observations confirm that, at least for small cyclic peptides, the energy terms which contribute to the closing of the ring and to the avoidance of steric overlap largely dominate the other terms, and put sharp limits on the possible range of values of the dihedral

TABLE 2

BACKBONE DIHEDRAL ANGLES OF STATIONARY POINT CONFORMATIONS OF RIBONUCLEASE LOOP*

Starting point† for minimization:	1		2		3		4		5		6		7		8		9	
Final structure:‡	φ	ψ	φ	ψ	φ	ψ	φ	ψ	φ	ψ	φ	ψ	φ	ψ	φ	ψ	φ	ψ
Residue																		
1. ½-Cystine	88.4	103.1	90.3	351.0	87.8	240.4	81.8	133.7	76.0	327.1	107.4	280.6	113.6	293.0	87.1	3.4	87.1	8.7
2. Lysine	82.0	127.6	131.1	134.3	133.9	134.2	108.5	128.6	124.8	132.2	101.8	142.1	134.6	128.3	108.1	143.3	91.0	120.4
3. Asparagine	111.4	130.5	88.4	311.2	51.8	114.8	82.3	124.9	48.9	306.8	38.0	280.3	63.0	315.8	54.0	211.2	128.9	127.5
4. Glycine	263.5	235.9	270.8	302.7	259.7	348.9	276.0	135.2	74.1	310.8	84.1	228.7	324.2	41.5	83.5	240.2	247.9	231.7
5. Glutamine	101.6	105.6	131.4	123.9	104.9	112.7	101.9	113.7	46.8	113.4	126.8	127.8	15.6	118.4	119.3	123.4	112.5	119.7
6. Threonine	31.9	128.3	55.4	127.1	72.7	152.7	34.6	126.0	37.3	123.8	51.3	129.7	43.9	123.0	81.4	138.7	34.0	134.8
7. Asparagine	79.2	129.2	88.7	304.3	24.3	123.4	37.1	320.3	36.7	283.9	38.9	313.0	91.0	313.4	38.6	318.0	54.9	123.2
8. ½-Cystine	1.7	129.3	89.8	135.0	54.8	117.6	110.9	131.8	132.4	124.5	133.7	124.0	49.5	120.7	128.8	122.6	32.2	120.3
Final energy (kcal/mole)	−0.36		3.25		20.74		4.09		5.43		7.53		8.04		−7.74		−7.58	

Starting point† for minimization:	10		11		12		13		14		15		α		β	
Final structure:‡	φ	ψ	φ	ψ	φ	ψ	φ	ψ	φ	ψ	φ	ψ	φ	ψ	φ	ψ
Residue																
1. ½-Cystine	74.5	327.0	79.3	232.9	110.0	283.0	131.3	289.6	50.1	339.7	110.2	283.1	118.1	291.5	75.9	325.5
2. Lysine	87.3	121.7	59.7	126.3	75.2	129.3	33.0	111.3	86.6	121.5	60.9	115.5	111.3	124.7	119.3	311.4
3. Asparagine	82.9	292.3	83.3	125.6	45.6	344.5	74.9	254.4	44.5	334.8	75.0	319.0	78.0	198.1	130.9	299.4
4. Glycine	276.6	222.9	271.1	230.9	294.0	37.3	354.3	289.0	336.2	333.8	239.8	10.1	76.6	190.6	342.4	347.1
5. Glutamine	110.6	296.5	112.7	128.9	47.1	122.6	127.8	123.3	107.5	122.2	46.4	320.3	127.9	142.1	125.7	274.9
6. Threonine	45.3	123.3	34.7	122.4	46.7	122.3	75.5	143.7	34.2	122.5	47.1	129.5	128.7	124.9	28.7	309.8
7. Asparagine	66.7	119.1	41.2	359.2	121.0	308.3	90.9	275.2	90.8	311.5	61.4	130.0	92.1	102.3	79.2	246.6
8. ½-Cystine	120.2	290.6	70.5	117.2	42.7	118.8	104.2	151.2	88.2	136.5	116.1	294.5	134.0	121.0	99.9	269.4
Final energy (kcal/mole)	−5.31		−3.77		−5.34		−1.34		1.41		−5.82		16.31		0.24	

* For clarity, only backbone dihedral angles are shown.
† Numbers refer to allowed conformations given in ref. 2, Table VII; α and β refer to approximate α-helix and β-conformation (cf. ref. 1).
‡ For conventions defining dihedral angles, see ref. 12.

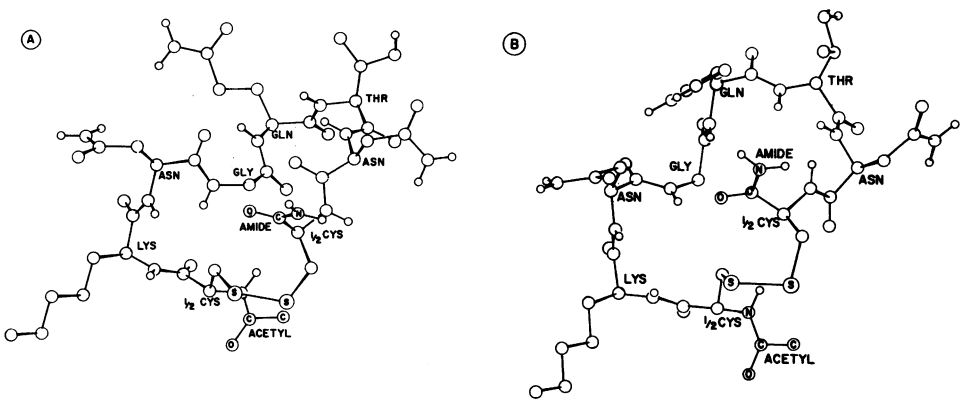

FIG. 2.—Lowest-energy computed structures of the octapeptide loop of bovine ribonuclease. (*A*) Conformation 8; (*B*) conformation 9 of Table 2.

angles at the energy minima. Earlier calculations with hard sphere potentials[2, 15–17] were based on this assumption. However, our results suggest also that the smaller contributions to the total energy, viz. the electrostatic, hydrogen bond, dispersion, torsional, and solvent terms, may tend to act cooperatively rather than to oppose one another in minimum energy conformations of naturally occurring peptides. This would account for the fact that almost the same conformation was found for the oxytocin ring when minimizations were carried out with different relative weights attached to some of these contributions, whereas if the conformation at the minimum were determined by a subtle balance between conflicting energy contributions, a much larger variation should have been observed. Thus, it is likely that computations of polypeptide structure by energy minimization will lead to results that are approximately correct even though there are still many uncertainties and assumptions in the expressions used for the energy. It must be emphasized that these remarks apply only to small peptides, and that, at this stage, we cannot say what effect the various terms in the energy expression will have in determining the stable conformations of proteins. In particular, we would expect the solvent free energy terms to have a marked effect on the orientation of the polar and nonpolar groups in a large polypeptide (presumably favoring the orientation of polar groups on the outside, in contact with water, and the orientation of nonpolar groups on the inside, on the average); indeed, this will provide an experimental test of the validity of our approach.

Our thanks are due to Nancy Moxley for her valuable programing assistance. Nearly all the computations described here were carried out on the CDC-6600 computer at the A.E.C. Computing Center at New York University. The remainder were performed on the CDC-1604 computer at Cornell.

* This work was supported by a research grant (GB-4766) from the National Science Foundation, and by a research grant (GM-14312) from the National Institute of General Medical Sciences of the National Institutes of Health, U.S. Public Health Service.

† To whom requests for reprints should be addressed.

[1] Gibson, K. D., and H. A. Scheraga, these PROCEEDINGS, **58**, 420 (1967).

[2] Nemethy, G., and H. A. Scheraga, *Biopolymers*, **3**, 155 (1965).

[3] Bergson, G., and L. Schotte, *Arkiv Kemi*, **13**, 43 (1958).
[4] Yakel, H. L., Jr., and E. W. Hughes, *Acta Cryst.*, **7**, 291 (1954).
[5] Davidon, W. C., AEC Research and Development Report, ANL-5990 (1959).
[6] Brant, D. A., and P. J. Flory, *J. Am. Chem. Soc.*, **87**, 663, 2791 (1965).
[7] Scott, R. A., and H. A. Scheraga, *J. Chem. Phys.*, **45**, 2091 (1966).
[8] Ooi, T., R. A. Scott, G. Vanderkooi, R. F. Epand, and H. A. Scheraga, *J. Am. Chem. Soc.*, **88**, 5680 (1966).
[9] Brant, D. A., W. G. Miller, and P. J. Flory, *J. Mol. Biol.*, **23**, 47 (1967).
[10] Ooi, T., R. A. Scott, G. Vanderkooi, and H. A. Scheraga, *J. Chem. Phys.*, **46**, 4410 (1967).
[11] Poland, D., and H. A. Scheraga, *Biochemistry*, submitted.
[12] Edsall, J. T., P. J. Flory, J. C. Kendrew, A. M. Liquori, G. Nemethy, G. N. Ramachandran, and H. A. Scheraga, *Biopolymers*, **4**, 121 (1966); *J. Biol. Chem.*, **241**, 1004 (1966); *J. Mol. Biol.*, **15**, 399 (1966).
[13] Johnson, C. K., *ORNL-3794* (Oak Ridge National Laboratory, 1965).
[14] Kartha, G., J. Bello, and D. Harker, *Nature*, **214**, 862 (1967).
[15] Leach, S. J., G. Nemethy, and H. A. Scheraga, *Biopolymers*, **4**, 369 (1966).
[16] *Ibid.*, p. 887.
[17] Scheraga, H. A., S. J. Leach, R. A. Scott, and G. Nemethy, *Disc. Faraday Soc.*, **40**, 268 (1965).