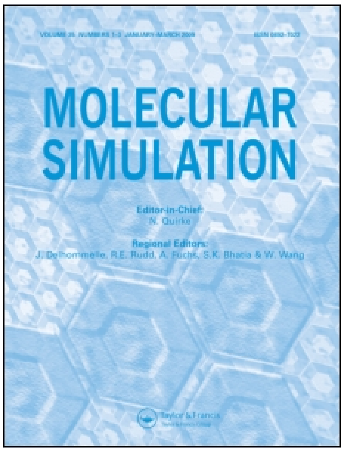

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Polarization Force Fields for Peptides Implemented in ECEPP2 and MM2

Francisco Torrens[a]
[a] Departament de Química Física, Facultat de Química, Universitat de Valéncia, València, Spain

To cite this Article Torrens, Francisco(2000) 'Polarization Force Fields for Peptides Implemented in ECEPP2 and MM2', Molecular Simulation, 24: 4, 391 — 410
To link to this Article: DOI: 10.1080/08927020008022384
URL: http://dx.doi.org/10.1080/08927020008022384

Molecular Simulation, 2000, Vol. 24, pp. 391–410
Reprints available directly from the publisher
Photocopying permitted by license only

Published by license under
the Gordon and Breach Science
Publishers imprint.
Printed in Malaysia.

POLARIZATION FORCE FIELDS FOR PEPTIDES IMPLEMENTED IN ECEPP2 AND MM2

FRANCISCO TORRENS

*Departament de Química Física, Facultat de Química, Universitat de València,
Dr. Moliner 50, E-46100-Burjassot (València), Spain*

(Received April 1999; accepted May 1999)

The empirical conformational energy program for peptides (ECEPP2) and molecular mechanics (MM2) have been used for the simulation of the For-Gly-NH_2 backbone. I propose two different methods for the calculation of the polarization energy term: the polarization procedure by non-interacting induced dipoles (NID) which assumes scalar isotropic point polarizabilities and the polarization scheme by interacting induced dipoles (ID) which calculates tensor effective anisotropic point polarizabilities (method of Applequist). I present a comparative study of ECEPP2 and MM2 + polarization. I discuss molecular mechanics results including the total energy differences, partitional analyses of the total *steric* energies and torsion dihedral angles. The γ global and the α, β and δ local minima are stabilized by intramolecular hydrogen bonds. Although ECEPP2-based calculations rather under or over-estimate the relative energy of some local minima, the ID polarization energy term represents a significant correction to the total relative energy.

Keywords: Polarization; polarizability; peptide models; amino acids; conformations; stable; molecular mechanics

ACS Classification Codes: Amino acids, peptides and proteins; biomolecules and their synthetic analogs; physical organic chemistry; general physical chemistry

INTRODUCTION

The shapes of peptides and proteins have been a central theme of scientific investigation for decades [1]. Several papers have been reported on the topological description of the potential energy surfaces (PES) of diamides (glycine [2–4], alanine [2–6], serine [7] and valine [3,4,8] derivatives) and their relevance to protein folding [9]. In this article, I deal with

N-formylglycinamide (For-Gly-NH_2), which is taken as a model for a single dipeptide [1]. I have selected this molecule as a test in my calculations for two reasons: (1) With the lack of a side chain, PES $E(\phi, \psi)$ is a function of only two variables. (2) It has been established from first principles that the idealized 2D-Ramachandran PES for a diamide has to have nine minima (α_L, α_D, β, γ_L, γ_D, δ_L, δ_D, ε_L and ε_D). Since glycine is achiral, instead of 9 only 5 unique conformations can occur (*i.e.*, $\alpha_L = \alpha_D$, $\gamma_L = \gamma_D$, $\delta_L = \delta_D$ and $\varepsilon_L = \varepsilon_D$); see Figure 1.

The knowledge of electronic polarizabilities describing the distortion of a molecular electronic cloud by an external field is very important to describe intermolecular energies (induction term). In addition, the intramolecular force fields used in mechanical computations are usually based on the assumption that the interaction between non-bonded atoms can be analyzed in the same way as intermolecular interactions. Therefore, a possible improvement of modern force fields is expected from the introduction of an induction term. Unfortunately accurate experimental data on the electronic polarizabilities of molecules are very rare, especially if one is interested by a whole tensorial quantity instead of the usual mean value. It seems that modern theoretical chemistry, with the help of modern

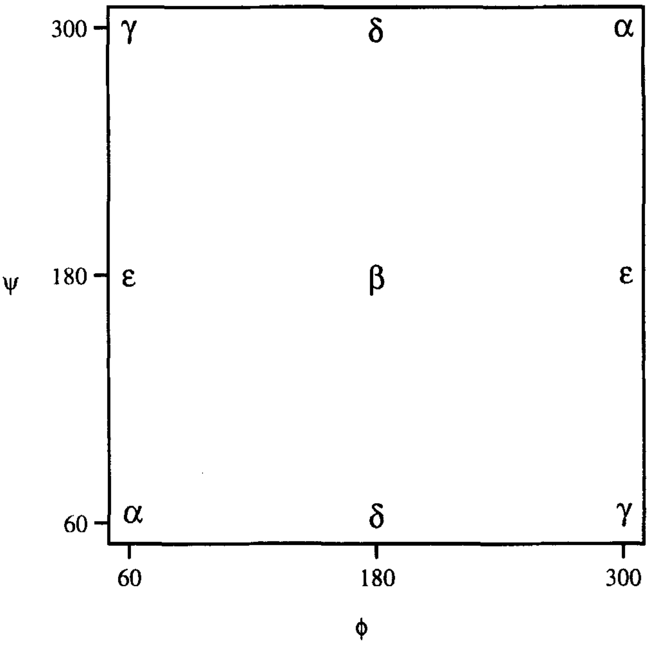

FIGURE 1 A schematic Ramachandran type map, $E = E(\phi, \psi)$, for For-Gly-NH_2 indicating the idealized locations of the expected five minima.

computers is now able to become a substitute of difficult experimental determinations.

For the implementation of the induction term in the interaction potential between small molecules, one assumes that the molecular polarizabilities are known. Molecular polarizabilities are the tensors that allow to compute several moments induced in a molecule when it is placed in a non-homogeneous electric field, which can be expanded in a point using the field at this point and its several derivatives. When passing from small molecules to macromolecules and from inter- to intramolecular interactions, one can assume that the atomic or bond increments (*i.e.*, atomic or bond polarizabilities) are available. These increments allow one to divide the global effect into a sum of local effects. With regard to the easy-handle calculation of large molecules, I have chosen the atomic increments.

In previous articles, I implemented an induction term in the MM2 force field [10–12] for the study of dipeptide models [13] and applied it to the study of the conformational aspects of some asymmetric Diels-Alder reactions [14] and dipeptide models (For-Gly-NH_2) [15]. In the present work, I describe the implementation of the induction term in the ECEPP2 force field [16–18] and present a comparative study of both force fields applied to For-Gly-NH_2.

In the next section I present the main features of the implementation of the polarization method in the ECEPP2 force field. Following that, the electric polarization energy models are presented. Next, the interacting induced dipoles polarization model (method of Applequist) is described. The next section is devoted to the results for For-Gly-NH_2 and their discussion. A comparative study of ECEPP2 and MM2 calculations is presented and the results are compared with *ab initio* CPHF calculations. The last section summarizes my conclusions.

METHOD

In my studies of the For-Gly-NH_2 backbone conformation I have used the molecular mechanics methods [10]. Molecular mechanics techniques aim at the description of macroscopic properties of physical systems in terms of microscopic particle interactions. There are many force field programs available today. The most frequently used method in organic chemistry is MM2 [10–12]. In a previous article, I applied the MM2+polarization force field to the study of For-Gly-NH_2 [15].

Another frequently used molecular mechanics method for peptides is the empirical conformational energy program for peptides version 2 (ECEPP2)

[16]. This force field describes the molecular *steric* energy as a sum of the electrostatic, non-bonded, torsional, cystine torsional and loop-closing energy components.

In order to obtain exact relative energy differences between conformations, it is necessary to optimize the geometry from appropriate starting dihedral angles to give the *gauche* and *trans* minima [19]. For this purpose I have implemented the SIMPLEX MS-3 optimization algorithm [20–26] in program ECEPP2.

The dipole moment $\bar{\mu}$ has been calculated from the point (atomic) distribution of net charges taken from the ECEPP2 database [16].

ELECTRIC POLARIZATION ENERGY MODELS

The appropriate treatment of non-additive many-body forces, in particular the inductive interactions in polar polarizable molecules, has been a long-standing problem in computer simulations of molecular conformations [27]. The inductive interactions are modelled by interacting atomic dipole polarizabilities. In energy minimization procedures one needs to know the forces acting on the molecules or molecular parts. A direct approach to the analytical evaluation of the forces and force constants, by calculating the variation of the total electrostatic energy and of the forces under small virtual displacements of the interaction centres has been proposed. An important advantage of the calculation of the forces by considering the variation of the energy expression with respect to the displacement of particles is that one can discuss those cases where the induced moments are not calculated self-consistently. In effect in such circumstances there are possible dangers of using inconsistent approximate energy and force expressions.

The familiar effective pairwise additive interaction potentials used in most molecular simulations have been extensively refined, and several internally consistent parameter sets are available [12, 19, 28–31]. This is not the case with model potentials that explicitly include polarization. Consider a system of N polarizable point charges $\{q^i, \alpha^i\}$, where q^i represent point charges and α^i scalar isotropic point polarizabilities in the absence of an external electric field [32–35]. (I consider here only point charges as the simplest case; generalization of the formulae for higher permanent moments is straightforward). The local field at position i, $\bar{E}^i$, introduces an induced dipole $\bar{\mu}^i$ in the direction of the local field,

$$\bar{\mu}^i = \alpha^i \cdot \bar{E}^i$$

This local electric field is due to the electric field due to all charges except q^i and the electric field due to all induced dipoles except $\bar{\mu}^i$. It should be noted that the local field is calculated only from the charges at 1–4 (and further) atoms because I am interested in the correction of only the torsion component of the ECEPP2 and MM2 force fields. The electric polarization energy U is the energy required to make the induced dipoles,

$$U = \sum_i^N U^i = -\frac{1}{2}\sum_i^N \frac{\bar{\mu}^i \bullet \bar{\mu}^i}{\alpha^i}$$

The explicit treatment of the induced dipole moments, as far as they are generated by the electric field due to the permanent charges and dipoles of the molecule, has been suggested [13]. Two different methods for the calculation of the effect of the induced dipole moments on the polarization energy term have been proposed [15]: the polarization procedure by non-interacting induced dipoles (NID), and the polarization scheme by interacting induced dipoles (ID).

The NID polarization method assumes scalar isotropic point (atomic) polarizabilities. Since the electric field at each position has to be evaluated to determine the induced dipoles, the polarization energy is most easily evaluated from this field,

$$U = -\frac{1}{2}\sum_i^N \alpha^i \cdot \bar{E}^i \bullet \bar{E}^i$$

The most important aspect by which my treatment differs from earlier methods is that the local electric field from which the induced dipole moments are evaluated does not depend on the polarizabilities $\{\alpha^i\}$. On the other hand, the polarization procedure by interacting dipoles (ID) allows the interaction of the induced dipole moments by means of tensor effective anisotropic point polarizabilities (method of Applequist [36–40]).

Both NID and ID algorithms have been implemented in the ECEPP2 and MM2 programs. A database of atomic polarizabilities [41] has been implemented in the *block data* of both programs.

The atomic polarizabilities used (NID methods) and obtained (ID procedures) in the calculation of the polarization energy term of For-Gly-NH_2 are reported in Table I. On one hand, for ECEPP2 the total molecular polarizabilities are greater for the ID calculations than for the NID results. Moreover, the atomic polarizabilities of the H, C and N atoms are greater with ID. However, the atomic contributions from the O atom are greater

TABLE I Atomic polarizabilities ($\AA^3$) used (ECEPP2 + NID)[a] and obtained (ECEPP2 + ID)[b] in the calculation of the polarization energy for For-Gly-NH_2 conformations

	ECEPP2 + NID		ECEPP2 + ID					MM2 + NID	MM2 + ID	
Residue	*Atom*	α-γ	α	β	γ	δ	ε	β, γ	β	γ
Formyl	H	0.407	1.666	1.676	1.665	1.669	1.669	0.407	1.671	1.665
	O	1.395	0.361	0.369	0.362	0.366	0.361	1.395	0.364	0.362
	C	0.075	0.635	0.642	0.635	0.636	0.634	0.075	0.636	0.634
Glycine	N	0.628	1.056	1.050	1.052	1.054	1.050	2.255	0.212	0.211
	HN	0.092	0.100	0.100	0.099	0.101	0.100	0.010	0.162	0.162
	CA	1.027	1.349	1.353	1.349	1.351	1.351	1.027	1.349	1.345
	HA	0.407	1.698	1.716	1.713	1.710	1.720	0.407	1.684	1.685
	HA	0.407	1.688	1.716	1.686	1.699	1.697	0.407	1.684	1.676
	C′	0.075	0.628	0.628	0.628	0.628	0.628	0.075	0.628	0.627
	O	1.395	0.357	0.358	0.357	0.360	0.358	1.395	0.357	0.356
Carboxyl-NH_2	N	0.628	1.057	1.073	1.059	1.055	1.060	2.255	0.213	0.211
	H2	0.092	0.098	0.102	0.099	0.098	0.101	0.010	0.163	0.159
	H2	0.092	0.097	0.099	0.097	0.098	0.098	0.010	0.159	0.157
Total	–	6.721	10.792	10.883	10.802	10.824	10.827	9.729	9.281	9.250

[a] NID: polarization by non-interacting induced dipoles.

[b] ID: polarization by interacting induced dipoles.

with NID. When the five α-ε ID results are compared, similar atomic and total molecular polarizabilities are obtained. On the other hand, for MM2 the total molecular polarizabilities are greater for NID than for ID. Furthermore, the atomic polarizabilities of the N and O atoms are greater with NID. However, the atomic contributions from the H and C atoms are greater with ID. When both β and γ ID results are compared, similar total molecular polarizabilities are obtained. The comparison of ECEPP2 and MM2 + polarization calculations show that the dispersion of the results between both NID methods (3 Å^3) is halved when the ID polarization is used.

INTERACTING INDUCED DIPOLES POLARIZATION MODEL (METHOD OF APPLEQUIST)

One considers the molecule as being up of N atoms (represented by $i, j, k, \ldots$), each of which acts as a point particle located at the nucleus, and responds to an electric field only by the induction of a dipole moment, which is a linear function of the local field [39, 40]. If a Cartesian coordinate of the field due to the permanent multipole moments is E_a^i, then the induced moment μ_a^i in atom i is

$$\mu_a^i = \alpha^i \cdot \left(E_a^i + \sum_{j(\neq i)}^{N} T_{ab}^{ij} \cdot \mu_b^j \right) \tag{1}$$

where α^i is the polarizability of atom i and T_{ab}^{ij} is the symmetrical field gradient tensor, $T_{ab}^{ij} = (1/e) V_a^i \cdot E_b^j$, e is the charge of the proton, and the subscripts $a, b, c, \ldots$ are standing for the Cartesian components x, y, z. The expression in parentheses in Eq. (1) is the total electric field at atom i, consisting of the external field plus the fields of all the other induced dipoles in the molecule.

The set of coupled linear equations (1) for the induced dipole moments can conveniently be expressed in compact matrix equation form, if one introduces the $3N \times 3N$ matrices $\bar{\bar{T}}$ and $\bar{\bar{\alpha}}$, with elements T_{ab}^{ij} and $\alpha_{ab}^i \cdot \delta^{ij}$ (δ^{ij} being the Kronecker δ), respectively. In order to supress the restriction in the sum, the elements T_{ab}^{ii} are defined as zero. Similarly $\bar{E}$ and $\bar{\mu}$ are $3N \times 1$ column vectors with elements E_a^i and μ_a^i. Equation (1) is thus written in matrix form,

$$\bar{\mu} = \bar{\bar{\alpha}} \cdot \left(\bar{\bar{I}} \cdot \bar{E} + \bar{\bar{T}} \cdot \bar{\mu} \right) = \bar{\bar{\alpha}} \cdot \bar{\bar{I}} \cdot \bar{E} + \bar{\bar{\alpha}} \cdot \bar{\bar{T}} \cdot \bar{\mu}$$

where $\bar{\bar{I}}$ is the $3N \times 3N$-dimensional unit matrix. This matrix equation can be solved for the induced dipole as

$$\bar{\mu} = \left(\bar{\bar{I}} - \bar{\bar{\alpha}} \cdot \bar{\bar{T}}\right)^{-1} \cdot \bar{\bar{\alpha}} \cdot \bar{E} = \bar{\bar{A}} \cdot \bar{E}$$

Here the symmetrical *many-body* polarizability matrix $\bar{\bar{A}}$ has been introduced:

$$\bar{\bar{A}} = \left(\bar{\bar{I}} - \bar{\bar{\alpha}} \cdot \bar{\bar{T}}\right)^{-1} \cdot \bar{\bar{\alpha}}$$

It is important to emphasize that the many-body polarizability matrix describes a non-local polarizability response of the system. In effect, an electrostatic field at point j gives rise not only to an induced moment at j (proportional to the diagonal block element A_{ab}^{jj}), but it contributes (*non-locally*) to the induced moments of all the other sites as well.

CALCULATION RESULTS AND DISCUSSION

I have calculated the five minima of the conformational PES of For-Gly-NH_2 with a version of the ECEPP2 program which includes, besides the standard algorithm, both polarization methods (NID and ID). These minima were described by Perczel *et al.* [2] at the ECEPP2 level, using the grid method for the geometry optimizations [42]. These minima were also described by Perczel *et al.* [2] and Viviani *et al.* [3, 4] at the *ab initio* level using second derivative techniques for the geometry optimizations [43, 44]. These conformations are shown in Figure 2. The α, γ, δ and ε minima are *folded* conformations while the β minimum is fully planar. The ECEPP2 + polarization calculations have been optimized with the SIMPLEX MS-3 algorithm [22].

The total energy differences are shown in Table II and compared with MM2 molecular mechanics and *ab initio* SCF 3-21G (optimized geometry) references [3, 4]. Five structures are found with the ECEPP2 methods, two with the MM2 methods and four with the *ab initio* method. The three types of methods show at the same time only two structures (β and γ). These are the only minima with MM2 and the two main minima with ECEPP2 and *ab initio*. The ECEPP2 + polariation relative energies of the β local minimum are in agreement with the reference calculations, lying between the MM2 + ID and *ab initio* values.

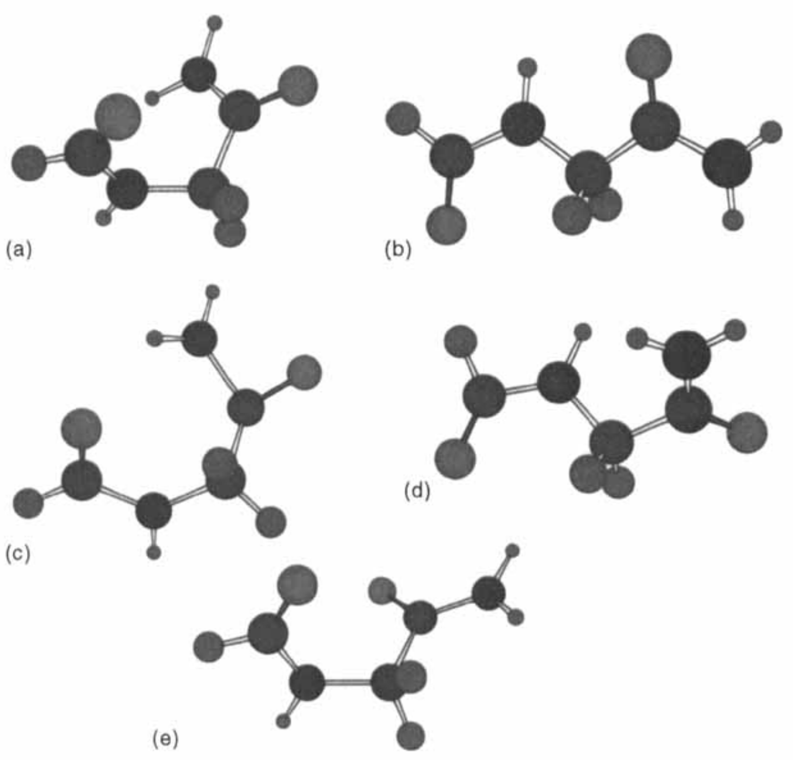

FIGURE 2 Molecular mechanics (ECEPP2) structure for For-Gly-NH_2. (a) α minimum, (b) fully-planar β minimum, (c) γ minimum, (d) δ minimum and (e) ε minimum. The α, γ, δ and ε minima are *folded* conformations. (See Color Plate XVIII).

It is well known that intramolecular hydrogen bonds contribute to stabilization of the γ conformers independently of the nature of the side chain [2, 5]. The variation of this contribution to the stability of the various γ conformers may be anticipated from the variations of the length of the hydrogen bond which occurs between a hydrogen atom bonded to the terminal nitrogen and the oxygen of the carboxylic group at the other end. The formation of this hydrogen bond between both ends of the backbone results in the formation of a cycle in which the optimum hydrogen bond length is modified by the angular tension.

The α and β local minima are stabilized by one intramolecular hydrogen bond forming a 5-membered ring $N—H\cdots N$ (α) or $N—H\cdots O$ (β). Moreover, the γ global and δ local minima show two shared intramolecular hydrogen bonds forming a 5-membered ring $N—H\cdots N$ and closing a ring with 7 members $N—H\cdots O$ (γ), or forming two shared 5-membered

TABLE II Molecular mechanics (ECEPP2) results for For-Gly-NH_2 conformations. Number of hydrogen bonds and total energy differences in $kJ \cdot mol^{-1}$

Conformation	*No. of H-bonds*	*ECEPP2*	*ECEPP2 + NID*[a]	*ECEPP2 + ID*[b]	*MM2*	*MM2 + NID*	*MM2 + ID*	*Reference*[c]
α	1	6.3	6.1	0.2	–[d]	–[d]	–[d]	18.6
β	1	6.1	7.0	9.2	21.5	22.0	16.2	2.6
γ	2 (shared)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
δ	2 (shared)	6.9	7.9	5.9	–[d]	–[d]	–[d]	13.7
ε	0	9.2	6.3	9.0	–[d]	–[d]	–[d]	–[d]

[a] NID: polarization by non-interacting induced dipoles.

[b] ID: polarization by interacting induced dipoles.

[c] Reference: *ab initio* SCF 3 – 21G (optimized geometry) taken from Ref. [2].

[d] A dash (–) indicates no local minimum for this conformation.

rings N—H···N (δ). On the other hand, the ε local minimum shows no intramolecular hydrogen bond. It should be pointed out that the energy differences among minima are expected to be smaller for For-Gly-NH_2 than for the remaining For-X-NH_2 amino acid dipeptide models (X stands for any amino acid) because glycine (Gly) is the most flexible amino acid (note that Gly-Gly-X-Gly-Gly pentapeptides are used as models of random coil arrangements around amino acid X in nuclear magnetic resonance) [45].

The partitional analyses of the ECEPP2 total *steric* energies for the five conformers are reported in Table III. No contribution of the torsional, nor cystine torsional nor loop-closing energy components has been observed. The γ global minimum is stabilized with respect to the remaining minima by the electrostatic energy term. It should be remarked that the polarization energy term is rather big when compared with the electrostatic and non-bonded component parts and with the total steric energy. The ECEPP2 + NID polarization energy represents *ca.* 20% of the total energy and the ECEPP2 + ID induction energy represents *ca.* 44% of the total energy. For MM2, the MM2 + NID polarization component is *ca.* 9% and the MM2 + ID polarization is *ca.* 16% of the total energy. Note that both total energies and their component parts cannot be compared directly between different force fields because only the relative total energy differences have physical meaning [10].

The dipole moments μ for the five conformers are reported in Table IV. The ECEPP2 + polarization μ values compare better than the MM2 + polarization ones with the *ab initio* reference calculations. ECEPP2 + polarization rather underestimates μ while MM2 + polarization overestimates them. The dispersion of the results is smaller for ECEPP2 + polarization because in this version of the ECEPP2 methods, the bond distances and angles are frozen during the optimization process, and only torsions around the dihedral angles are allowed [17]. In particular, the μ results for the β fully-planar local minimum show a constant value independently of the inclusion of the polarization energy terms in the force field. It should be remarked that, for both types of force fields, the effect of polarization corrects the dipole moments in the direction of the reference calculations, slightly increasing them for ECEPP2 and rather decreasing them for MM2.

The torsion dihedral angles (ϕ and ψ) for both families of force fields has been reported in Table V. The effect of ECEPP2 + polarization on ϕ and ψ is rather big (*e.g.*, 28 – 47° for local minimum δ). The effect on the remaining geometric parameters has a small individual effect on the chain

TABLE III Molecular mechanics (ECEPP2) results for For-Gly-NH_2 conformations. Partitional analyses of the total *steric* energies in kJ · mol^{-1}[a]

Conformation[b]	*Method*	*Electrostatic*	*Non-bonded*	*Polarization*	*Total*
α	ECEPP2	−21.8	−9.9	–	−31.7
	ECEPP2 + NID[c]	−22.0	−9.6	−8.3	−39.9
	ECEPP2 + ID[d]	−21.4	−10.1	−30.0	−61.5
β	ECEPP2	−23.4	−8.5	–	−31.9
	ECEPP2 + NID[c]	−23.4	−8.5	−7.0	−38.9
	ECEPP2 + ID[d]	−23.3	−8.5	−20.6	−52.4
γ	ECEPP2	−28.6	−9.4	–	−38.0
	ECEPP2 + NID[c]	−28.8	−9.2	−8.0	−46.0
	ECEPP2 + ID[d]	−28.9	−9.0	−23.8	−61.7
δ	ECEPP2	−22.6	−8.5	–	−31.1
	ECEPP2 + NID[c]	−22.7	−8.3	−7.0	−38.0
	ECEPP2 + ID[d]	−22.3	−8.7	−24.8	−55.8
ε	ECEPP2	−21.5	−7.3	–	−28.8
	ECEPP2 + NID[c]	−20.9	−7.8	−11.0	−39.7
	ECEPP2 + ID[d]	−20.3	−7.9	−24.5	−52.7

BB[b]	*Method*	*Stretching*	*Bending*	*Stretching – Bending*	*van der Waals*	*Torsion*	*Electrostatic*	*Polarization*	*Total*
β	MM2	1.2	3.5	−0.2	−9.3	8.3	−89.5	–	−86.2
	MM2 + NID[c]	1.2	3.4	−0.2	−9.0	8.3	−89.7	−8.7	−94.8
	MM2 + ID[d]	1.2	3.4	−0.2	−9.1	8.3	−89.5	−20.4	−106.3
γ	MM2	1.3	2.4	−0.1	−10.7	−3.4	−97.2	–	−107.7
	MM2 + NID[c]	1.3	2.2	−0.1	−9.9	−3.3	−97.8	−9.2	−116.8
	MM2 + ID[d]	1.3	2.2	−0.1	−10.1	−3.3	−97.4	−15.1	−122.5

[a] The torsional energy component is always below 0.1 kJ · mol^{-1} and is not shown.
[b] BB: backbone conformation.
[c] NID: polarization by non-interacting induced dipoles.
[d] ID: polarization by interacting induced dipoles.

TABLE IV Molecular mechanics (ECEPP) results for For-Gly-NH_2 conformations. Dipole moments in Debyes

Conformation	*ECEPP2*	*ECEPP2 + NID*[a]	*ECEPP2 + ID*[b]	*MM2*	*MM2 + NID*[a]	*MM2 + ID*[b]	*Reference*[c]
α	4.684	4.590	4.810	–[d]	–[d]	–[d]	6.560
β	3.048	3.048	3.048	7.694	6.268	4.717	3.360
γ	0.947	0.985	1.003	6.938	4.280	5.809	3.540
δ	3.433	3.290	3.661	–[d]	–[d]	–[d]	4.910
ε	3.573	3.866	4.188	–[d]	–[d]	–[d]	5.470[e]

[a] NID: polarization by non-interacting induced dipoles.

[b] ID: polarization by interacting induced dipoles.

[c] Reference: *ab initio* SCF 3–21G (optimized geometry) taken from Ref. [4].

[d] A dash (–) indicates no local minimum for this conformation.

[e] In order to evaluate the above quantity for the missing ε minimum, its geometry was assumed to be equivalent to the geometry of the optimized corresponding conformation of alanine.

TABLE V Molecular mechanics (ECEPP) results for For-Gly-NH_2 conformations. Torsion dihedral angles (ϕ, ψ) in degrees

	ϕ						
Conformation	*ECEPP2*	*ECEPP2 + NID*[a]	*ECEPP2 + ID*[b]	*MM2*	*MM2 + NID*[a]	*MM2 + ID*[b]	*Reference*[c]
α	−72.5	−73.0	−70.2	–[d]	–[d]	–[d]	−83.2
β	180.0	180.0	180.0	−179.8	−179.9	−179.9	180.0
γ	−78.9	−78.2	−77.9	−80.7	−81.4	−81.3	−83.9
δ	−165.9	−161.4	−173.1	–[d]	–[d]	–[d]	−126.0
ε	−77.1	−78.7	−76.6	–[d]	–[d]	–[d]	–[d]

	ψ						
Conformation	*ECEPP2*	*ECEPP2 + NID*[a]	*ECEPP2 + ID*[b]	*MM2*	*MM2 + NID*[a]	*MM2 + ID*[b]	*Reference*[c]
α	−35.1	−31.3	−40.8	–[d]	–[d]	–[d]	−14.2
β	180.0	180.0	180.0	179.9	180.0	179.9	180.0
γ	74.1	73.2	72.7	48.8	49.6	49.9	67.8
δ	53.3	52.7	52.9	–[d]	–[d]	–[d]	25.5
ε	154.2	164.2	171.5	–[d]	–[d]	–[d]	–[d]

[a] BB: backbone conformation.
[b] NID: polarization by non-interacting induced dipoles.
[c] ID: polarization by interacting induced dipoles.
[d] Reference: *ab initio* SCF 3–21G (optimized geometry) taken from Ref. [2].
[e] A dash (−) indicates no local minimum for this conformation.

conformation because the amide bonds are known to be planar around the sp^2 nitrogen atoms. Moreover, the stretching and bending movements are more rigid than the dihedral angle changes, and carry out shorter movements. However, the combined effect on the overall conformation of the changes associated to the 33 conformational degrees of freedom corresponding to For-Gly-NH_2, is difficult to predict.

CONCLUSIONS

From the preceding discussion the following conclusions can be drawn.

1. The results of the present work clearly indicate that due to the creation and annihilation of critical points, caused by excessive repulsive or excessive attractive interactions, it may become necessary to recalibrate empirical force fields. It appears that the results of good-quality *ab initio* calculations might be suitable as primary standards for such a calibration. At least in the present study for For-Gly-NH_2, molecular mechanics was able to predict, without and with the inclusion of the polarization energy term, the global minimum of the conformational PES. It is also clear that, in order to see at least a good portion of the total picture, the *N*-formylamides of all naturally occurring L-amino acids must be studied by *ab initio* methods as well as the amides of dipeptides, tripeptides and tetrapeptides.
2. The comparison of ECEPP2 and MM2 + polarization calculations for For-Gly-NH_2 shows that ECEPP2 gives a better description of the global conformational PES, while in MM2 the inclusion of the polarization energy gives a better description of the relative energy difference between the two main minima.
3. The polarization energy term has not been revealed necessary in the ECEPP2 geometry optimization for For-Gly-NH_2, but it influences the energy differences between the conformers. In particular, this difference for only the polarization component is 44% of the total energy difference, resulting of the same order of magnitude that the energy differences reported in the *ab initio* reference calculations. An extension of the present study to alanine, serine and valine analogues of For-Gly-NH_2 would give an insight into a possible generality of this conclusion. For these *N*-formylamides the changes of geometry due to polarization are expected to be even smaller because all the amino acids are more rigid than glycine [45].

4. I want to deal with the analytical calculation of the overall conformational change induced by the electric polarization energy. Further work is in progress dealing with the calculation of the changes induced by polarization in the molecular and atomic indices of the For-Gly-NH_2 conformers. This work will take advantage of the capabilities of the TOPO program [46–48] written for the simulation of the molecular shape through the calculation of electrostatic and geometric descriptors and topological indices of molecules, including the fractal dimension of the accessible surface.

Acknowledgement

The author acknowledges financial support of the Dirección General de Enseñanza Superior of the Spanish MEC (Project PB97-1383).

APPENDIX

Hamiltonians

In the MM2 + polarization methods, the geometry optimization is carried out with the *block diagonal* Newton-Raphson *method* of the program MM2 [10]. The diagonal of the Hessian is calculated analytically [49]. In this work, I have used the MM2 (1980) set of parameters [12]. The parameters for hydrogen bonding and for amides, polypeptides and proteins have been included according to Refs. [28] and [30], respectively. For both NID and ID polarization formulations, a scheme for the analytical polarization energy, gradient and Hessian has been developed based on the electric field value, its gradient and Hessian [13]. The polarization interactions are taken into account only for atoms located at position 4 (and further) referred to bonds between atoms placed at 1–2. A database of atomic polarizabilities has been built up from coupled Hartree-Fock (CPHF) calculations of the polarizability tensors of a series of molecules containing the main functional groups which are present in peptides [41].

In the ECEPP2 + polarization methods, the energy minimization is performed with the SIMPLEX MS-3 optimization algorithm [22], which has been implemented in the code for this work. The original parameter set (ECEPP version 2) [17] has been used along the calculations. The polarization interactions are considered only for atoms located at 1–4 (and further).

Algorithms

FORTRAN programs MMID2 [13–15] and ECEPPID2 include the polarization procedure by non-interacting induced dipoles (NID) and the polarization scheme by interacting included dipoles (ID). Both programs are available from the author and are free for academics.

Expressions

The expression for both NID and ID polarization schemes are given in the sections "Electric polarization energy models" and "Interacting induced dipoles polarization model (method of Applequist)", respectively.

Validation

Results of checks undertaken to validate the simulation were published for For-Gly-NH_2 with MM2 + NID [3] and MM2 + ID [15]. Also, for Diels-Alder dienophiles, with MM2 + ID [14]. Reference calculations for For-Gly-NH_2 with ECEPP2 were published by Perczel *et al.* [2].

The SIMPLEX MS-3 [20–22] optimization algorithm was tested in the program AMYR for the theoretical simulation of molecular associations and reactions that was applied to benzene dimers [50, 51], *s*-terazine dimers, benzene-*s*-tetrazine dimers [51], azine dimers [52, 53], phthalocyanine (dimers [54–57], clusters [57] and *β*-crystal [57, 58]) and coronene clusters [59, 60]).

Other Details

One only molecule (For-Gly-NH_2) has been presented in this work in 5 different conformations. The CPU time for a typical run is 4s for MMID2 and 9s with ECEPPID2 in an IBM RISC System/6000 590. Starting geometries were taken from the *ab initio* and ECEPP2 optimized geometries from Refs. [2] to [4]. An improved DRIVE optimization scheme has been implemented in the MMID2 program. Details and results are reported in Ref. [15].

References

[1] Pullman, B. and Pullman, A. (1974). *Advances in Protein Chemistry*, Anfinsen, C. B., Edsall, J. T. and Richards, F. M. Eds., Academic Press, New York, 1974, **28**, 347–526.

[2] Perczel, A., Ángyán, J. G., Kajtár, M., Viviani, W., Rivail, J.-L., Marcoccia, J.-F. and Csizmadia, I. G. (1991). "Peptide models. 1. Topology of selected peptide conformational potential energy surfaces (glycine and alanine derivates)", *J. Am. Chem. Soc.*, **113**, 6256.

[3] Viviani, W., Rivail, J.-L. and Csizmadia, I. G. (1991). "*Ab-initio* SCF calculations on conformational space of peptides models", *J. Chim. Phys. Phys.-Chim. Biol.*, **88**, 2505.

[4] Viviani, W., Rivail, J.-L. and Csizmadia, I. G. (1993). "Peptide models II. Intramolecular interactions and stable conformations of glycine, alanine, and valine peptide analogues", *Theor. Chim. Acta (Berlin)*, **85**, 189.

[5] Head-Gordon, T., Head-Gordon, M., Frisch, M. J., Brooks III, C. and Pople, J. (1989). "A theoretical study of alanine dipeptide and analogs", *Int. J. Quantum Chem., Quantum Biol. Symp.*, **16**, 311.

[6] Head-Gordon, T., Head-Gordon, M., Frisch, M. J., Brooks III, C. and Pople, J. (1991). "Theoretical study of blocked glycine and peptide analogues", *J. Am. Chem. Soc.*, **113**, 5989.

[7] Perczel, A., Daudel, R., Ángyán, J. G. and Csizmadia, I. G. (1990). "A study on the backbone/side-chain interaction in *N*-formyl-(L)-serineamide", *Can. J. Chem.*, **68**, 1882.

[8] Viviani, W., Rivail, J.-L., Perczel, A. and Csizmadia, I. G. (1993). "Peptide models. 3. Conformational potential energy hypersurface of formyl-L-valinamide", *J. Am. Chem. Soc.*, **115**, 8321.

[9] Perczel, A., Rivail, J.-L. and Csizmadia, I. G. (1992). "Peptide conformational potential energy surfaces and their relevance to protein folding", In: *The Role of Computational Models and Theories in Biotechnologies*, Bertrán, J. Ed., Kluwer, Dordrecht, pp. 28–76.

[10] Burkert, U. and Allinger, N. L. (1982). *Molecular Mechanics*, ACS Monograph No. 177, American Chemical Society, Washington D.C.

[11] Allinger, N. L. (1977). "Conformational analysis. 130 MM2. A hydrocarbon force field utilizing V_1 and V_2 torsional terms", *J. Am. Chem. Soc.*, **99**, 8127.

[12] Allinger, N. L. and Yuh, Y. H. (1980). "MM2", QCPE Program No. 395.

[13] Torrens, F., Voisin, C. and Rivail, J.-L. (1991). "Electric polarization in a force field for the study of dipeptide models", In: *Computing Methods in Applied Sciences and Engineering*, Glowinski, R. Ed., Nova Science, New York, pp. 249–258.

[14] Torrens, F., Ruiz-López, M., Cativiela, C., García, J. I. and Mayoral, J. A. (1992). "Conformational aspects of some asymmetric Diels-Alder reactions", A molecular mechanics + polarization study", *Tetrahedron*, **48**, 5209.

[15] Torrens, F., Sánchez-Marín, J. and Rivail, J.-L. (1994). "Interacting induced dipoles polarization in a force field for dipeptide models (glycine derivative)", *An. Física (Madrid)*, **90**, 197.

[16] Némethy, G., Pottle, M. S. and Scheraga, H. A. (1983). "Energy parameters in polypeptides. 9. Updating of geometrical parameters, nonbonded interactions, and hydrogen bond interactions for the naturally occurring amino acids", *J. Phys. Chem.*, **87**, 1883.

[17] Némethy, G., Pottle, M. S. and Scheraga, H. A. (1983). "ECEPP2", *QCPE Program No. 454*.

[18] Sippl, M. J., Némethy, G. and Scheraga, H. A. (1984). "Intermolecular potentials from crystal data 6. Determination of empirical potentials for $O—H \cdots O=C$ hydrogen bonds from packing configuration configurations", *J. Phys. Chem.*, **88**, 6231.

[19] Abraham, R. J. and Grant, G. H. (1988). "A molecular mechanics study of the Si—O bond and alkyl-silanes", *J. Comput. Chem.*, **9**, 709.

[20] Deming, S. N. and Morgan, S. L. (1973). "Simplex optimization of variables in analytical chemistry", *Anal. Chem.*, **45**, 278-A.

[21] Morgan, S. L. and Deming, S. N. (1974). "Simplex optimization of analytical chemical methods", *Anal. Chem.*, **46**, 1170.

[22] Routh, M. W., Swartz, P. A. and Denton, M. B. (1977). "Performance of the Super Modified Simplex", *Anal. Chem.*, **49**, 1422.

[23] Leggett, D. J. (1983). "Instrumental simplex optimization", *J. Chem. Educ.*, **60**, 707.

[24] Walters, F. H., Parker, L. J. Jr., Morgan, S. L. and Deming, S. N. (1991). *Sequential Simplex Optimization*, CRC, Boca Raton.

[25] Deming, S. N. and Morgan, S. L. (1988). "Simplex-V: An interactive computer program for experimental optimization", Statistical Programs, Houston.

[26] Walters, F. H., Parker, L. J. Jr., Morgan, S. L. and Deming, S. N. (1999). "MultiSimplex 2.0/98", MultiSimplex, Karlskrona.

[27] Ángyán, J. G., Colonna-Cesari, F. and Tapia, O. (1990). "Analytical first and second energy and derivatives in the polarization model", *Chem. Phys. Lett.*, **166**, 180.

[28] Allinger, N. L., Kok, R. A. and Imam, M. R. (1988). "Hydrogen bonding in MM2", *J. Comput. Chem.*, **9**, 591.

[29] Lii, J. H., Galliot, S., Bender, C., Wikström, H., Allinger, N. L., Flurchick, K. M. and Teeter, M. M. (1989). "Molecular mechanics (MM2) calculations on peptides and on the protein crambin using the Cyber 205", *J. Comput. Chem.*, **10**, 503.

[30] Lii, J. H. and Allinger, N. L. (1991). "The MM3 force field for amides, polypeptides and proteins", *J. Comput. Chem.*, **12**, 186.

[31] Burkert, U. and Allinger, N. L. (1982). "Pitfalls in the use of the torsion angle driving method for the calculation of conformational interconversions", *J. Comput. Chem.*, **3**, 40.

[32] Straatsma, T. P. and McCammon, J. A. (1990). *Mol. Simul.*, **5**, 181.

[33] Straatsma, T. P. and McCammon, J. A. (1990). "ARGOS, a vectorized general molecular dynamics program", *J. Comput. Chem.*, **11**, 943.

[34] Straatsma, T. P. and McCammon, J. A. (1990). "Free energy thermodynamic intergrations in molecular dynamics simulations using a noniterative method to include electronic polarization", *Chem. Phys. Lett.*, **167**, 252.

[35] Straatsma, T. P. and McCammon, J. A. (1991). "Free energy evaluation from molecular dynamics simulations using force fields including electronic polarization", *Chem. Phys. Lett.*, **177**, 433.

[36] Silberstein, L. (1917). *Philos. Mag.*, **33**, 92.

[37] Silberstein, L. (1917). *Philos. Mag.*, **33**, 215.

[38] Silberstein, L. (1917). *Philos. Mag.*, **33**, 521.

[39] Applequist, J., Carl, J. R. and Fung, K. K. (1972). "An atom dipole interaction model for molecular polarizability. Application to polyatomic molecules and determination of atom polarizabilities", *J. Am. Chem. Soc.*, **94**, 2952.

[40] Applequist, J. (1993). "Atom charge transfer in molecular polarizabilities. Application of the Olson-Sundberg model to aliphatic and aromatic hydrocarbons", *J. Phys. Chem.*, **97**, 6016.

[41] Voisin, C., Cartier, A. and Rivail, J.-L. (1992). "Computation of accurate electronic molecular polarizabilities", *J. Phys. Chem.*, **96**, 7966.

[42] Yeh, Y.-H. and Fink, W. H. (1986). "A strategy for the regional characterization of potential energy surfaces", *J. Comput. Chem.*, **7**, 539.

[43] Murtagh, B. A. and Sargent, R. W. H. (1970). "Computational experience with quadratically convergent minimisation methods", *Comput. J.*, **13**, 185.

[44] McIver, J. W. Jr. and Komornicki, A. (1971). "Rapid geometry optimization for semiempirical molecular orbital methods", *Chem. Phys. Lett.*, **10**, 303.

[45] Celda, B., Personal communication.

[46] Torrens, F., Ortí, E. and Sánchez-Marín, J. (1991). "Vectorized TOPO program for the theoretical simulation of molecular shape", *J. Chim. Phys. Phys.-Chim. Biol.*, **88**, 2435.

[47] Torrens, F., Ortí, E. and Sánchez-Marín, J., "Calculation of geometrical descriptors and topological indices of molecules. A vectorized program, In: *High Performance Computing II*, Durand, M. and El Dabaghy, F. Eds., Elsevier, Amsterdam, 1991, pp. 549–560.

[48] Torrens, F., Sánchez-Marín, J. and Nebot-Gil, I. (1998). "Characterizing cavities in model inclusion molecules: A comparative study", *J. Mol. Graphics Mod.*, **16**, 57.

[49] Clark, T., *A Handbook of Computational Chemistry*, Wiley, New York, 1985, Ch. 2, p. 18.

[50] Rubio, M., Torrens, F. and Sánchez-Marín, J. (1993). "Are most of the stationary points in a molecular association minima? An application of Fraga's potential to benzene-benzene", *J. Comput. Chem.*, **14**, 647.

[51] Torrens, F., Sánchez-Marín, J., Ortí, E. and Nebot-Gil, I. (1987). "Incorporation of a dispersion energy term to Fraga's atom-atom pair intermolecular potential. Application to benzene, *s*-tetrazine, and their mixed dimers", *J. Chem. Soc., Perkin Trans. 2*, 943.

[52] Torrens, F., Sánchez de Merás, A. M. and Sánchez-Marín, J. (1988). "The use of *ab initio* net charges to improve Fraga's atom-atom pair potential for molecular association", *J. Mol. Struct. (Theochem)*, **166**, 135.

[53] Torrens, F., Sánchez-Marín, J. and Tomás, F. (1990). "Intermolecular calculations on azine dimer", *J. Chem. Res. (S)*, p. 176.

[54] Torrens, F., Ortí, E. and Sánchez-Marín, J., "Structural and charge effects on the phthalocyanine dimer", In: *Physical Chemistry Modelling of Molecular Structures and*

Properties XLIV, Rivail, J. L. Ed., *Stud. Phys. Theor. Chem.*, No. 71, Elsevier, Amsterdam, 1990, pp. 221–228.

[55] Torrens, F., Ortí, E. and Sánchez-Marín, J. (1991). "Pair potential calculation of molecular associations. A vectorized version", *Comput. Phys. Commun.*, **66**, 341.

[56] Torrens, F., Ortí, E. and Sánchez-Marín, J. (1991). "Improved AMYR program: An algorithm for the theoretical simulation of molecular associations, including geometrical and topological characterization of the dimers", *J. Mol. Graphics.*, **9**, 254.

[57] Torrens, F., Ortí, E. and Sánchez-Marín, J., "Electrically conductive phthalocyanine assemblies. Structural and non-integer oxidation number considerations", In: *Lower-Dimensional Systems and Molecular Electronics*, Metzger, R. M., Day, P. and Papavassiliou, G. C. Eds., NATO-ASI Ser. B No. 248, Plenum Press, New York, 1991, pp. 461–466.

[58] Torrens, F., Montañana, R. and Sánchez-Marín, J., "Vectorizing pair-potential AMYR program for the study of molecular associations", In: *High Performance Computing I*, Delhaye, J. L. and Gelenbe, E. Eds., Elsevier Amsterdam, 1989, pp. 299–308.

[59] Rodríguez, J., Sánchez-Marín, J., Torrens, F. and Ruette, F. (1992). "Molecular aggregation of polycyclic aromatic hydrocarbons. A theoretical modelling of coronene aggregation", *J. Mol. Struct.* (*Theochem*), **254**, 429.

[60] Torrens, F., Rubio, M. and Sánchez-Marín, J. (1998). "AMYR 2: A new version of a computer program for pair potential calculation of molecular associations", *Comput. Phys. Commun.*, **115**, 87.