



J. Serb. Chem. Soc. 78 (11) 1789–1795 (2013)
JSCS–4533

SHORT COMMUNICATION

An *ab-initio* study of pyrrole and imidazole arylamides

ARA M. ABRAMYAN, ZHIWEI LIU and VOJISLAVA POPHRISTIC*

*Department of Chemistry & Biochemistry, University of the Sciences, Philadelphia,
PA 19104, USA*

(Received 29 September, revised 9 October 2013)

Abstract: Arylamide foldamers have been shown to have a number of biological and medicinal applications. For example, a class of pyrrole–imidazole polyamide foldamers is capable of binding specific DNA sequences and preventing development of various gene disorders, most importantly cancer. Molecular dynamics (MD) simulations can provide crucial details in understanding the atomic level events related to foldamer/DNA binding. An important first step in the accurate simulation of these foldamer/DNA systems is the reparametrization of force field parameters for torsion around the aryl–amide bonds. Herein, the density functional theory (DFT) potential energy profiles and the derived force field parameters for four types of aryl–amide bonds for pyrrole and imidazole building blocks, which have been extensively employed in foldamer design for DNA-binding polyamides, will be highlighted. These results contribute to the development of computational tools for an appropriate molecular modeling of pyrrole–imidazole polyamide/DNA binding, and provide an insight into the chemical factors that influence the flexibility of pyrrole–imidazole polyamides and their binding to DNA.

Keywords: foldamer, torsional energy profiles, force field reparametrization, DNA-binding polyamide.

INTRODUCTION

Foldamers are synthetic oligomers that adopt stable secondary structures in solution.¹ Their functionality strongly depends on their conformation in solution. A number of important medicinal applications for foldamers have been demonstrated.^{2,3} A particular class of aromatic polyamide foldamers, consisting of pyrrole and imidazole monomers, has been shown to bind DNA sequences.^{4,5} Due to their ability to compete with DNA binding proteins, these polyamides have been used to block development of various diseases, including cancer.⁶ These DNA binders are composed of double-stranded pyrrole and imidazole carboxamides

* Corresponding author. E-mail: v.pophri@uscience.edu
doi: 10.2298/JSC130929104A

(Fig. 1). Dervan and co-workers⁴ revealed that *N*-methylimidazole (Im) paired with *N*-methylpyrrole (Py) binds G–C base pairs, and that the Py–Im pair binds C–G base pairs, whereas the Py–Py pair binds both A–T and T–A pairs. Thus, pyrrole–imidazole oligomers can be designed for DNA sequence-specific targeting.

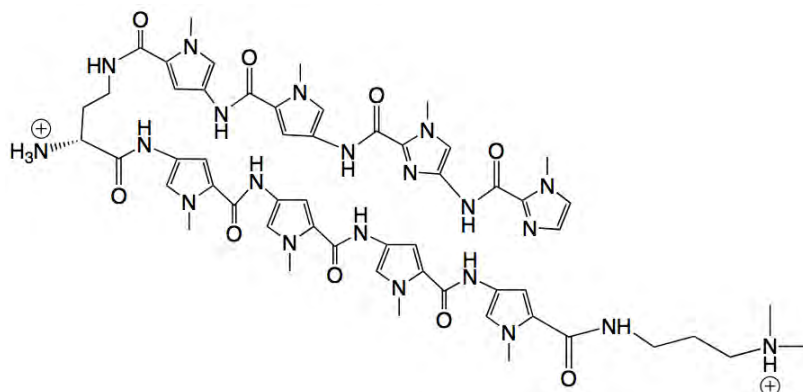


Fig. 1. Structure of a hairpin pyrrole–imidazole polyamide synthesized by Dervan *et al.*⁶ as an antitumor candidate.

Despite the importance of these polyamides, little is known about their dynamics in solution, in either free or in DNA-bound states. Molecular dynamics (MD) simulations can be used to shed light on the structural and dynamical characteristics of these DNA binding foldamers provided that accurate force fields are used.⁷ As the aromatic rings and peptide bonds are planar and rigid, the conformations of these polyamides are determined by only two types of bond rotations: those around the $C_{\text{aromatic}}-C_{\text{peptide}}$ (C_a-C_p) and $C_{\text{aromatic}}-N_{\text{peptide}}$ (C_a-N_p) bonds (see Fig. 2). In a previous work,⁷ it was shown that the general AMBER force field (GAFF)⁸ overestimates the torsional potentials for rotations around the aryl–amide bond (the GAFF torsional barrier for the C_a-C_p rotation corresponds to 29 kcal mol⁻¹,⁸ whereas quantum mechanically (QM) derived barriers are ≈ 8 kcal mol⁻¹).⁷ In addition, further studies^{9–12} showed that specific structural features significantly influence the rotational degree of freedom of an aryl–amide bond. Examples include an intramolecular hydrogen bond (H-bond) between the amide NH and an *ortho* aromatic substituent or an *ortho* endocyclic oxygen or nitrogen atoms. Therefore, to represent accurately the structure and dynamics of these polyamides by MD simulations, first the effects of various structural characteristics and energetics of aryl–amide torsions were studied using high-level QM methods. Reparametrization of the building block specific torsional parameters for GAFF was then realized using these QM potential energy profiles. In this paper, the determined QM potential energy profiles of four different aryl–amide bond types and the derived force field parameters are reported.

COMPUTATIONAL METHODS

The torsional profiles of four small molecules: *N*-(1-methyl-1*H*-imidazol-4-yl)acetamide (Im-N); *N*,1-dimethyl-1*H*-imidazole-2-carboxamide (Im-C); *N*-(1-methyl-1*H*-pyrrol-3-yl)acetamide (Py-N) and *N*,1-dimethyl-1*H*-pyrrol-2-carboxamide (Py-C) are shown in Fig. 2. These molecules were used for the reparameterization of the four different types of dihedral angles in the polyamides shown in Fig. 1. First, the molecules were optimized at several different levels of theory (HF, MP2 and B3LYP), and with basis sets ranging from B3LYP/6-31G(d,p) to B3LYP/6-311G(3df,2p). A convergence study revealed that B3LYP/6-311G(d,p) agrees well with the higher levels of calculation, as was also shown in previous studies.^{7,10} Therefore, the potential energy profiles were obtained by scanning the C_a-C_p and C_a-N_p torsional surface in 20° increments at the B3LYP/6-311G(d,p) level of theory. At each scan point, all internal coordinates except the dihedral angle in question were optimized. The calculations were realized using Gaussian 03.¹³ Then these torsional profiles were employed to reparametrize the GAFF parameters for the torsions around the C_a-C_p and C_a-N_p bonds. The new torsional parameters were obtained by the procedure described in a previous study,¹⁰ *i.e.*, by subtracting the non-bonded energies (calculated with restrained electrostatic potential, RESP,¹⁴ charges) from the *ab initio* torsional potential profile and then least square fitting of the GAFF torsional function to the “subtracted” *ab initio* torsional potential profile. Charges reported in this study were obtained by the RESP¹⁴ procedure.

RESULTS AND DISCUSSION

Torsional energy profiles with respect to the four types of dihedral angles are shown in Fig. 2. For compound Im-C, the global minimum is found with a dihedral angle N_{a1}-C_a-C_p-N_p at 180°. This is clearly due to the favorable H-bonding between the amide proton and the endocyclic unsaturated nitrogen (with a N_{a2}-H distance of 2.28 Å). Upon rotation of the C_a-C_p bond, an energy maximum is reached at around ±90°. The related increase in energy (12.97 kcal mol⁻¹) can be attributed to the loss of both the H-bond and the π-electron delocalization between the aromatic ring and the peptide group, which are co-planar in the global minimum. The highest energy point corresponds to a N_{a1}-C_a-C_p-N_p dihedral angle value of 0°, which is at 15.42 kcal mol⁻¹ with respect to its global minimum. This high energy barrier arises from several sources: a) destabilization due to the loss of the intramolecular H-bond discussed above; b) electrostatic repulsion between the carbonyl oxygen (O_{peptide}) and the endocyclic nitrogen (N_{a2}) in the highest energy conformer (the O_{peptide} to N_{a2} distance is 2.80 Å) and c) the steric repulsion between the amide proton and the N_{a1}-methyl group (the distance between amide proton and the carbon atom in the N_{a1}-methyl group (C_{methyl}) is 2.49 Å) in the highest energy conformer. The distance between the O_{peptide} and the carbon atom of the N_{a1}-methyl group in the equilibrium conformer is 2.87 Å, which is longer than the distance between the amide proton and the carbon atom in the N-methyl group in the highest energy conformer. A relatively flat energy region is observed at dihedral angles between 90° and 30° (or between -90 and -30°) which is due to a counter balancing effect of the

stabilizing (π -electron delocalization) and destabilizing factors (electrostatic and steric repulsions described above).

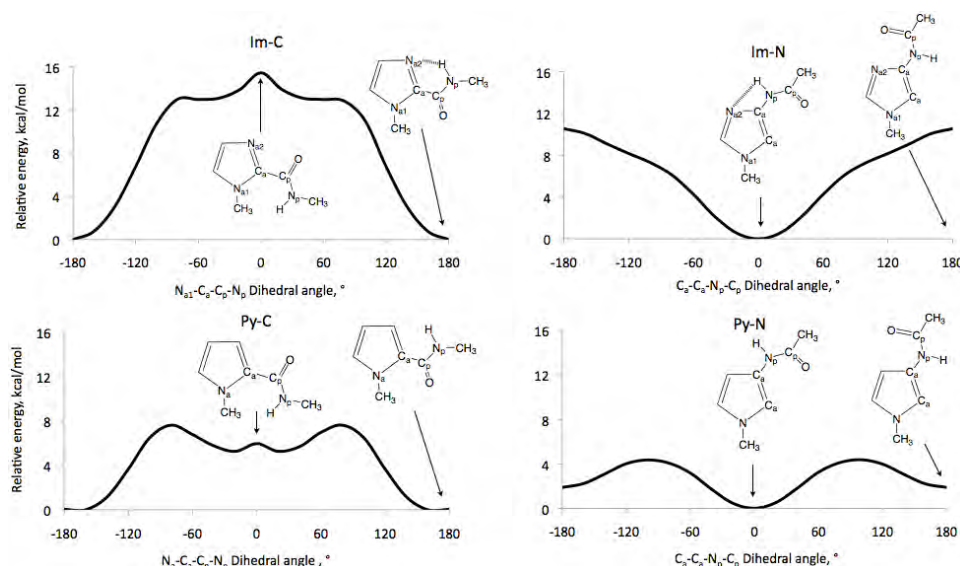


Fig. 2. Torsional energy profiles of the pyrrole and imidazole connected to the peptide bond through C and N atoms (Py-C, Py-N, Im-C, Im-N, respectively). Structures shown correspond to 0 and 180° torsional angle conformations. Dashed lines represent H-bonds.

1 kcal mol⁻¹ = 4.184 kJ mol⁻¹.

The global minimum of compound Im-N has the dihedral angle $C_a-C_a-N_p-C_p$ at 0°. Similarly to Im-C, this is due to stabilization by an H-bond between the amide proton and the endocyclic unsaturated nitrogen (with an $N_{a2}-H$ distance of 2.48 Å). Rotation around C_a-N_p leads to the energy maximum at 180°. The barrier of 10.55 kcal mol⁻¹ with respect to the minimum arises from: a) the loss of the H-bond and b) the electrostatic repulsion between the peptide oxygen and the endocyclic nitrogen (the distance between them is 2.93 Å) in the 180° conformer.

Compound Py-C has its minimum when the dihedral angle $N_a-C_a-C_p-N_p$ is at 180°. At this conformation, the arrangement of its functional groups is the most favorable due to: a) stabilization by electron delocalization between the aromatic ring and the peptide bond and b) the fact that the peptide oxygen and the N_a -methyl group (the distance between them is 2.85 Å) are not close enough for steric repulsion to have significant impact on the energy. Similarly to Im-C, rotation of the C_a-C_p bond leads to energy maximum at $\pm 90^\circ$ due to the loss of π -electron delocalization. The energy barrier of 7.63 kcal mol⁻¹ for this maximum is lower than that of the Im-C compound. This is because there is no H-bond in the global minimum of Py-C. The potential energy decreases slightly

due to the increase of π -electron delocalization when the dihedral angle $N_a-C_a-C_p-N_p$ changes from 90° to 20° (or from -90° to -20°) to reach a second minimum at $5.27 \text{ kcal mol}^{-1}$. Another energy maximum with the $N_a-C_a-C_p-N_p$ dihedral angle at 0° is observed, which is attributed to the steric repulsion between the amide proton and the N_a -methyl group ($H-C_{\text{methyl}}$ distance of 2.38 \AA).

Finally, compound Py-N has the lowest energy barrier ($4.38 \text{ kcal mol}^{-1}$) of the four considered compounds because it has no H-bond acceptors on the aromatic ring or other substituents that influence rotation around the $C_a-C_a-N_p-C_p$ dihedral angle. The maximum is at $\pm 90^\circ$ because this conformer does not have π -electron delocalization between the aromatic ring and the peptide group.

To improve the molecular mechanics description of the compounds, the *ab initio* torsional energy profiles were used to reparametrize the torsional parameters around the C_a-C_p and C_a-N_p bonds in GAFF, as described in the computational methods section. The atomic partial charges and the torsional parameters for the dihedral angles around the C_a-C_p and C_a-N_p bonds are summarized in Table I. These parameters can be used for accurate simulations of the pyrrole-imidazole polyamides and other structures containing the four torsional angles reparameterized herein.

TABLE I. Selected atomic charges and the fitted torsional parameters (bottom row) for the dihedral angles around the C_a-C_p and C_a-N_p bonds

Im-C		Im-N		Py-C		Py-N	
N_{a2}	-0.5843	N_{a2}	-0.4984	C_a	-0.1346	C_a	0.0574
C_a	0.1267	C_a	0.0529	C_p	0.5968	N_p	-0.3493
C_p	0.5496	N_p	-0.4625	O_p	-0.5458	H_p	0.2879
O_p	-0.5646	H_p	0.2780	N_p	-0.4267	C_p	0.5791
N_p	-0.3734	C_p	0.5848	H_p	0.2994	O_p	-0.5676
H_p	0.2855	O_p	-0.5627	V_n	7.63	V_n	3.0
V_n^a	10.0	V_n	3.0				

^aIn kcal mol^{-1} ($1 \text{ kcal mol}^{-1} = 4.184 \text{ kJ mol}^{-1}$). $E_{\text{torsion}} = \frac{V_n}{2} [1 + \cos(n\phi - f)]$, $n = 2$ and $f = 180^\circ$

CONCLUSIONS

Generally, higher torsional barriers were observed for aryl-amide torsions in imidazoles than in pyrroles. For example, the barrier of Im-C is 2 times higher than that of Py-C ($\approx 16 \text{ kcal mol}^{-1}$ vs. $\approx 8 \text{ kcal mol}^{-1}$), while the barrier of Im-N is ≈ 3 times that of Py-N (≈ 11 vs. $\approx 4 \text{ kcal mol}^{-1}$). This is due to the presence of the unsaturated endocyclic nitrogen (N_{a2}) in the imidazoles, which results in: 1) an H-bond between the amide proton and N_{a2} in the global minimum and 2) the electrostatic repulsion between the O_{peptide} and N_{a2} in the energy maximum. The stabilization of the minimum by the H-bond and destabilization of the maximum by the repulsion together result in a much higher barrier for the imidazole compounds.

Higher energy barriers were also observed in compounds Im-C and Py-C relative to those in Im-N and Py-N. There are two structural differences between the Im-C (or Py-C) and Im-N (or Py-N) molecules that need to be considered to account for the difference in the energy barriers. First, they differ in the type of connection (C_a-C_p vs. C_a-N_p). The C_a-C_p connection type in Im-C results in a more favorable H-bond between the amide H and the endocyclic N (the $H\cdots N_{a2}$ distance is 2.28 Å and the $N_p-H\cdots N_{a2}$ angle is 107.5°) than that in the minimum of Im-N (the $H\cdots N_{a2}$ distance is 2.48 Å and the $N_p-H\cdots N_{a2}$ angle is 72.2°). Therefore, it requires more energy to break the more favorable H-bond in Im-C than in Im-N. Secondly, the N_{a1} -methyl is *ortho* to the aryl-amide bond in the Im-C and Py-C molecules, but *meta* in the Im-N and Py-N molecules. The latter difference results in a lower (or no) steric repulsion between the amide proton and the N-methyl group in the high-energy conformers of Im-N and Py-N than in the high-energy conformers of Im-C and Py-C.

The force field parameters provided here are of immediate importance for proper modeling and binding studies of systems developed by Dervan and others.^{5,6} As has been shown in the past (with arylamides⁷), binding modes that result from MD simulations can be strikingly different, depending on the suitability of the torsional parameters. Consequently, a realistic description of the flexibility of an oligoamide backbone requires carefully derived torsional parameters.

This series of monomers allows for fine-tuning of the flexibility of oligomer backbones, as Im-C (15.42 kcal mol⁻¹), Im-N (10.55 kcal mol⁻¹), Py-C (7.63 kcal mol⁻¹) and Py-N (4.38 kcal mol⁻¹) have distinct torsional barriers for the bonds involved in a polyamide scaffold. Thus, the choice of the monomer sequence could result in various levels of backbone flexibility. This design strategy, however, is of limited benefit when sequence specific binding is sought. In such cases, quantitative information about the relative torsional barriers provides an opportunity to use alternative design features (such as “closed” double stranded structures⁵) in a better-informed manner, to affect backbone flexibility.

From a broader perspective, this and previous studies suggest caution against simply relying on chemical intuition in molecular design, since often the exact outcome of the interplay of several forces driving conformational stability cannot be predicted. Therefore, highly accurate QM calculations of the energetics of molecular building blocks and their analysis can be an important tool in molecular design.

Acknowledgments. This work was supported by the NSF MSN-1049771 and MRI CHE-1229564 grants. One of the authors (V. P.) would like to thank her former professors, members of the Faculty of Physical Chemistry, University of Belgrade, for the excellent education in physical chemistry and strong preparation for a research career. In particular, she is

grateful to her undergraduate mentor, Prof. Dr. Miljenko Perić, for introducing her to the field of computational chemistry and for his support in her development as a scientist.

ИЗВОД

AB INITIO СТУДИЈА ПИРОЛ- И ИМИДАЗОЛ-АРИЛАМИДА

ARA M. ABRAMYAN, ZHIWEI LIU и ВОЈИСЛАВА ПОПХРИСТИЋ

Department of Chemistry & Biochemistry, University of the Sciences, Philadelphia, PA 19104, USA

Фолдамери базирани на ариламидима се користе у биолошке и медицинске сврхе. На пример, једна класа пирол–имидазол полиамид–фолдамера се везује за одређене секвенције ДНК, и на тај начин спречава развијање разних генетских проблема и болести, укључујући и рак. Симулације молекулске динамике (МД) пружају могућност за увид у атомске детаље процеса који се одвијају при везивању фолдамера за ДНК. Од нарочите важности за поуздане МД симулације фолдамер/ДНК система је употреба исправних торзионих параметара за арил–амид везе. У овом раду, ми описујемо DFT (од енг. *density functional theory*; теорија функционала густине) торзионе потенцијалне површине и на основу њих изведене торзионе параметре, за четири типа арил–амид веза, које се појављују у пирол–имидазол мономерима коришћеним у дизајнирању фолдамер/ДНК комплекса. Овај рад описује једно додатно средство које омогућава поуздане МД симулације пирол/имидазол фолдамер система. Такође, резултати представљени у овом раду дефинишу на који начин одређени структурни фактори утичу на флексибилност полиамида и њихово везивање за ДНК.

(Примљено 29. септембра, ревидирано 9. октобра 2013)

REFERENCES

1. S. H. Gellman, *Acc. Chem. Res.* **31** (1998) 173
2. C. M. Goodman, S. Choi, S. Shandler, W. F. DeGrado, *Nat. Chem. Biol.* **3** (2007) 252
3. G. N. Tew, R. W. Scott, M. L. Klein, W. F. DeGrado, *Acc. Chem. Res.* **43** (2010) 30
4. S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, *Nature* **391** (1998) 468
5. D. M. Chenoweth, P. B. Dervan, *J. Am. Chem. Soc.* **132** (2010) 14521
6. J. A. Raskatov, J. L. Meier, J. W. Puckett, F. Yang, P. Ramakrishnan, P. B. Dervan, *Proc. Natl. Acad. Sci. USA.* **109** (2012) 1023
7. V. Pophristic; S. Vemparala, I. Ivanov, Z. Liu, M. L. Klein, W. F. DeGrado, *J. Phys. Chem., B* **110** (2006) 3517
8. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **25** (2004) 1157
9. G. F. Galan, C. N. Tang, S. Chakrabarty, Z. Liu, G. Moyna, V. Pophristic, *Phys. Chem. Chem. Phys.* **15** (2013) 11883
10. Z. Liu, A. Teslja, V. Pophristic, *J. Comput. Chem.* **32** (2011) 1846
11. G. F. Galan, J. Brown, J. L. Wildin, Z. Liu, D. Liu, G. Moyna, V. Pophristic, *J. Phys. Chem., B* **113** (2009) 12809
12. Z. Liu, R. C. Remsing, D. Liu, G. Moyna, V. Pophristic, *J. Phys. Chem., B* **113** (2009) 7041
13. Gaussian 03, Revision C.02, Gaussian, Inc., Wallingford, CT, 2004
14. W. D. Cornell, P. Cieplak, C. I. Bayly, P. A. Kollmann, *J. Am. Chem. Soc.* **115** (1993) 9620.