

Quantum Mechanical Approach to the Conformational Analysis of Macromolecules in Ground and Excited States

ROALD HOFFMANN and AKIRA IMAMURA,* *Department of
Chemistry, Cornell University, Ithaca, New York 14850*

Synopsis

Molecular orbital calculations of the extended Hückel type have been used to study the conformations of glycyl and alanyl residues in ground and excited states. The ground-state surfaces show features similar to those obtained with the standard calculational methods in which the total energy is partitioned into components such as torsions, nonbonded and electrostatic interactions. The molecular orbital calculations provide the first independent theoretical check on such calculations. The excited-state surfaces, uniquely available from the molecular orbital calculations, exhibit a better definition and sharpening of potential minima in the sterically allowed regions.

The problem of computing the stability of various conformations of polypeptides has been extensively studied by several laboratories.¹⁻¹⁰ In all of these calculations, the total energy has been partitioned into additive components such as nonbonded interactions, barriers to internal rotations, dipole-dipole and dipole-induced dipole interactions, and hydrogen bonding. The calculations have differed only in the type of energy component admitted and in the parameters defining the strength and distance dependence of each type of interaction. These calculations have been reasonably successful in rationalizing the observed peptide structures and carry still untested promise in the prediction of unknown structures. In this contribution we provide a totally independent check on some of the detailed features of the simplest conformational maps computed by other workers. We are also able to derive the conformational preference of the lower excited electronic state of these molecules.

The method of calculation is the molecular orbital variant known as the extended Hückel theory. This procedure has been widely used in calculations on realistic organic and inorganic molecules,¹¹⁻¹⁴ but since its application to biopolymers is novel we will describe it briefly here.

Molecular orbitals for X_nH_m (where X is some first row atom) are written as linear combinations of the valence orbitals of the molecule. These orbitals are taken as Slater-type exponential functions. There are thus four orbitals per X (2s and three 2p functions) and one (1s) per H.

* Present address: National Cancer Research Institute of Japan, Tsukiji, Chuo-Ku, Tokyo, Japan.

Each molecular orbital is thus a linear combination of $4n + m$ atomic orbitals.

$$\psi_k = \sum_i C_{ki} \phi_i$$

and the coefficients and one-electron energy levels are by the variational theorem determined from the set of secular equations

$$\sum_j [H_{ij} - ES_{ij}]C_j = 0 \quad i = 1, 2, \dots, 4n + m$$

where S_{ij} is the overlap integral $\int \phi_i^* \phi_j d\tau$ and H_{ij} is the Hamiltonian matrix element $\int \phi_i^* H \phi_j d\tau$. All the S_{ij} are retained and computed exactly.

The Hamiltonian is the usual unspecified effective one-electron operator of semiempirical molecular orbital theory. Its diagonal matrix elements are assigned as the valence state ionization potentials, some typical values being H(1s), -13.6 eV; C(2s), -21.4 eV; C(2p), -11.4 eV; N(2s), -26.0 eV; N(2p), -13.4 eV; O(2s), -32.3 eV; O(2p), -14.8 eV.

The off-diagonal matrix elements are taken proportional to the overlap

$$H_{ij} = K[(H_{ii} + H_{jj})/2]S_{ij}$$

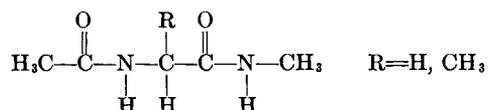
The parameter K has been maintained at the value of 1.75 consistently used by us.

The solution of the eigenvalue problem yields the wave functions and one electron energy levels. The total energy is taken as a simple sum of the occupied one-electron energy levels. The wave functions are subjected to a Mulliken population analysis.¹⁵ This yields gross atomic populations and overlap populations, analogous to the familiar electron densities and bond orders.

This method is clearly an approximation to a true *ab initio* calculation of the electronic structure of a molecule. As such it yields a total energy E for a given molecular geometry. This energy cannot be partitioned in any way into such components as nonbonded repulsions, dipole-dipole interactions, etc. Yet to the extent that it is a good quantum mechanical calculation all such energy components are presumably included. While the energy is not *ipso facto* partitioned into chemically familiar components, the presence of all such components may be easily probed. Torsional barriers of a reasonable magnitude are clearly predicted.¹¹ Nonbonded repulsions are inferred from the manifest rise in energy as two stable molecules, such as methanes, are pushed close to each other. Electronic factors, such as those determining the preferred planarity of ethylene or amide groups, are present. Reasonable parameters for the depth and shape of a hydrogen bond potential function have been obtained.¹² In general, these calculations have proven reliable in predicting ground and excited state equilibrium geometries and the shape of potential surfaces for simple reactions.

The rate-determining step in these calculations is the diagonalization of a matrix of the order of the number of orbitals in the computation, i.e., $4n + m$. These calculations thus cannot hope to be competitive with the classical methods. For a dipeptide residue of the type treated in this paper the computation time is roughly 1 min on an IBM 360/65 computer per point on the potential surface. This is approximately 1000 times slower than a comparable classical calculation. Our computation times, moreover, increase as the third power of the number of orbitals, making multi-dimensional explorations of the potential surface prohibitive. Nevertheless we believe the calculations are significant, both in the independent check they provide on the energy-partitioned methods, and in their capability to explore novel features, such as excited-state potential surfaces.

In this work we have studied the conformational energies of two simple dipeptides, or more precisely the glycylyl and alanyl residues, *N*-acetyl-*N'*-methyl-glycylamide and *N*-acetyl-*N'*-methyl-alaninamide.



The fixed bond distances and angles were those used by Scott and Scheraga.³ The corresponding Cartesian coordinates of the atoms, which form the input to our program were kindly supplied to us by K. D. Gibson. We studied the energy of conformations defined by the dihedral angles ϕ and ψ .¹⁶ (The standard convention has been used.) A grid of points at 30° increments of ϕ and ψ was scrutinized. The results for the ground states of the glycylyl and alanyl residues are presented in Figures 1 and 2 as contour maps of the energy relative to an energy zero at the most stable conformation. The first electronic transition of these residues corresponds in our calculations to an excitation of an electron from an orbital identifiable as a combination of lone pairs on both carbonyl groups (considerably delocalized to nearby atoms) to a π^* -type orbital of both amide groups. Figures 3 and 4 show the energy contours for these glycylyl and alanyl residue excited states. It is assumed that the average energy of the excited configuration computed here will show the same conformational dependence for both singlet and triplet states arising from that configuration.

The general resemblance of our calculated ground-state curves to those of previous hard-sphere^{2,5-7} and extended-interaction⁸ models is good. For the glycylyl residue we obtain a large low-energy basin extending over the range $\phi = 0^\circ$ to 110° and $\psi = -70^\circ$ to $+70^\circ$ (the symmetry related region for $\phi > 180^\circ$ will not be explicitly referred to in our discussions of the glycylyl residue surface). The actual absolute minimum of our calculations lies along the line $\psi = 0^\circ$, but the energies within the basin defined above are all within 1 kcal of the minimum. Our potential energy surface approaches mirror symmetry about the $\psi = 180^\circ$ line more closely than other calculations. The alanyl residue surface differs in minor ways from

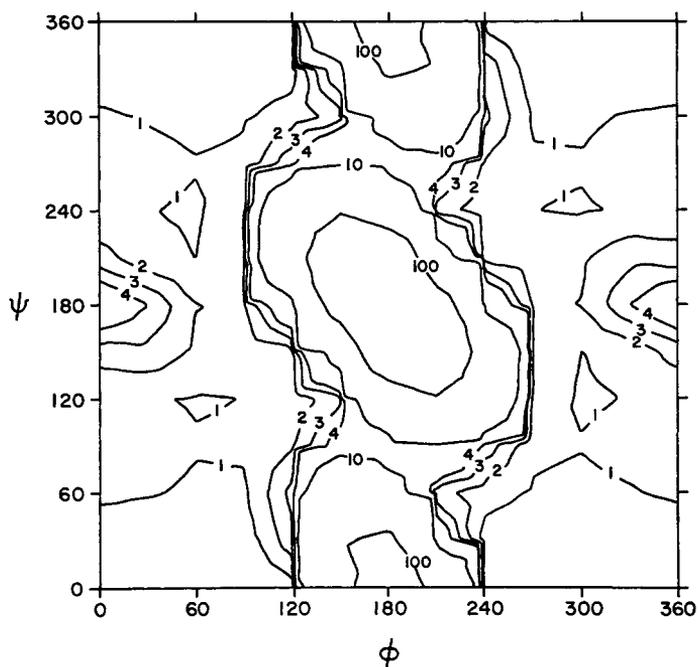


Fig. 1. Calculated energy contours for the ground state of a glycyly residue. The contours are labeled in kilocalories per mole relative to a zero of energy at the absolute minimum of the calculation.

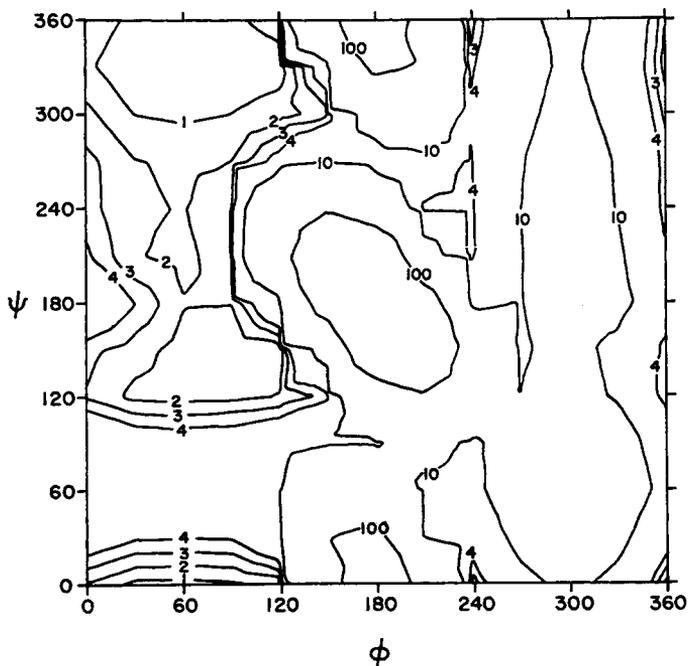


Fig. 2. Calculated energy curves for the ground state of an alanyl residue. The contours are labeled in kilocalories per mole relative to a zero of energy at the absolute minimum of the calculation.

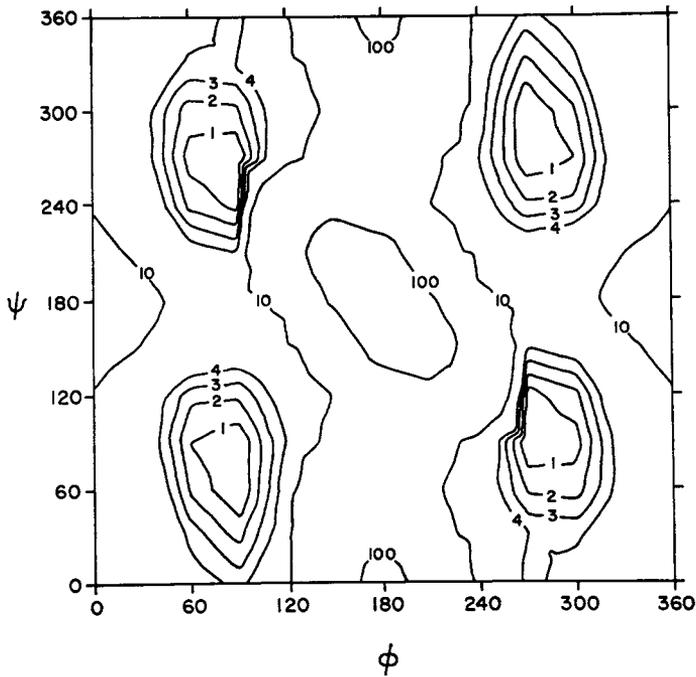


Fig. 3. Calculated energy curves for the first excited state of a glycyI residue. The contours are labeled in kilocalories per mole relative to a zero of energy at the absolute minimum of the excited state calculation.

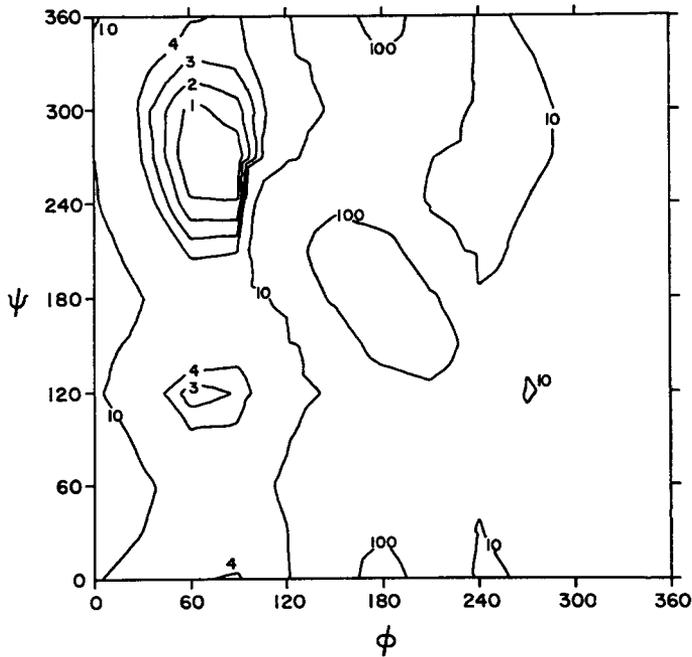


Fig. 4. Calculated energy curves for the first excited state of an alanyl residue. The contours are labeled in kilocalories per mole relative to a zero of energy at the absolute minimum of the excited state calculation.

other calculations.^{3,8} The minimum energy is at greater ψ and the local minimum near $\phi = \psi = 240^\circ$ is more unstable with respect to the absolute minimum than in the calculations of Scott and Scheraga.³ The sterically forbidden regions are remarkably similar to those obtained in other calculations, emphasizing once again the predominant role of nonbonded repulsions in determining those regions.

Brant and Schimmel¹⁷ have compared the experimental distribution of conformations in hen egg-white lysozyme¹⁸ with calculated conformational maps. A similar comparison for myoglobin has been made by Watson.¹⁹ Given all the reservations on the validity of the comparison noted by previous authors,¹⁷ the conformational map presented here for the alanyl residue accommodates at low energy a slightly greater set of experimental points than previous calculations.⁸ The greatest improvement is in the region $150^\circ < \psi < 210^\circ$, $45^\circ < \phi < 90^\circ$, where our calculations give only a low saddle point. It should be noted that other calculations provide a lower energy for this pass when geometrical restraints are further relaxed.^{20,21}

The ground-state forbidden regions would be expected to remain forbidden on the excited-state surfaces. This is indeed so, and the most interesting effects are in the sterically allowed regions. There one finds in our calculations a general sharpening and clearer definition of the potential minima. The excited glycyI residue surface (Fig. 3) has well separated "volcanic" minima at $\phi = 80^\circ$, $\psi = 80^\circ$, 270° . The alanyl residue retains virtually unchanged one of the above minima, at $\phi = 80^\circ$, $\psi = 270^\circ$. The other one is partially affected and is reduced in stability. The excited state calculations must be considered provisional until we examine the effect of distortions corresponding to further degrees of freedom in the excited state. In particular we plan to investigate the possible noncoplanarity of the carbonyl and amine segments of an amide group.

Conclusions

The approximate molecular orbital calculations produce potential surfaces for glycyI and alanyl residue ground states substantially similar to those obtained from calculations in which the energy is partitioned into various additive contributions such as nonbonded repulsions, torsions and electrostatic interactions. An independent check has been provided on the partitioned energy potential surfaces. The appearance of serious discrepancies must await a comparison of surfaces for oligopeptides in which helical ordering has been attained. Unfortunately the computation times required for the molecular orbital calculations leave such a comparison for the future. We do however anticipate studies of a polypeptide chain long enough to incorporate a single intramolecular hydrogen bond. We also have begun some studies of microscopic solvation of these structures, in which we investigate similar potential surfaces in the presence of discrete molecules of water or methanol. Calculations on excited-state surfaces are accessible only from the molecular orbital viewpoint and we

will report in the future on geometrical distortions, energy transfer, and the optical rotatory strengths of these molecules.

We are grateful to K. D. Gibson for generating atomic coordinates of the residues for us. This work was generously supported by the National Institutes of Health, Grant GM 13468.

References

1. G. Némethy and H. A. Scheraga, *Biopolymers*, **3**, 155 (1965).
2. S. J. Leach, G. Némethy and H. A. Scheraga, *Biopolymers*, **4**, 369, 887 (1966).
3. R. A. Scott and H. A. Scheraga, *J. Chem. Phys.*, **45**, 2091 (1966).
4. T. Ooi, R. A. Scott, G. Vanderkooi and H. A. Scheraga, *J. Chem. Phys.*, **46**, 4410 (1967).
5. G. N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan, *J. Mol. Biol.*, **7**, 95 (1963).
6. C. Ramakrishnan, *Proc. Indian Acad. Sci.*, **A59**, 327 (1964).
7. C. Ramakrishnan and G. N. Ramachandran, *Biol. J.*, **5**, 909 (1965), and references therein.
8. D. E. Brant and P. J. Flory, *J. Amer. Chem. Soc.*, **87**, 663, 2791 (1965).
9. P. DeSantis, E. Giglio, A. M. Liquori and A. Ripamonti, *Nuovo Cimento*, **26**, 616 (1962); *J. Polym. Sci. A*, **1**, 1383 (1963); *Nature*, **206**, 456 (1965).
10. A. M. Liquori, in *Perspectives in Polymer Science (J. Polym. Sci., C, 12)*, E. S. Proskauer, E. H. Immergut, and C. G. Overberger, Eds., Interscience, New York, 1966, p. 209.
11. R. Hoffmann, *J. Chem. Phys.*, **39**, 1397 (1963); *ibid.*, **40**, 2474, 2480, 2745 (1964); *Tetrahedron*, **22**, 521, 539 (1966) and subsequent papers.
12. W. Adam, A. Grimison, R. Hoffmann, and C. Zuazaga de Ortiz, *J. Amer. Chem. Soc.*, **90**, 1509 (1968).
13. G. Blyholder and C. A. Coulson, *Theor. Chim. Acta*, **10**, 316 (1968).
14. L. C. Allen and J. D. Russell, *J. Chem. Phys.*, **46**, 1029 (1967).
15. R. S. Mulliken, *J. Chem. Phys.*, **23**, 1833, 1841, 2338, 2343 (1955).
16. J. T. Edsall, P. J. Flory, J. C. Kendrew, A. M. Liquori, G. Némethy, G. N. Ramachandran, H. A. Scheraga, *Biopolymers*, **4**, 121 (1966); *J. Biol. Chem.*, **241**, 1004 (1966); *J. Mol. Biol.*, **15**, 399 (1966).
17. D. A. Brant and P. R. Schimmel, *Proc. Natl. Acad. Sci. U.S.*, **58**, 429 (1967).
18. D. C. Phillips, *Proc. Natl. Acad. Sci. U.S.*, **57**, 484 (1967).
19. H. C. Watson, personal communication to K. D. Gibson and H. A. Scheraga, cited in ref. 20; also Fig. 2 of ref. 21.
20. K. D. Gibson and H. A. Scheraga, *Biopolymers*, **4**, 709 (1966).
21. G. N. Ramachandran, C. M. Venkatachalan, and S. Krimm, *Biophys. J.*, **6**, 849 (1966).

Received August 15, 1968