A New General Form of Molecular Force Fields. Application to Intra- and Interresidue Interactions in Peptides

Pietro Amodeo and Vincenzo Barone*

Contribution from the Dipartimento di Chimica, Università di Napoli, via Mezzocannone 4, I-80134 Napoli, Italy. *Received January 27, 1992*

Abstract: The introduction of suitable expansion variables solves some nagging problems of current molecular force fields and offers a number of computational advantages. A parametrization for neutral peptides shows a remarkable agreement with refined ab initio computations, concerning both structural and energetic features of dipeptides. The onset of different helical structures in representative polypeptides is also correctly reproduced. This validates the balanced treatment of intra- and interresidue interactions. Entropic contributions evaluated from harmonic frequencies appear reasonable and not negligible in the evaluation of relative stabilities. A key role in all these achievements is played by the improved description of short-range repulsions between nonbonded atoms.

I. Introduction

Numerical simulations provide an invaluable complement to experiment in the study of structures and dynamics of biomolecules. These techniques rest on the hypothesis that the potential energy of a molecule or assembly of molecules can be reproduced by a low-order Taylor expansion in terms of internal coordinates. Several studies have been devoted to the refinement of parameters and to the analysis of their transferability from small model systems.[1-22] As an outcome, the most recent force fields (FFs) are perfectly adequate to the refinement of energy minima. However, the study of flexible molecules and of their conformational transitions requires a correct description of a much wider region of potential energy surfaces. Several attempts[23-26] convinced us that simple reparametrization of standard FFs is not sufficient to extend their range of application away from equilibrium and/or to crowded systems. Some inaccuracy is probably unavoidable if one wishes to retain the simplifying assumptions (e.g., use of pairwise interactions) necessary for computational efficiency.[1-7,21] We wonder, however, if the errors could be significantly reduced by an improved choice of expansion variables. To this end, we have introduced a new molecular force field and investigated its performances with special reference to the prediction of the secondary structure of peptides and proteins. Our approach rests on a simple, but flexible parametrization scheme, which leaves room for improvements when further reference data will be available. The lack of definite experimental information for some building blocks of proteins suggests, in this case, the use of ab initio quantum mechanical methods for purposes of reference. Of course, the level of reference computations must be sufficient to describe, in a balanced way, the interactions governing the physico-chemical properties of interest. Computations of this kind have recently been performed for some key systems[27-31] and will be integrated in this work by other computations especially devised to allow a better parametrization. The consistent results thus available for neutral dipeptide analogues with progressively larger steric hindrance at the C^{α} provide a reliable benchmark for empirical computations. Since dipeptides at most mimic local interactions of proteins, further insight in the overall balance of different effects can be gained by the study of longer polypeptides. As a first step, we analyze regular structures of oligomers of L-alanine (Ala) and α-methylalanine (Aib), which give rise to different helical structures.[32,33] The influence of dielectric screening and entropy on the overall conformational stabilities of model peptides is also investigated.

II. Methods

The structure, atom labeling, and main geometrical parameters of peptides are shown in Figure 1.

Model compounds of general forms Ac-$(X)_n$-NHMe and F-$(X)_n$-NH_2 have been considered, where Ac, NHMe, F, and NH_2 stand for acetyl, amidic $NHCH_3$, formyl, and amidic NH_2

(1) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361-2381.
(2) Nemethy, G.; Pottle, M. S.; Scheraga, H. A. *J. Phys. Chem.* **1983**, *87*, 1883-1887.
(3) Sippl, M. J.; Nemethy, G.; Scheraga, H. A. *J. Phys. Chem.* **1984**, *88*, 6231-6233.
(4) Weiner, P.; Kollman, P. A. *J. Comput. Chem.* **1981**, *2*, 287-309.
(5) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Chandra Singh, U.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765-784.
(6) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230-252.
(7) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187-217.
(8) Momany, F. A.; Klimkowsky, V. J.; Schafer, L. *J. Comput. Chem.* **1990**, *11*, 654-662.
(9) Allinger, N. L.; Burkert, U. *Molecular Mechanics*; ACS Monograph 177; American Chemical Society: Washington, DC, 1982.
(10) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127-8133.
(11) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8551-8566.
(12) Lii, J. H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8566-8575.
(13) Lii, J. H.; Allinger, N. L. *J. Comput. Chem.* **1991**, *12*, 186-199.
(14) Hermans, J.; Berendsen, H. J. C.; van Gunsteren, W. F.; Postma, J. P. M. *Biopolymers* **1984**, *23*, 1513-1518.
(15) Jorgensen, W. L.; Swenson, C. J. *J. Am. Chem. Soc.* **1985**, *107*, 569-578.
(16) Jorgensen, W. L.; Swenson, C. J. *J. Am. Chem. Soc.* **1985**, *107*, 1489-1496.
(17) Hagler, A. T.; Lifson, S.; Huler, E. *J. Am. Chem. Soc.* **1974**, *96*, 5319-5335.
(18) Lifson, S.; Hagler, A. T.; Dauber, P. *J. Am. Chem. Soc.* **1979**, *101*, 5111-5121.
(19) Hagler, A. T.; Stern, P. S.; Sharon, R.; Becker, J. M.; Naider, F. *J. Am. Chem. Soc.* **1979**, *101*, 6842-6852.
(20) Lifson, S.; Stern, P. S. *J. Chem. Phys.* **1982**, *77*, 4542-4550.
(21) Dauber-Osguthorpe, P.; Roberts, V. A.; Orguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. *Proteins: Struct. Funct. Genetics* **1988**, *4*, 31.
(22) Mayo, S. L.; Olafson, B. D.; Goddard, W. A., III *J. Phys. Chem.* **1990**, *94*, 8897-8909.
(23) Amodeo, P. *Tesi di Dottorato*, Naples, Italy (1989).
(24) Barone, V.; Fraternali, F.; Cristinziano, P. L.; Lelj, F.; Rosa, A. *Biopolymers* **1988**, *27*, 1673-1685.
(25) Barone, V.; Fraternali, F.; Cristinziano, P. L. *Macromolecules* **1990**, *23*, 2038-2044.
(26) Amodeo, P.; Barone, V.; Fraternali, F. *Thermochim. Acta* **1990**, *162*, 141-154.
(27) Schafer, L.; Van Alsenoy, C.; Scarsdale, J. N. *J. Chem. Phys.* **1982**, *76*, 1439-1444.
(28) Klimkowski, V. J.; Schafer, L.; Momany, F. A.; Van Alsenoy, C. *J. Mol. Struct.* **1985**, *124*, 143-153.
(29) Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C. L., III; Pople, J. *Int. J. Quantum Chem.:Quantum Biol. Symp.* **1989**, *16*, 311-322.
(30) Head-Gordon, T.; Head-Gordon, M.; Frisch, M. J.; Brooks, C. L., III; Pople, J. *J. Am. Chem. Soc.* **1991**, *113*, 5989-5997.
(31) Böhm, H. J.; Brode, S. *J. Am. Chem. Soc.* **1991**, *113*, 7129-7135.
(32) Toniolo, C.; Crisma, M.; Bonora, G. M.; Benedetti, E.; Di Blasio, B.; Pavone, V.; Pedone, C.; Santini, A. *Biopolymers* **1991**, *31*, 129-138.
(33) Hodgkin, E. E.; Clark, J. D.; Miller, K. R.; Marshall, G. R. *Biopolymers* **1990**, *30*, 533-546.

0002-7863/92/1514-9085$03.00/0

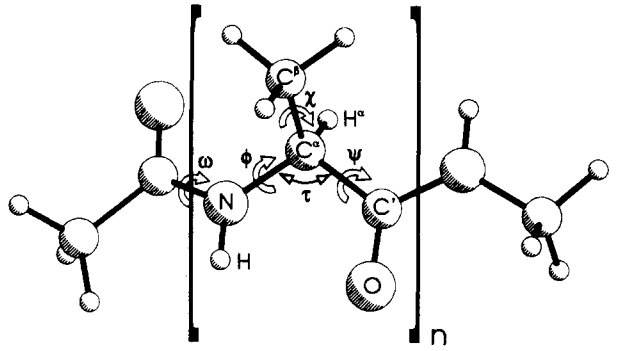

Figure 1. Atom labeling and principal geometric parameters of amino acid residues.

groups, respectively; X can be glycine (Gly), alanine (Ala), or α-methylalanine (Aib), and n ranges from 1 and 13. Standard geometries of residues and terminal groups were taken from refs 1–3, 34, and 35.

Regular structures are classified in terms of the number of atoms forming the "cycle" closed by an H-bond between NH and OC groups belonging to the backbone. According to this scheme, the extended conformation typical of β-sheets is labeled C_5; normal and reverse γ-turns are labeled C_{7ax} and C_{7eq}, respectively; a family of helices, whose most important example is the 3_{10} one, is labeled C_{10}; α-helices and their "variations", which differ in the exact number of residues per turn (from 3.7 in an ideal α-helix up to 4) are labeled C_{13}. Standard labels are, instead, used for semiextended structures, namely α' for $\phi = 180°$, $\psi = \pm 60°$, and P_{II} for $\phi = \pm 60°$, $\psi = 180°$.

Ab initio computations were performed by the GAUSSIAN 90 package[36,37] using the Huzinaga–Dunning (9,5;4)/[3,2;2] split valence basis set augmented by polarization functions on all atoms[38] (hereafter referred to as HD) and by the 4-21G basis set used in several previous studies of peptides.[27,28] Empirical computations were performed by the ICER package[25,39] and by a version of the AMBER 3.0 package modified to allow the automatic building of FR maps and to include the new potential energy functions and their derivatives.

In the case of the so-called dipeptide analogues ($n = 1$), ϕ,ψ energy maps were generated in two ways: (1) by varying only ϕ and ψ, with the peptide bond dihedral angles (ω) set to 180° and, where appropriate, with staggered methyl groups (this model will be referred to as rigid rotor (RR)); (2) by maintaining ϕ and ψ at fixed values (by rigid constraints) and minimizing the energy while allowing all other variables to change freely, (this model will be referred to as flexible rotor (FR)).

Unconstrained energy minimizations (EM) for dipeptides and oligopeptide analogues were also performed using both RR and FR models. The final Hessian matrix was used in the evaluation of harmonic frequencies and thermodynamic functions.[40]

III. The Force Field

Our force field is defined, as usual, over a set of redundant internal coordinates including interatomic distances, valence, and dihedral angles. Although a Taylor expansion has the formal flexibility of reproducing, within its radius of convergence, any given form of the potential energy surface, practical considerations (i.e., limitations in the number of coefficients which can be evaluated) make it necessary to employ expansion variables leading to rapid convergence. In the case of distances, it is well known that the Morse[41] or SPF[42] variables are superior to simple displacements because they increase more rapidly for bond compressions than stretches and because they have the correct asymptotic limit for large distances. We prefer Morse variables since only this case does a simple quadratic expansion turn out to be well suited to describe both covalent and van der Waals interactions. Furthermore, the physical soundness and computational efficiency of Morse variables is well known.[23,43,44] Trigonometric variables of the form $(\sin\theta/2 - \sin\theta_0/2)$ satisfies all the most important requirements for the angular dependence of the potential energy. In particular they (a) are symmetric around 180°, (b) have symmetric wells with minima at 180° for linear species, and (c) have two asymmetric wells, in which it is easier to open than to close the bond angle, for nonlinear molecules. Of course, the standard variables $\theta-\theta_0$ satisfy only condition b. Finally, the periodic parts of the potential can be represented by low-order Fourier series.

The total energy can be partitioned as follows:

$$E = E^{00}{}_{2b} + E^{qq}{}_{2b} + E_{3b} + E_{4b} \quad (1)$$

where the subscripts indicate the number of "bodies" (generally atoms) involved in each kind of interaction. The superscripts denote effective zero charge (00) and electrostatic (qq) interactions, respectively.

As mentioned above, a general quadratic expansion is used for $E^{00}{}_{2b}$

$$E^{00}{}_{2b} = \sum_{AB} \epsilon^0{}_{ab}[y^2{}_{AB} - 2y_{AB}] \quad (2)$$

where the Morse variable y_{AB} is defined in terms of the distance between A and B atoms

$$y_{AB} = \exp\left[\beta_{ab}\left(1 - \frac{r_{AB}}{r^0{}_{ab}}\right)\right] \quad (3)$$

Of course, different sets of parameters (ϵ^0, r^0, and β) are introduced for bond, geminal, vicinal, and distal interactions. Note that here and in the following, capital letters indicate atoms, and small letters denote atom types.

Electrostatic interactions between nonbonded atoms are well represented by point charge models. Those between directly bonded and geminal atoms can be implicitly included in the Morse or bending parameters since in conformational problems the relative displacements of these atoms are small. The situation is, however, more involved for vicinal atoms whose distances can vary from well above to well below the sum of van der Waals radii. Since the simple Coulomb law is not equally valid in these situations, we introduce a formally correct quantum mechanical ansatz for the charge dependent $E^{qq}{}_{2b}$ term

$$E^{qq}{}_{2b} = \frac{Z_A Z_B}{r_{AB}} + P_A P_B\,(s_A s_A|s_B s_B) - Z_A P_B\,(s_A s_A|r_{AB}) - P_A Z_B\,(s_B s_B|r_{AB}) \quad (4)$$

where Z's are core charges, P's are electron populations, and s atomic orbitals have been used to obtain effective two-center, two-electron and electron-core attraction integrals. Further[45]

$$(s_A s_A|r_{AB}) \approx (s_B s_B|r_{AB}) \approx (s_A s_A|s_B s_B) \quad (5)$$

Using next

$$Z_A Z_B/r_{AB} = Z_A Z_B(s_A s_A|s_B s_B) + \Delta_{AB} \quad (6)$$

(34) Benedetti, E. In *Chemistry & Biochemistry of Amino Acids, Peptides and Proteins*, Weinstein: New York, 1982, pp 105–184.
(35) Benedetti, E.; Bavoso, A.; Di Blasio, B.; Pavone, V.; Pedone, C.; Crisma, M.; Bonora, G. M.; Toniolo, C. *J. Am. Chem. Soc.* **1982**, *104*, 2437–2444.
(36) Frisch, M. J.; Head-Gordon, M.; Trucks, G. W.; Foresman, J. B.; Schlegel, H. B.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Defrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Kahn, L. R.; Stewartt, J. J. P.; Fluder, E. M.; Topiol, S.; Pople, J. A. *Gaussian 90*; Gaussian Inc.: Pittsburgh, PA.
(37) Schlegel, H. B. In *Ab-Initio Methods in Quantum Chemistry*, Part 1; Lawley, K. P., Ed.; John Wiley: New York, 1987; pp 249–286.
(38) Hay, P. J.; Dunning, T. H., Jr. In *Modern Theoretical Chemistry, Methods of Electronic Structure Theory*; Schafer, H. F.; Ed.; Plenum Press: New York, 1977; Vol. III.
(39) Barone, V.; Cristinziano, P. L.; Lelj, F. Unpublished.
(40) Hill, T. L. *An Introduction to Statistical Thermodynamics*; Addison-Wesley: Reading, MA, 1960.
(41) Morse, P. M. *Phys. Rev.* **1929**, *39*, 57–64.
(42) Simons, G.; Parr, R. G.; Finlan, J. M. *J. Chem. Phys.* **1973**, *59*, 3229–3234.
(43) Lifson, Sh.; Stern, P. S. *J. Chem. Phys.* **1982**, *77*, 4542–4550.
(44) Abraham, R. J.; Stolevik, R. *Chem. Phys. Lett.* **1978**, *58*, 622–624.
(45) Pople, J. A.; Beveridge, D. L. *Approximate Molecular Orbital Theory*; McGraw Hill: New York, 1970.

we obtain

$$E^{qq}{}_{2b} \approx q_A q_B (s_A s_A | s_B s_B) + \Delta_{AB} \quad (7)$$

where $q = Z - P$ are net charges. The short range repulsive term Δ_{AB} does not depend on electron populations and can be absorbed in $E^{00}{}_{2b}$. Two-center, two-electron integrals can be approximated as[46]

$$(s_A s_A | s_B s_B) = 1/(\rho_{\alpha\beta} + r_{AB}) \quad (8)$$

where $\rho_{ab} = 0.5(\rho_a + \rho_b)$ is an average orbital radius.[46]

Three-body interactions are considered only for geminal atoms:

$$E_{3b} = \frac{1}{2}\sum_{A}\sum_{B(A)}\sum_{C(A)} \Theta_{abc}\left(\sin\frac{\theta_{ABC}}{2} - \sin\frac{\theta^0{}_{abc}}{2}\right)^2 \quad (9)$$

where θ_{ABC} is the angle between bonds AB and BC. Here and in the following J(I) means that atom J is bonded to atom I. The constant Θ_{abc} in eq 9 is related to the harmonic bending force constant K_θ

$$\Theta_{abc} = 4K_\theta/\cos^2(\theta^0{}_{abc}/2) \quad (10)$$

Four-body interactions are considered only for vicinal atoms:

$$E_{4b} = \frac{1}{2}\sum_{A}\sum_{B(A)}\sum_{C(B)}\sum_{D(C)} V_{abcd}\,[1 - \cos(n_{abcd}(\phi - \phi^0{}_{abcd}))] \quad (11)$$

where ϕ is the angle between the ABC and BCD planes, n_{abcd} is the periodicity, V_{abcd} is the barrier to rotation for equilibrium bond lengths, and $\phi^0{}_{abcd}$ is the equilibrium angle. V_{abcd} is the total barrier after adding all possible A and D terms to the energy expression, but the energy is renormalized by the total number of terms having a common B and C. Inversion motions can be described by the same equation if D is bonded to B rather than to C, and ϕ is the angle between ABC and ABD planes.

IV. Parametrization

The first step is the definition of suitable atom classes, whose number results from a balance between accuracy and ease of parametrization. As a general rule, two atoms belong to the same class if they have the same atomic number and form the same number of σ bonds. For instance, all sp^3 carbon atoms belong to the class labeled C4, and hydrogens potentially involved in hydrogen bridges to the class H2. Second-row atoms involved in five- or six-membered aromatic rings belong to specific classes labeled by the numbers 5 and 6. The planned extension of the parametrization to other systems prompted us to reduce the number of parameters as far as possible. After several trials, we succeeded in enforcing the following features without significantly worsening the agreement with reference data.

(1) Interaction centers always coincide with atoms. As a consequence lone pair sites and foreshortening of XH bonds have not been employed.

(2) van der Waals and hydrogen bond interactions are described by the same parameter set. The parameters for interactions between atoms belonging to different classes are obtained by the following generalized combination rules:

$$r^0{}_{ab} = \frac{r^0{}_{aa} + r^0{}_{bb}}{2} - \Delta r(\delta_a^{acc}\,\delta_b^{don} + \delta_b^{acc}\,\delta_a^{don}) \quad (12a)$$

$$\epsilon^0{}_{ab} = (\epsilon^0{}_{aa}\epsilon^0{}_{bb})^{1/2} - \Delta\epsilon\,(\delta_a^{acc}\,\delta_b^{don} + \delta_b^{acc}\,\delta_a^{don}) \quad (12b)$$

where the δ's are 1 for potentially acceptor (δ^{acc}) or donor (δ^{don}) atoms and 0 otherwise. The fitted values of the other constants are $\Delta r = 0.9$ Å, and $\Delta\epsilon = 4.184$ kJ mol^{-1}.

(3) Charges are computed by additive bond increments $\gamma^0{}_{ab}$:

$$q_A = q^0{}_a + \sum_{B\ \text{bonded}} \gamma^0{}_{ab} \quad (13)$$

(4) A single average orbital radius is used in eq 8 ($\rho = 3$).

(5) The parameters of three-body interactions only depend on the class of the central atom.

(6) The parameters of four-body interactions only depend on the classes of the atoms forming the central bond.

The only departures from this scheme concern the N3C4C3X moiety (i.e., the ψ torsional angle), some valence angle involving the C^α atom, and the valence angles of the nonstandard terminal groups HCO and NH_2 (see Table III).

The bond parameters involving C4 and H1 classes are those of ref 43. The equilibrium distances (r^0) and harmonic force constants for other atom types were taken from the AMBER FF,[6] and dissociation energies (ϵ^0) from ref 47. From these data we obtained the β parameters of Table II. Although the anharmonicity of stretchings is relatively unimportant in determining molecular structures and relative stabilities, this is not the case for vibrational frequencies. In particular, only Morse functions reproduce the shift in NH and CO stretching frequencies, connected to the formation of intramolecular H bonds.

The parameters $\gamma^0{}_{ab}$ of Table I were obtained by a best fit of experimental dipole moments for formamide, *N*-methylacetamide, and peptide unit, and of the IR-derived charges for nonpolar aliphatic and aromatic hydrogen atoms.[48,49]

Table I. One-Center Parameters[a]

atom class	$\epsilon^{0\,b}$ (kJ mol^{-1})	$r^{0\,b}$ (Å)	δ^{acc}, δ^{don}	$^1/_2K_\theta{}^c$ (kJ mol^{-1} deg^{-2})	$\theta^{0\,c}$ (deg)
C4	0.215	3.970	0, 0	146.44	109.47
C3 = C6	0.215	3.853	0, 0	334.72	120.00
N[d]	0.269	3.623	0, 1	209.20	120.00
O2	0.245	3.461	0, 1	230.12	109.47
O1	0.323	3.454	0, 1		
H2	0.176	2.386	1, 0		
H1	0.230	2.826	0, 0		

[a] See text for definitions. [b] Used for nonbonded atoms. [c] Used for all terminal atoms, except those explicitly given in Table III. [d] All types.

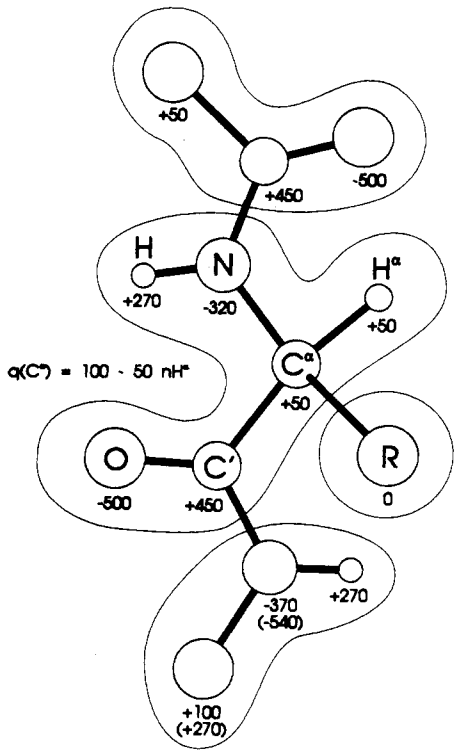

Figure 2. Atomic charges (in |me|) obtained with the parameters of Table I for standard amino acid residues and terminal groups. The values in parentheses refer to NH_2 terminal.

(46) (a) Ohno, K. *Theor. Chim. Acta* **1964**, *2*, 219–227. (b) Klopman, G. *J. Am. Chem. Soc.* **1964**, *86*, 4550–4557.

(47) Benson, S. W. *J. Chem. Ed.*, **1965**, *42*, 502–518.

(48) Gussoni, M.; Castiglioni, C.; Zerbi, G. *J. Chem. Phys.* **1984**, *80*, 1377–1381.

(49) Gussoni, M.; Castiglioni, C.; Zerbi, G. *THEOCHEM* **1986**, *138*, 203–212.

Table II. Pair Parameters[a]

atom pair (a–b)	$\gamma^0{}_{ab}$[b]	$\epsilon^0{}_{AB}$[c] (kJ mol^{-1})	$r^0{}_{ab}$[c] (Å)	β_{ab}[c,d]	V_{ab} (kJ mol^{-1})	n_{ab}	$\phi^0{}_{ab}$ (deg)
C4–C4	0.00	368.2	1.538	2.864	10.9	3	180.0
C4–C3	0.05	368.2	1.538	2.864	0.0	0	0.0
C4–C6	0.05	368.2	1.530	2.864	0.0	0	0.0
C4–N3	0.10	330.5	1.460	2.646	8.0	3	180.0
C4–O2	0.50	380.7	1.410	2.646	4.0	3	180.0
C4–H1	0.05	446.4	1.117	1.980			
C3–C3	0.00	682.0	1.400	2.374	44.4	2	180.0
C3–C6	0.00	682.0	1.400	2.374	44.4	2	180.0
C3–N3	0.00	543.9	1.349	2.091	83.7	2	180.0
C3–O1[e]	0.50	732.2	1.227	2.331	87.9	2	180.0
C3–H1[e]	−0.10	418.4	1.100	1.980	8.4	2	180.0
C6–H1[e]	−0.12	438.2	1.100	1.980	8.4	2	180.0
N3–H2[e]	−0.27	443.5	1.000	2.110	8.4	2	180.0

[a] See text for definitions. [b] $\gamma^0{}_{ba} = -\gamma^0{}_{ab}$. [c] Used for directly bonded atoms. [d] Obtained from force constants K_{ab} by $\beta_{ab} = (K_{ab}/2r^0{}_{ab}{}^2\epsilon^0{}_{ab})^{1/2}$. [e] V, n, and ϕ^0 parameters refer to improper torsions around this bond (see text).

Table III. Special Three-Body Parameters[a]

atom types	$^1/_2K_\theta$ (kJ mol^{-1} deg^{-2})	$\theta^0{}_{abc}$ (deg)
C4C4C4	167.4	109.5
C4C4C3	264.9	111.1
C4C4N3	334.7	109.5
C3C4N3	251.0	114.5
N3C3O1	418.4	126.1
N3C3H1	209.2	109.5
O1C3H1	209.2	117.5
C4N3H2	159.0	118.4
H2N3H2	92.0	112.3

[a] See text for definitions.

Our charges (see Figure 2) agree rather satisfactorily with point charges derived from the ab initio molecular electrostatic potential (MEP) of glycine and alanine dipeptide analogues.[31,50] The only significant difference is the rather low absolute value used in our FF for the charges of amidic carbon and nitrogen. Charge redistributions among "internal" atoms have, however, only a negligible influence on the MEP because of efficient shielding by "external atoms". Thus we prefer to accept this minor discrepancy in order to obtain the partitioning of peptides in simple "neutral groups". This allows the implementation of a particularly efficient strategy for the computation of long-range electrostatic interactions.[51,52] Figure 2 points out the fairly small dimensions of our neutral groups. The parameters of the Morse functions describing van der Waals and hydrogen bond interactions were obtained from a least-squares fitting to a number of reference data. Since Morse functions are quite "soft", a higher weight was assigned to repulsive regions. The reference data for van der Waals interactions were provided by computations using the potential functions developed by Williams and co-workers for the study of intermolecular interactions.[53,54] The results obtained by the above functions are in excellent agreement with refined ab initio computations of small model systems.[55,56] Some modifications were only introduced to enhance agreement with ab initio computations for sterically hindered compounds characterized by short O1–H1 and H1–H1 interactions. The reference data for hydrogen bridges were provided by refined ab initio computations of amide dimers, with just some consistency check on crystal packing of amides.[23]

Vicinal interactions generally give an overall repulsive contribution to total energy; this repulsion must be somehow reduced in order to obtain correct stabilities and trends for valence angles. Even in RR calculations, where geometrical effects on valence angles are not involved, the predicted stability of some conformation (e.g., C_5 in peptides) strongly depends on the parametrization of 1–4 interactions. Pending further theoretical analysis, we have adopted an approach shared by several FFs, namely, the use of the same parameters as for nonbonded interactions, but with a scale factor of 0.5.

V. Results for Dipeptide Analogues

The quantitative reproduction of hydrogen bond strengths requires large, multiply polarized basis sets, and proper account of many body effects.[57,58] Since this level of computation is out of question for peptides, we investigated whether some error compensation allows the use of less refined procedures. The set of simple hydrogen-bonded 1:1 complexes reported in Table IV was used as a benchmark. Hartree–Fock computations with the HD basis set (hereafter referred to as HF/HD) appear remarkably accurate concerning both structural and energetic aspects. This can be ascribed to some error compensation (for instance, the basis set superposition error and the many-body contribution for the water dimer are −4 and 3 kJ mol^{-1}, respectively), but, from a practical point of view, it gives strong support to the quantitative accuracy of this kind of computations. Recent studies[23] have shown that the same level of computation provides an accurate description of the conformational behavior of saturated and aromatic systems not involving hydrogen bridges. At least judging from the results of Table IV, minimal or split valence basis sets are not reliable enough for the determination of conformer stabilities in peptides. This kind of computation can be, however, useful in the analysis of structural changes and frequency shifts induced by conformational variations.[27,28]

Some energy minimizations of Gly and Ala dipeptide analogues have been performed at the HF/HD or HF/6-31+G** levels,[29–31] but no conformational maps are available to the best of our knowledge. Since the key intra-residue interactions can be investigated using a model as simple as F–Gly–NH_2, we computed a ϕ,ψ map of this molecule at the HF/HD level. A comparison with the corresponding map obtained by our FF is shown in Figure 3. If an upper limit of 20–25 kJ mol^{-1} is used to define allowed conformations, both maps define a trapezoidal region (shadowed in Figure 3) limited by C_5 ($\phi = \psi = 180°$) and bridge ($\phi = \pm 90°$, $\psi = 0°$) structures, and including the C_7 conformer ($\phi \approx \pm 90°$, $\psi \approx \mp 90°$). Standard helical conformations ($\phi \approx \pm 60°$, $\psi \approx \pm 30°$) do not correspond to local energy minima, lie outside the allowed region, and are less stable than bridge structures. Two energy minima are found in the maps, corresponding to the C_5 and C_7 conformers, but the stability order of the two minima is quite uncertain. In particular, Hartree–Fock computations (Table V and refs 29–31) by polarized basis sets favor the extended structure, whereas correlation energy slightly favors the C_7

(50) Williams, D. E. *Biopolymers* **1990**, *29*, 1367–1386.
(51) Berendsen, H. J. C.; van Gusteren, W. F. In *Molecular Liquids, Dynamics and Interactions*, Barnes, A. J., et al., Eds.; NATO ASI Series C135; Reidel: Dordrecht, 1984; pp 475–500.
(52) Van Gusteren, W. F.; Berendsen, H. J. C. *Groningen Molecular Simulations (GROMOS) Library Manual, Biomos*, Nijenborgh 16, Groningen, The Netherlands, 1987.
(53) Cox, S. R.; Hsu, L. Y.; Williams, D. E. *Acta Cryst.* **1981**, *A37*, 293–301.
(54) Williams, D. E.; Weller, R. R. *J. Am. Chem. Soc.* **1983**, *105*, 4143–4148.
(55) Starr, T. L.; Williams, D. E.; *J. Chem. Phys.* **1977**, *66*, 2054–2057.
(56) Huiszoon, C.; Mulder, F.; *Mol. Phys.* **1979**, *38*, 1497–1515; **1980**, *40*, 249–251.
(57) Szalewicz, K.; Cole, S. J.; Kolos, W.; Bartlett, R. J. *J. Chem. Phys.* **1988**, *89*, 3662–3671.
(58) Latajka, Z. *THEOCHEM* **1991**, *251*, 245–260.

Table IV. Comparison between Heavy Atom Distances (R in Å) and Stabilization Energies (ΔE in kJ mol^{-1}) of Representative Hydrogen Bonded Complexes Obtained by Different Methods

	STO-3G		4-21G		HD		6-31+G**			exper	
system	R	ΔE	R	ΔE	R	ΔE	R	ΔE	MP2[a] ΔE	R	ΔE
$(HF)_2$	2.57	23.0	2.72	32.2	2.79	21.3	2.79	19.7	20.5	2.79	25.0
$(H_2O)_2$	2.73	25.1	2.87	33.9	2.98	20.9	2.97	20.5	25.9	2.98	21.3
$(NH_3)_2$	3.08	15.9	3.31	17.2	3.44	10.0	3.38	11.7	16.7	3.34	12.6
$FH\cdots OH_2$	2.63	31.4	2.63	56.1	2.70	35.1	2.72	36.0	41.4	2.66	35.6
$FH\cdots NH_3$	2.77	34.7	2.69	68.2	2.73	51.5				2.66	49.4
$HOH\cdots NH_3$	2.89	17.2	3.21	17.2	3.05	24.3	3.04	25.1	32.2	2.98	25.9
$FAM\cdots HOH$	2.75	33.5	3.10	33.5	3.00	24.3				3.00	25.9
$H_2O\cdots FAM$	2.70	32.6	3.25	32.6	3.10	21.8				3.10	23.0
FAM_2	2.64	73.6	2.90	76.6	2.99	59.4				2.95	58.6

[a] 6-31+G** basis set at HF/6-31+G** geometry.

Table V. Results for Peptides

	flexible rotor							rigid rotor		
label	ϕ	ψ	τ	ΔE_c	ΔE_v	$T\Delta S_v$	ΔG	ϕ	ψ	ΔE_c
					F–Gly–NH_2					
C5	180.0	180.0	107.6	0.0	0.0	0.0	0.0	180.0	180.0	0.0
C7	±74.2	∓67.3	111.0	0.0	0.4	−3.3	3.7	±79.1	∓69.0	−4.9
bridge	(±90)	(0)	112.7	21.6				(±90)	(0)	19.1
α	(±57)	(±47)	111.5	27.8				(±57)	(±47)	21.8
3_{10}	(±60)	(±30)	112.6	22.4				(±60)	(±30)	17.6
					F–Gly–NH_2 (HF/HD)					
C5	180.0	180.0	109.3	0.0	0.0	0.0	0.0	180.0	180.0	0.0
C7	±87.2	∓70.1	112.6	4.3	1.6	−3.4	9.3	±86.1	∓70.9	2.3
bridge	(±90)	(0)	116.0	14.3				(±90)	(0)	23.0
α								(±57)	(±47)	37.7
3_{10}								(±60)	(±30)	30.5
					Ac–Gly–$NHCH_3$					
C5	180.0	180.0	107.6	0.0	0.0	0.0	0.0	180.0	180	0.0
C7	±73.7	±73.7	∓71.9	0.7	0.3	−3.3	4.3	±76.2	∓74.2	−8.0
					Ac–Ala–$NHCH_3$					
C_5	−157.0	169.1	106.7	0.0	0.0	0.0	0.0	−170.0	166.0	0.0
C_{7eq}	−73.7	71.9	110.6	−0.4	0.2	−2.5	2.3	−76.3	72.7	−5.8
C_{7ax}	64.8	−67.9	112.9	5.5	0.7	−2.8	9.0	53.9	−85.2	17.3
α'	−168.6	−57.7	109.8	32.9	−0.5	0.7	31.7	−158.1	−59.0	29.8
helix^L								49.0	31.3	25.5
					Ac–Aib–$NHCH_3$					
C_5	180.0	180.0	105.6	0.0	0.0	0.0	0.0	180.0	180.0	0.0
C_7	±64.9	∓67.3	111.3	2.8	0.3	−1.5	4.6	±76.7	∓61.7	1.1
α'	166.3	−53.6	107.5	28.5	0.9	2.2	27.2	173.1	−57.1	20.1
helix								±49.2	±30.5	5.0
P_{II}								±58.8	∓155.4	10.0

[a] Angles are in degrees, and thermodynamic data (at 298 K) in kJ mol^{-1}. Structures with dihedral angles in parentheses are not local minima.

structure.[30] The results obtained by our FF are intermediate between the two extrema. It is noteworthy that extended and bridge structures are probably stabilized by hyperconjugative interactions, which cannot be reproduced by standard parametrizations. Agreement with quantum mechanical findings can be obtained only adopting a 2-fold torsional potential on ψ. This has been suggested long time ago[59] and has been also adopted in the last revision of the CHARMm FF.[8] Thermodynamic functions derived from scaled 4-21G harmonic frequencies[60] show (see Table V) that entropy disfavors the structure (C_7) with the strongest hydrogen bridge. The entropic contribution is larger than the zero-point one, thus suggesting a strong effect of hydrogen bonding on low-frequency torsional motions. This is also apparent on inspection of individual frequencies and shows that the shape of minima in the (ϕ,ψ) space is not less important than their depth. Although our FF is too simple to provide reliable frequencies, it is gratifying that the entropic contributions to the relative stabilities of C_5 and C_7 conformations are in remarkable agreement with ab initio results.

Alanine is the standard benchmark of new theoretical methods in the field of peptides and proteins. Our results for the model Ac–Ala–NHMe are collected in Figure 4 and Table V. Five relative minima are present in the maps, with an energy spreading of less than 40 kJ mol^{-1}, the two lowest ones corresponding to C_{7eq} and C_5 conformers. A comparison between RR and FR results shows that full geometry relaxation mainly affects the relative stability of the C_{7ax} conformation. In particular, it determines a general energy lowering and a broadening of the region including this conformer. This is due to a partial reduction of the steric strain associated with the axial methyl group in the seven-member "ring" characteristic of this conformation. As a consequence of the increased stability of this region, the α^L minimum observed in RR calculation collapses into the C_{7ax} minimum. Owing to the absence of sterical strain in the C_{7eq} conformer, right helical conformations never correspond to energy minima. A further effect of full geometry relaxation is the stabilization of the C_5 conformer in comparison with the C_{7eq} one. In fact, extended structures generally involve $NC^{\alpha}C'$ valence angles significantly narrower (106.7° in the present case) than those characteristic of folded conformations which, in turn, are close to the standard value of 110.5° used in RR studies.

A comparison with ab initio results shows a substantial agreement with our results. This is particularly evident if the FR map of F–Ala–NH_2 of Figure 5 is compared with the 3-21G ab

(59) Ramachandran, G. N. In *Peptides, Polypeptides, and Proteins*; Blout, E. R., Bovey, F. A., Goodman, M., Lotan, N., Eds.; John Wiley: New York, 1974.

(60) Arnaud, R.; Subra, R.; Barone, V.; Lelj, F.; Olivella, S.; Solè, A.; Russo, N. *J. Chem. Soc., Perkin Trans. 2*, **1986**, 1517–1524.

Table VI. Relative Stabilities of Different Conformers of Ac–Ala–NHMe (in kJ mol^{-1}) According to different Methods[a]

label	HF/HD FR[b]	present FR[c]	AMBER FR[c,d]	CHARMm FR[c,d]	CVFF FR[c]	present RR[c]	ECEPP RR[d]
C_5	2.1	0.4	23.4	8.8	(10.0)	5.8	2.9
C_{7eq}	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C_{7ax}	12.5	5.5	4.6	10.0	5.4	17.3	30.5
α^R	(13.5)	(26.2)	23.8	(30.0)	17.6	(20.1)	3.3
α^L	18.5	(28.9)	28.9	(40.0)	(30.5)	25.5	10.0
α'	23.8	32.9				29.8	

[a] Structures with dihedral angles in parentheses are not local minima. [b] From ref 31. [c] Using a unitary dielectric constant. [d] From ref 64.

Table VII. Relative Stabilities of Different Conformers of Ac–Aib–NHMe (in kJ mol^{-1}) According to Different Methods

label	STO-3G FR[a]	HD FR[b]	present FR	AMBER FR[a]	STO-3G RR[a]	HD RR[b]	present RR	ECEPP RR[a]
C_5	0.0	0.0	0.0	18.0	0.8	0.0	0.0	42.7
C_7	0.0	4.0	2.8	0.0	0.0	3.2	1.1	27.2
helix	12.6	20.5		14.6	1.7	8.6	5.0	0.0
P_{II}	20.1				13.6		10.0	37.6

[a] From ref 25. [b] Estimated from the differences between the two basis sets in the analogous minima of F–Gly–NH_2.

initio map of Figure 4 in ref 30. The only minor discrepancy in the absence of left-handed helical minima in our FR calculations. However, the helical and β_2 minima characterized by basis sets not including polarization functions effectively vanish at more reliable levels. Therefore, the authors of ref 30 regard "the hydrogen-bonding conformations C_{7eq}, C_{7ax}, and C_5, and the α' structure as the only stable minima on the F–Ala–NH_2 surface". It is finally noteworthy that our FF correctly reproduces the conformation dependent trends of valence angles obtained by flexible geometry ab initio computations.

Table VI shows a comparison between the results obtained by the major current FFs; the improved agreement with ab initio results afforded by the present FF is quite apparent. The occurrence of C_5 and C_7 conformations has been experimentally confirmed for the isolated peptide in low-polarity solvents[61] and in argon matrix.[62] These findings are generally interpreted in terms of a C_7 conformation marginally more stable than the C_5 one. An increase in solvent polarity leads to the stabilization of helical structures, which in highly polar solvents, such as water, represent the absolute minima for dipeptide analogues. This is exactly the trend obtained in computations performed with different dielectric constants; already for $\epsilon = 4$, both right- and left-handed helices become relative energy minima (see Figure 4), and a 3_{10R}-like conformer becomes the absolute minimum when $\epsilon \geq 20$.

The Aib residue differs from the Ala one in a methyl group that substitutes a hydrogen atom on the C^α position. The difference in the conformational behavior of the two molecules is well evidenced by the ϕ,ψ maps of Figure 4. The accessible conformational space of Aib is severely restricted, but energy minima still occur in the same regions, namely, C_5, C_7, helical, and α'. The presence of the second β-methyl leads to a narrowing of low-energy regions rather than to significant displacements of their positions. In particular, the helical and C_7 energy minima of Aib strongly resemble the C_{7ax} and α^L ones of Ala since one of the two β-methyl groups of Aib always occupies the more sterically hindered "axial" position. As a consequence, the stabilization induced by full geometry optimization becomes particularly significant for this residue. The available ab initio results suggest that C_5, C_7, and helical structures have comparable energies.[25] Even more significant are the results of an experimental conformational study of Ac–Aib–NHMe in CCl_4.[63] Since only C_5 and C_7 structures were detected, the authors conclude that the study "has not provided evidence for a particular behavior of this residue due to eventual steric hindrances". It is, however, well documented that the Aib residue preferentially induces helical structures also for short polypeptides.[25,32,35] Our results allow a rationalization of the above data since, from one side, C_7 and C_5 structures remain the absolute energy minima, but, on the other side, the relative stability of helical structures is significantly enhanced in comparison with standard residues. No other current FF is able to give analogous results (see Table VII), probably owing to the overestimation of steric repulsions between methyl groups and carbonyl oxygen.[25] The overall effect of an increase

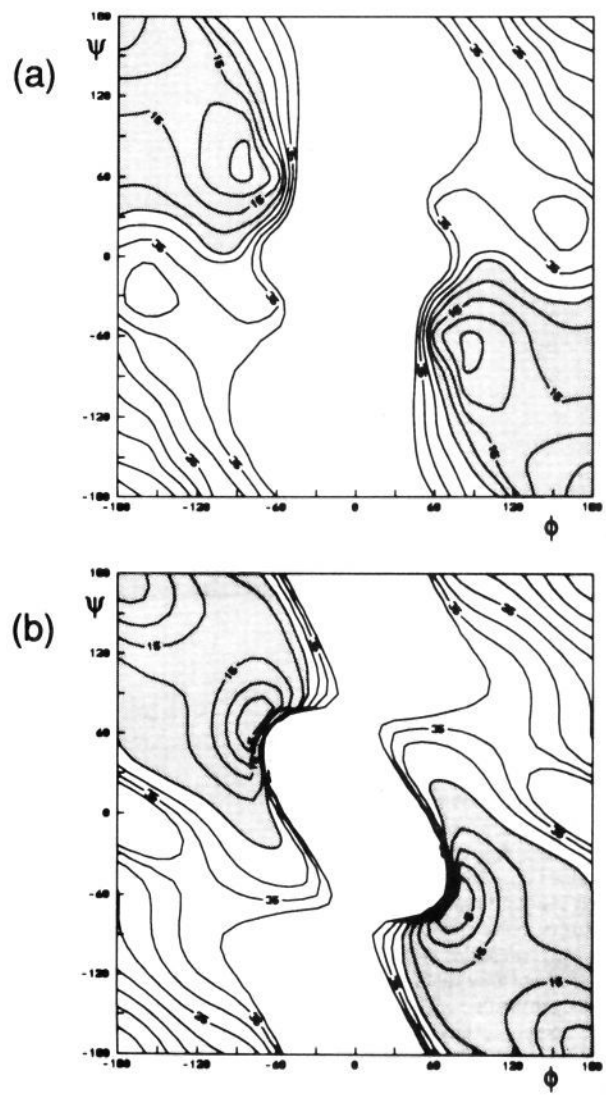

Figure 3. Rigid rotor (ϕ,ψ) maps of F–Gly–NH_2 obtained by (a) Hartree–Fock computations by the HD basis set; (b) the force field proposed in this work. Contour lines are drawn every 5 kJ mol^{-1} up to 40 kJ mol^{-1} above the absolute minimum.

(61) Yamazaki, T.; Abe, A. *Biopolymers* **1988**, *27*, 969–984.
(62) Grenie, Y.; Avignon, M.; Garrigou-Lagrange, C. *J. Mol. Struct.* **1975**, *24*, 293–307.
(63) Aubry, A.; Protas, J.; Boussard, G.; Marraud, M.; Neel, J. *Biopolymers* **1978**, *17*, 1693–1711.

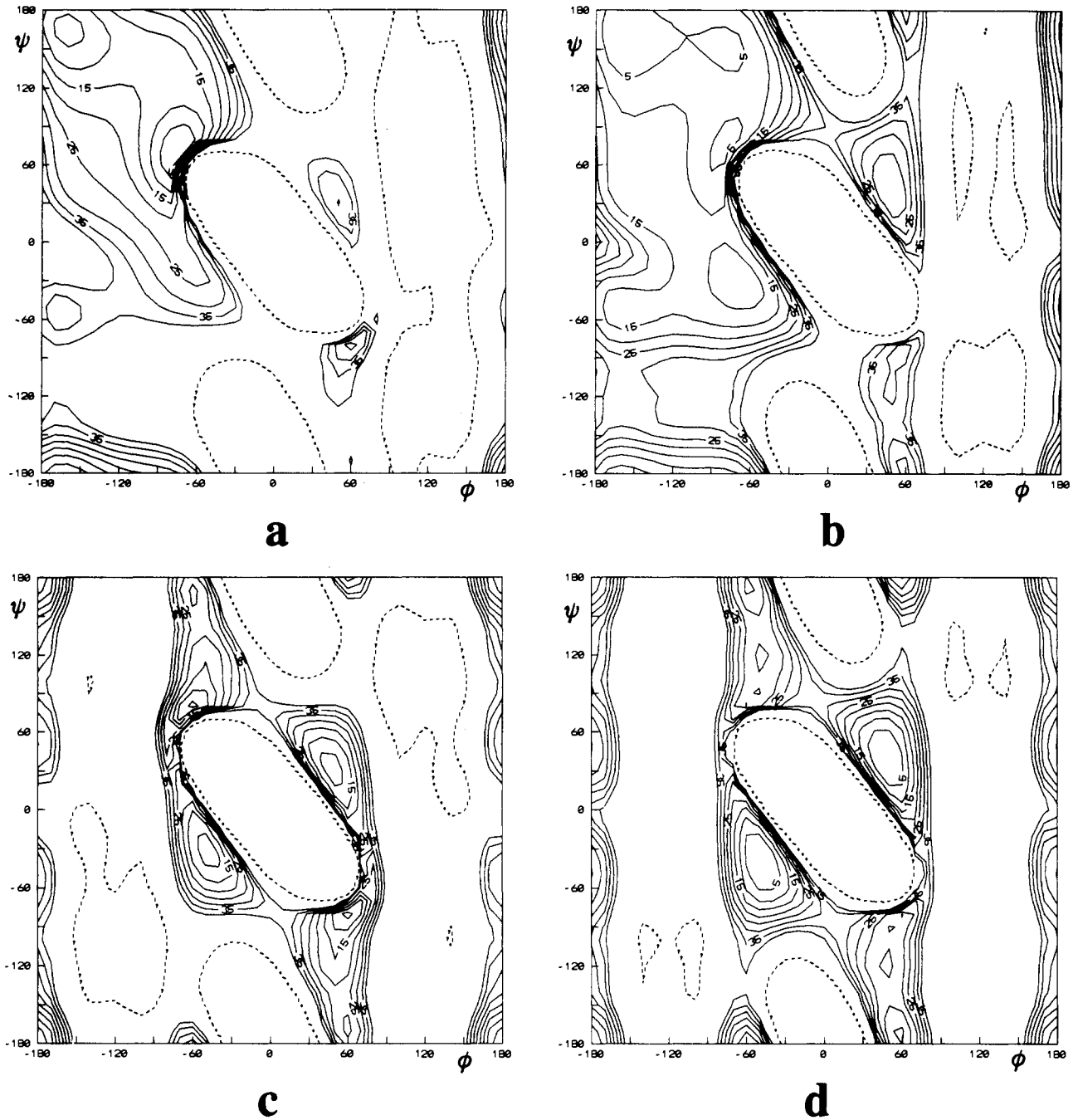

Figure 4. Rigid rotor (ϕ,ψ) maps of (a) Ac–Ala–$NHCH_3$ with a dielectric constant of 1; (b) Ac–Ala–$NHCH_3$ with a dielectric constant of 4; (c) Ac–Aib–$NHCH_3$ with a dielectric constant of 1; (d) Ac–Aib–$NHCH_3$ with a dielectric constant of 4. Continuous contour lines are drawn every 5 kJ mol^{-1} up to 40 kJ mol^{-1} above the absolute minimum. The dashed contour line is 80 kJ mol^{-1} above the minimum.

in the dielectric constant is the same as observed in Ala, namely, a relative stabilization of helical structures and a slight destabilization of C_7 versus C_5.

In summary, the general trends provided by our FF are consistent with all the quantum mechanical and experimental results available for simple dipeptide analogues. A number of evidences are now available to support the view that in these simple models there are no significant minima in the regions of conformational space corresponding to secondary structures of longer peptides and proteins. This, in turn, means that inter-residue effects must be treated on the same footing as intra-residue ones, and cannot be considered as simple perturbations.

VI. Results for Polypeptides

The balance between intra- and inter-residue interactions has been investigated in the case of regular structures of Ala and Aib homopolypeptides. In order to minimize end effects, particular attention has been devoted to the central residue, which in the longer polypeptides gives, even for folded structures, essentially converged inter-residue interactions.

Table VIII. Intra- and Interresidue Contributions to the Relative Stabilities of Different Conformers with Respect to C_5 Structures in L-Alanine and Aib[a] Tridecamers

	Ala				Aib	
interaction	C_{7eq}	C_{7ax}	C_{10R}	C_{13R}	C_{10}	C_{13}
i to $i+1$	−14.9	−12.7	−19.1	−12.8	−22.0	−11.3
i to $i+2$	−16.1	−14.9	−0.6	−1.8	−1.5	−0.8
i to $i+3$	0.0	0.0	−23.3	−13.0	−23.4	−12.7
i to $i+4$	0.0	0.0	−1.9	−22.0	−1.7	−22.3
others	0.0	0.0	−2.7	−5.2	−2.7	−4.5
inter	−31.0	−27.6	−47.6	−54.8	−51.3	−51.6
intra	31.2	34.9	35.3	37.1	33.9	37.0
total	0.2	7.3	−12.3	−17.7	−17.4	−14.6

[a] All the energies are in kJ mol^{-1}.

Table IX. Structural and Energetic Characteristics for Tridecamers of Alanine and α-Methylalanine[a]

label	ϕ	ψ	σ	h^b	n_R^c	ΔE_c	ΔE_v	$T\Delta S$	ΔG
				Ac–$(Ala)_{13}$–NHMe, $\epsilon = 1$					
C_5	−155.1	171.1	106.5			0.0	0.0	0.0	0.0
C_{10R}	−46.9	−32.9	112.9	1.86	3.31	−112.6	−4.2	−57.8	−59.0
C_{13R}	−59.4	−43.5	110.7	1.50	3.63	−145.7	−3.0	−78.8	−69.9
				Ac–$(Aib)_{13}$–NHMe, $\epsilon = 1$					
C_5	180	180	105.5			0.0	0.0	0.0	0.0
C_{10}	±41.6	±35.6	112.3	2.02	3.05	−172.1	−11.8	−42.1	−141.8
C_{13}	±52.3	±53.1	109.2	1.55	3.68	−146.5	−6.9	−75.3	−78.1
				Ac–$(Aib)_{13}$–NHMe, $\epsilon = 4$					
C_5	180	180	108.4			0.0	0.0	0.0	0.0
C_{10}	±46.4	±34.7	113.2	2.01	3.13	−130.6	−8.1	−33.9	−104.8
C_{13}	±53.9	±52.1	111.2	1.58	3.71	−131.4	−3.6	−70.9	−64.1

[a] Energies and other thermodynamic quantities are in kJ mol^{-1}; angles in deg. [b] Unit height (Å). [c] Number of residues per turn.

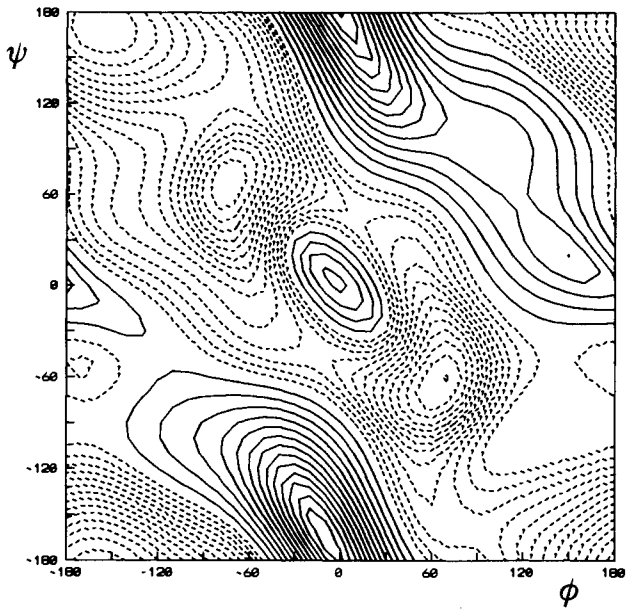

Figure 5. Flexible rotor (ϕ,ψ) map of F–Ala–NH_2. For a better comparison with the corresponding map in ref 30, dashed contour lines are drawn every 2.092 kJ mol^{-1} and extend up to 29.288 kJ mol^{-1} above the C_{7eq} absolute minimum. Continuous contour lines are drawn every 4.184 kJ mol^{-1} thereafter.

From a structural point of view, our results confirm the sensitivity of the valence angle τ (see Figure 1) to molecular conformation; in fact, values ranging between 105° and 113° are obtained for the central residue and even larger values occur in terminal ones. All the other stiff degrees of freedom are much less sensitive to conformation. It has been recently argued[64,65] that reliable FR studies cannot be performed by current FFs, since they predict a variation of valence angles for different conformations wider than the experimental range of about 3.5°. In the present FF only the τ valence angle has a spread larger than this threshold. The large value of τ obtained for C_{7ax} conformers is not statistically significant since these structures have never been detected, and the narrow value found in the C_5 conformation of Aib is fully supported by a recent experimental study.[66] Small values of τ have also been found in other nonstandard residues, which adopt extended conformations.[67,68] Furthermore, as discussed in the previous section, ab initio computations provide similar trends for τ in the case of dipeptide analogues.

A detailed analysis of regular polymers can be based on the data of Table VIII. The largest inter-residue interactions occur between nearest neighbor units, and stabilize by similar amounts all the folded structures versus the extended minimum. For C_5 and C_7 conformers interactions further apart than i to $i + 2$ are negligible and, at this level, C_5 and C_{7eq} structures are nearly isoenergetic on a *per unit* basis. Regular sequences of C_{7ax} structures are destabilized by inter-residue interactions. This explains why they are not experimentally observed despite the relative stability in the dipeptide model. It is apparent that the C_{13R} helix is the preferred structure for sufficiently long polymers, but it is also interesting to analyze the onset of different stable structures as a function of the number of residues in a polypeptide chain. Although only some results obtained for tridecamers are reported in Table IX for purposes of illustration, the following remarks are based on the ensemble of computations.

In the case of alanine, C_5 and C_{7eq} conformers are favored for oligopeptides of less than five to seven residues; starting from this point, the most stable regular structures correspond to various types of helices, which involve stronger inter-residue hydrogen bridges and lower steric repulsions. In particular, C_{10} and C_{13} right-handed helices correspond to true energy minima. Entropy effects account for up to 50% of the total free energy differences among different extended and folded structures. They stabilize conformers in the following order: $C_5 > C_7 > C_{10R} > C_{13R}$.

Also in the case of Aib, extended structures correspond to the absolute energy minima for short oligomers. The onset of helical structures is significantly faster than in the case of alanine, the 3_{10} helix becoming the absolute minimum already for the tetramer or pentamer. In agreement with experiment[35] and previous calculations,[25] helical parameters for Aib peptides differ from those of Ala and, in general, from those of standard residues. The secondary structure of Aib oligomers is dominated by the 3_{10} helix, the α helix corresponding to a true energy minimum only for the tridecamer. The dielectric screening induced stabilization of C_{13} structures versus C_{10} ones is not sufficient to reverse the predicted stability order of the two conformers, at least up to the tridecamer. It is noteworthy that also in experimental studies no α-helix structures have been found until now for oligomers up to tridecamers.

VII. Summary and Conclusion

In the present study we have suggested a new general force field able to describe in a balanced way the conformational behavior of peptides and proteins. Ab initio computations for dipeptide analogues (refs 29–31 and present work) have pointed out a number of significant conformational trends, concerning, in particular, the comparable stability of C_5 and C_7 structures and the lack of a well-defined energy minimum corresponding to an incipient helical structure. Bridge ($\psi \approx 0°$) conformations are, on the other hand, readily accessible due to the combined effect of a weak NH···N hydrogen bond and of hyperconjugative interactions between neighboring amidic moieties.

None of the major current FFs is able to reproduce all these trends. More generally, one of the most common drawbacks of

(64) Roterman, I. K.; Lambert, M. H.; Gibson, K. D.; Scheraga, H. A. *J. Biomol. Struct. Dynam.* **1989**, *7*, 391–419.
(65) Roterman, I. K.; Lambert, M. H.; Gibson, K. D.; Scheraga, H. A. *J. Biomol. Struct. Dynam.* **1989**, *7*, 421–453.
(66) Valle, G.; Toniolo, C.; Jung, G. *Gazz. Chim. Ital.* **1987**, *117*, 549–553.
(67) Benedetti, E.; Toniolo, C.; Hardy, P.; Barone, V.; Bavoso, A.; Di Blasio, B.; Grimaldi, P.; Lelj, F.; Pavone, V.; Pedone, C.; Bonora, G. M.; Lingham, I. *J. Am. Chem. Soc.* **1984**, *106*, 8146–8152.
(68) Benedetti, E.; Barone, V.; Bavoso, A.; Di Blasio, B.; Lelj, F.; Pavone, V.; Pedone, C.; Bonora, G. M.; Toniolo, C.; Leplawy, M. T.; Kaczmarek, K.; Redlinski, A. *Biopolymers* **1988**, *27*, 357–371.

the available FFs is their "intrinsic preference" for some conformational state (e.g., α-helix), irrespective of number of residues, nature of substituents at the C^{α}, and environmental effects. Several experimental studies have, however, unequivocally shown that the above parameters have a profound effect on the secondary structure of peptides and proteins.

All these discrepancies have been settled in our FF using a number of adjustable parameters lower than in other FFs. The keys to this success appear to be the physically based choice of potential energy functions, the use of softer nonbond interactions, and the re-examination of vicinal interactions. The above results and the reasonable trends obtained for the onset of different helical structures in representative polypeptides show that we have succeeded in obtain a well-balanced set of parameters. Furthermore, the flexibility of the parametrization leaves room for further improvements, especially concerning vibrational frequencies.

Acknowledgment. The authors acknowledge the Italian Research Council (Comitato Informatica) for financial support and Professor P. A. Kollman for a copy of Version 3.0 of the AMBER package.

Registry No. F-Gly-NH_2, 4238-57-7; Ac-Gly-$NHCH_3$, 7606-79-3; Ac-Ala-$NHCH_3$, 19701-83-8; Ac-Aib-$NHCH_3$, 42037-26-3; Ac-$(Ala)_{13}$-$NHCH_3$, 143142-48-7; Ac-$(Aib)_{13}$-$NHCH_3$, 143123-57-3.

Origin of the Regioselective Lithiation of 1,3-Disubstituted Heteroatom Aromatics. MNDO Evidence for Bidentate Complexation

José M. Saá,* Pere M. Deyá, Guillem A. Suñer, and Antoni Frontera

Contribution from the Departament de Química, Universitat de les Illes Balears, E-07071 Palma de Mallorca, Spain. *Received February 21, 1992*

Abstract: The regioselective lithiation of several prototype 1,3-disubstituted heteroatom aromatics, such as 1,3-dimethoxybenzene (DMB), 1,3-difluorobenzene (DFB), and 3-fluoroanisole (FA), has been studied by means of the semiempirical MNDO method. Calculations show clear-cut evidence for the intermediate formation of chelated species **6**, which formally derives from the bidentate coordination of the coordinatively unsaturated lithium base by the educt working as a "pair of tweezers". A very strong agostic interaction on C_2–H exists at this stage of the reaction coordinate, even though lithium atoms are formally pentacoordinated. In all cases studied the energy barriers corresponding to lithiation at C_2 by the bidentate coordination mode of approach are lower than those of the monodentate coordination mode leading either to lithiation at C_2 or C_6. Both the experimentally observed rate enhancement and regioselectivity are thus fully supported by the bidentate coordination mode of approach. These results can be satisfactorily explained by examining the "neighbor" and "non-neighbor" interactions being developed in reaching the alternative transition states. Thus, by partitioning the MNDO-calculated transition state total energy it can be recognized that five large, overall attractive, neighbor interactions (two O–Li, one C–H, and two C_2–Li) are being established in reaching the transition structure ($TS_{6\text{-}C2}$) by the bidentate coordination mode, but only three major interactions (one O–Li, one C–H, and one C_2–Li) develop throughout the monodentate coordination mode.

The so-called heteroatom-directed lithiation,[1] also termed ortho lithiation,[2] is a class of directed ortho metalations (DoM)[3] which allows for the regioselective introduction of lithium (hydrogen–lithium exchange) onto heteroatom-substituted aromatics. The lithiation of 1,3-disubstituted heteroatom aromatics, in particular, merits special mention because of its exceptional features. Not only does this reaction take place rapidly but, more importantly, it does so in a highly regioselective (C_2) manner. In other words, due to some poorly understood "cooperative effect",[3] lithiation at the C–H flanked by both heteroatom-based functional groups is kinetically and/or thermodynamically[4] favored over competitive lithiations at C_6 or C_4.

This large enhancing effect on rate shown by meta (relative to ortho or para) substituents such as fluoro, methoxyl, and dimethylamino was ascribed to the "coordinative involvement of the two moieties", from early work by Huisgen et al.[5] on the kinetics of the lithiation of substituted bromobenzenes. Unfortunately, no further kinetic data has been reported since then for other closely related compounds.[6] On other other hand, synthetic chemists have taken advantage of the high regioselective nature of the lithiation of this kind of substrates for the preparation of the difficult-to-synthesize 1,2,3-trisubstituted aromatics.[7]

The very recent work of Bauer and Schleyer[8] brought to light a comprehensive[9] mechanistic explanation for the formation of aromatic organolithium compounds by hydrogen–lithium exchange. On the basis of one-dimensional (^1H, ^{13}C, and ^6Li) and two-dimensional (HOESY ^6Li–^1H) NMR studies[10] as well as MNDO calculations, these authors proposed a generalized

(1) Gschwend, H. W.; Rodriguez, H. R. *Org. React. (N.Y.)* **1979**, *26*, 1.
(2) Excellent monographs and reviews on the subject are available. See, for example: (a) Wakefield, B. J. *The Chemistry of Organolithium Compounds*; Pergamon: Oxford, 1974. (b) Beak, P.; Snieckus, V. *Acc. Chem. Res.* **1982**, *15*, 306. (c) Wardell, J. L. In *Comprehensive Organic Chemistry*; Wilkinson, G., Stone, F. G. A., Abel, E. W., Eds.; Pergamon: Oxford, 1982; Vol. 1, p 57. (d) Narsimhan, N. S.; Mali, R. S. *Synthesis* **1983**, 957. (e) Beak, P.; Meyers, A. I. *Acc. Chem. Res.* **1986**, *19*, 356. (f) Narasisimhan, N. S.; Mali, R. S. *Top. Curr. Chem.* **1987**, *138*, 63. See also refs 1 and 3.
(3) Snieckus, V. *Chem. Rev.* **1990**, *90*, 879 and references therein.
(4) See: Ziegler, F. E.; Fruler, K. W. *J. Org. Chem.* **1976**, *41*, 1564.
(5) Huisgen, R.; Mack, W.; Herbig, K.; Ott, N.; Anneser, E. *Chem. Ber.* **1960**, *93*, 412.
(6) Only scattered, qualitative kinetic data is available for other 1,3-disubstituted aromatics. See, for example: Thornton, T. J.; Jarman, M. *Synthesis* **1990**, 295.
(7) For an extensive list of examples, see refs 1 and 3. For recent, interesting applications for the preparation of fluorosubstituted aromatics, see: Ladd, D. L.; Weinstock, J. *J. Org. Chem.* **1981**, *46*, 203. Adejare, A.; Miller, D. D. *Tetrahedron Lett.* **1984**, *25*, 5597.
(8) Bauer, W.; Schleyer, P. v. R. *J. Am. Chem. Soc.* **1989**, *111*, 7191.
(9) For previous studies, see: Ellison, R. A.; Kotsonis, F. N. *Tetrahedron* **1973**, *29*, 805; *J. Org. Chem.* **1973**, *38*, 4192. Graybill, B. M.; Shirley, D. A. *J. Org. Chem.* **1966**, *31*, 1221.
(10) Bauer, W.; Winchester, W. R.; Schleyer, P. v. R. *Organometallics* **1987**, *6*, 2371. Bauer, W.; Clark, T.; Schleyer, P. v. R. *J. Am. Chem. Soc.* **1987**, *109*, 970. Bauer, W.; Müller, G.; Pi, R.; Schleyer, P. v. R. *Angew. Chem., Int. Ed. Engl.* **1986**, *25*, 1103. See also ref 17.

0002-7863/92/1514-9093$03.00/0