Tetrahedron Letters 49 (2008) 6674–6678

Contents lists available at ScienceDirect

Tetrahedron Letters

journal homepage: www.elsevier.com/locate/tetlet

Structure and dynamics of pyrimidine-based macrocycles in solution

Artem V. Kozlov, Vyacheslav E. Semenov, Anatoliy S. Mikhailov, Vladimir S. Reznik, Shamil K. Latypov *

Institute of Organic and Physical Chemistry, Arbuzov Street 8, Kazan 420088, Russian Federation

ARTICLE INFO

Article history:
Received 1 July 2008
Revised 27 August 2008
Accepted 10 September 2008
Available online 13 September 2008

Keywords:
Pyrimidine
Uracil
Macrocycle
NMR
Hydrogen bond
Protonation

ABSTRACT

The conformational structure of macrocycles obtained from two thiopyrimidine and uracil nucleic acids linked by polymethylene spacers is determined by the length of the spacers, intramolecular NH bonding, pH and solvent. In $CDCl_3$, NH–O═C hydrogen bonding can impact the overall stabilization of the folded conformation, however spatial preorganization to such hydrogen bonding is a prerequisite. Protonation leads to disruption of intramolecular hydrogen bonds, destabilization of the folded conformation and to strong counterion assisted self-aggregation of macrocyles which can be destroyed in polar solvents.

One of the methods in rational design of new biologically active compounds is the combination of a priori active (pharmacophore) units and their spatial preorganization by different spacers.[1–3] Distinct architecture can be achieved using rigid linkers. The use of labile spacers results in macrocycles and acyclic structures with more conformational flexibility. In addition, the use of units prone to weak non-covalent interactions opens the way to fine tuning of the 3D and supramolecular structure of such systems. In this respect, macrocyclic derivatives of nucleic acids bonded by different spacers are very promising.[4,5]

However, for rational design of such compounds it is necessary to know their 3D structure, the main factors that determine their geometry and energy. Particularly significant is to predict correctly the influence of weak intra- and intermolecular interactions and solvent effects on the macrocyclic structure and on the structure of their complexes as well. It is important to have such data in solution as most metabolic processes and reactions occur in this state.

These non-covalent interactions are responsible for recognition and complexation of ligands (guests) with receptors/targets (DNA/RNA). The potential to recognize and bind is also determined by the conformational structure of the host and guest, which are also partly controlled by weak intramolecular dispersion interactions. Therefore, investigation of such interactions has been of significant interest for a long time and there have been a number of publications in recent years.[6–8] However, there are many computational studies and very limited experimental data, particularly in solution. Thus, in spite of a number of studies devoted to this problem, there is still a lot to be understood.

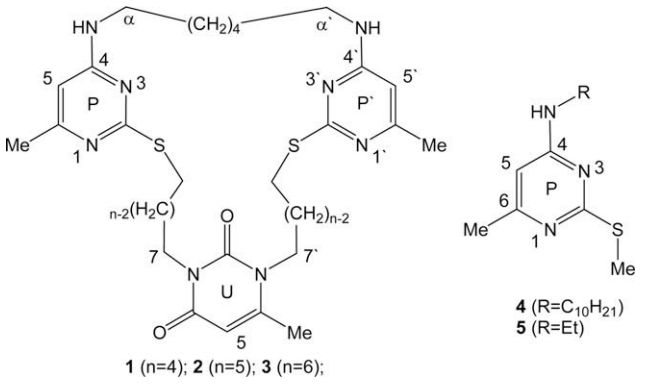
There are intrinsic obstacles to the investigation of such interactions because it is very difficult to 'measure' small indications of non-covalent interactions, and to separate the different terms of these interactions from each other. In addition, a problem arises due to the variety of tautomeric and charged forms which contribute to the net non-covalent interactions, which are solvent and pH dependent.[9–11] This complicates the overall picture and can lead to an enormous number of energy minima on the potential energy surface of such systems.

Nucleic acid-based macrocycles are of interest for their biorelevancy, selective complexation ability towards different functional groups, and there have been a number of biochemical investigations with such systems.[12–14] There are various examples of the design of macrocyclic and acyclic derivatives of thymine, uracil and pyrimidine fragments preorganized by aliphatic and other spacers.[15–17]

These macrocycles are also of interest from another point of view—as models to investigate intra- and intermolecular non-covalent interactions. A new type of pyrimidinophane containing uracil and thiopyrimidine moieties bridged by different aliphatic chains was elaborated (Scheme 1).[18,19] According to IR data in solution, these macrocycles have both free and intramolecular hydrogen bonded NH protons.[18] In general, this could lead to some folded conformation(s). The non equivalency of the H5 protons of the thiopyrimidine moieties in ^{1}H NMR spectra indirectly supports this hypothesis. That is, these protons are far from the asymmetric uracil moiety, therefore their non-equivalency might be ascribed to a folded conformation in solution that transfers the information

* Corresponding author. Tel.: +7 843 272 7484; fax: +7 843 273 2253.
E-mail address: lsk@iopc.knc.ru (Sh. K. Latypov).

0040-4039/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tetlet.2008.09.047



Scheme 1.

about the asymmetry to the thiopyrimidine rings. However, no support for the overall structure of the macrocycles has been provided yet.

Taking into account the above facts and considerations, we have undertaken to find out whether the folded structure of these macrocycles exists in solution. In this Letter, we present preliminary results on the conformational analysis of three macrocycles **1–3** by dynamic NMR and quantum chemical methods. In addition, model compounds that mimic different parts of the macrocycles were also prepared and investigated.

In the room temperature ^{1}H NMR spectra of the title macrocycles, there were only minimal differences. In all the compounds, there was only a small non-equivalence of the H5(P) protons and broadening of the CH_2 protons adjacent to NH (Fig. 1a for **1**). Attempts to obtain information on the folded structure at room temperature by NOE measurements were unsuccessful. In the 1D DPFGSE and 2D NOESY spectra (in $CDCl_3$) there were no 'non-trivial' NOEs that could be ascribed to such geometry.

Decreasing the temperature has a significant impact on the ^{1}H NMR spectra. At ca. 273–263 K there was coalescence of the spectra (Fig. 1b) and then at ca. 233 K slow exchange was observed (on the NMR time scale) (Fig. 1c). At a lower temperature, (ca. 223–213 K) there was an indication of the second dynamic process that mainly affected the NH and H5 signals (Fig. 1d). While the first process (broadening at ca. 273 K) is similar in all these compounds, the second process is less pronounced in **2** and more marked in **1** (Fig. 1e and d).

In order to clarify the nature of these processes, dynamic NMR experiments on the simpler thiopyrimidine-containing model **4** were carried out. ^{1}H (Fig. 2), ^{13}C, ^{1}H–^{13}C and ^{1}H–^{15}N 2D HSQC/HMBC and NOESY experiments at different temperatures allowed us to conclude that the first process is rotation around the C4–NH bond which produces two conformations: *Z* and *E* (Fig. 3), the latter being dominant (0.26 kcal/mol) in $CDCl_3$. This conclusion is supported additionally by the fact that the barrier of the process derived from line shape analysis for **4** ($\Delta H^{\#}$ = 15.8 kcal/mol) is in good agreement with the theoretical barrier (HF/6-31G) calculated for rotation around the C4–NH bond in the simpler model **5** ($\Delta E^{\#}$ = 14.4 kcal/mol).[20]

It is important that in **4**, indications of the second process were also seen at T = 223–213 K (Fig. 2d), which was very similar to those observed in compounds **1–3**. To explain this broadening of the signals, several likely processes might be invoked. With **4,**

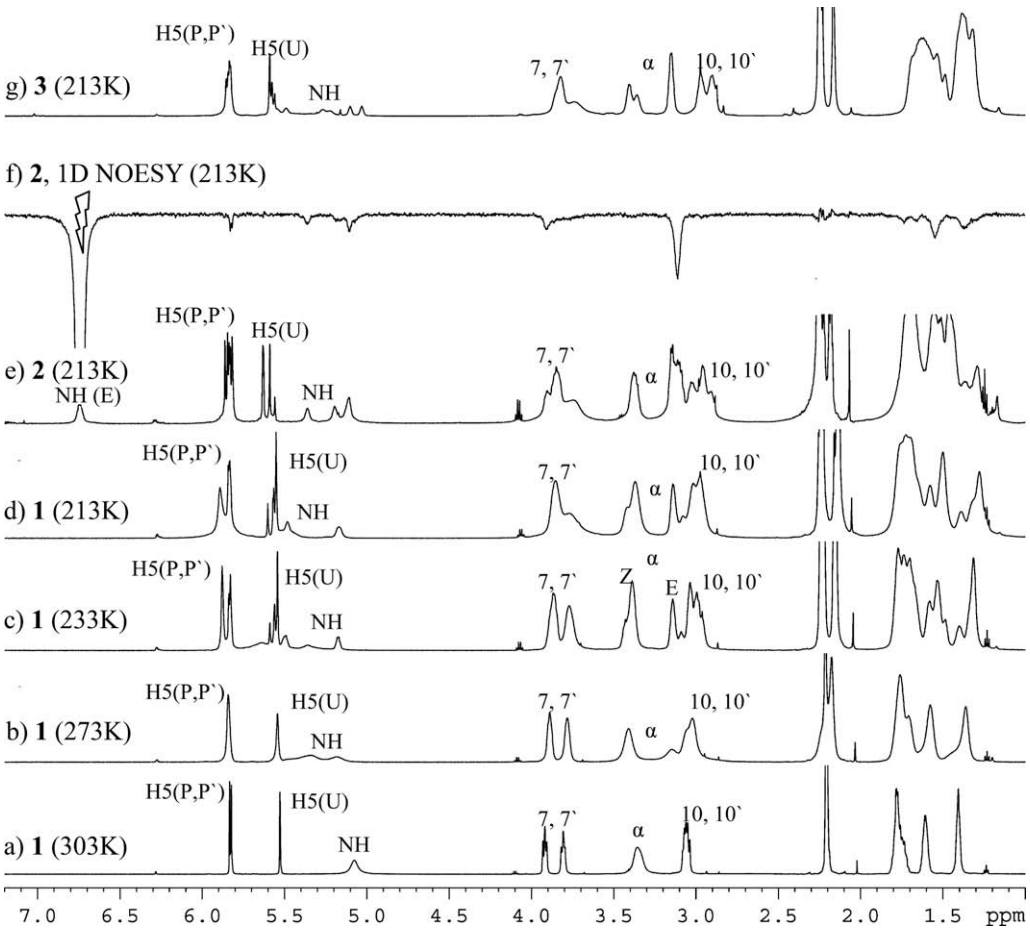


Figure 1. ^{1}H NMR spectra in $CDCl_3$ at different temperatures of **1** (a–d), **2** (e); 1D NOESY of **2** (f); **3** (g).

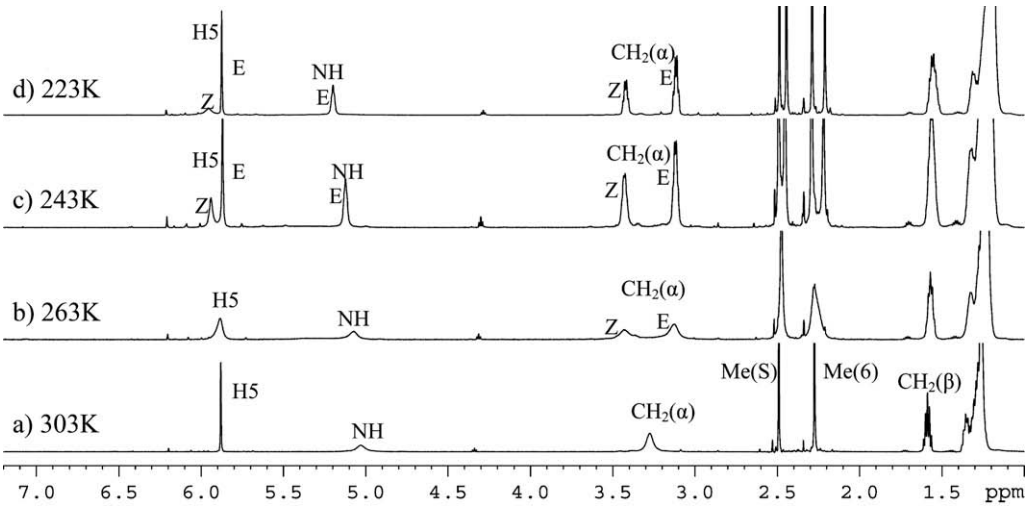


Figure 2. ^{1}H NMR spectra of **4** in $CDCl_3$ at different temperatures (*E* and *Z* are denoted according to Fig. 3).

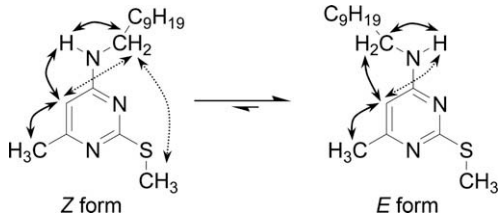


Figure 3. Main conformations due to rotation around the C4–NH bond in **4**, NOEs at $T = 213$ K in $CDCl_3$ (weak NOEs are shown by dashed lines).

the barriers of conformational exchange (rotation around the C–S or C–C bonds) are very low (<7 kcal/mol), and should not slow down at these temperatures on the NMR time scale. Any significant contribution from tautomeric equilibria or exchange was also excluded.[19] As a hypothetical reason for such NMR behaviour there might also be an impact due to the protonated form which even in small quantities can influence the ^{1}H NMR spectra.[16] Our hypothesis is that the compounds were partially protonated already in the course of reaction, that is, there was a small quantity (ca. 5%) of protonated form present. So, in the low temperature NMR spectrum there might be fast exchange between the neutral (dominant) and protonated (minor) forms. To check this hypothesis, we tried to increase or to neutralize these effects by adding acid or base, respectively. ^{1}H NMR experiments on **4** showed that these indications are enlarged (Fig. 4b–e) with an increase of the population of the protonated form when we titrated with trifluoroacetic acid (TFA).[21] At an equimolar concentration of TFA, the title compound is fully protonated.[19] In contrast, the addition of base (e.g., Et_3N, TEA) neutralizes the charge and diminishes the contribution of

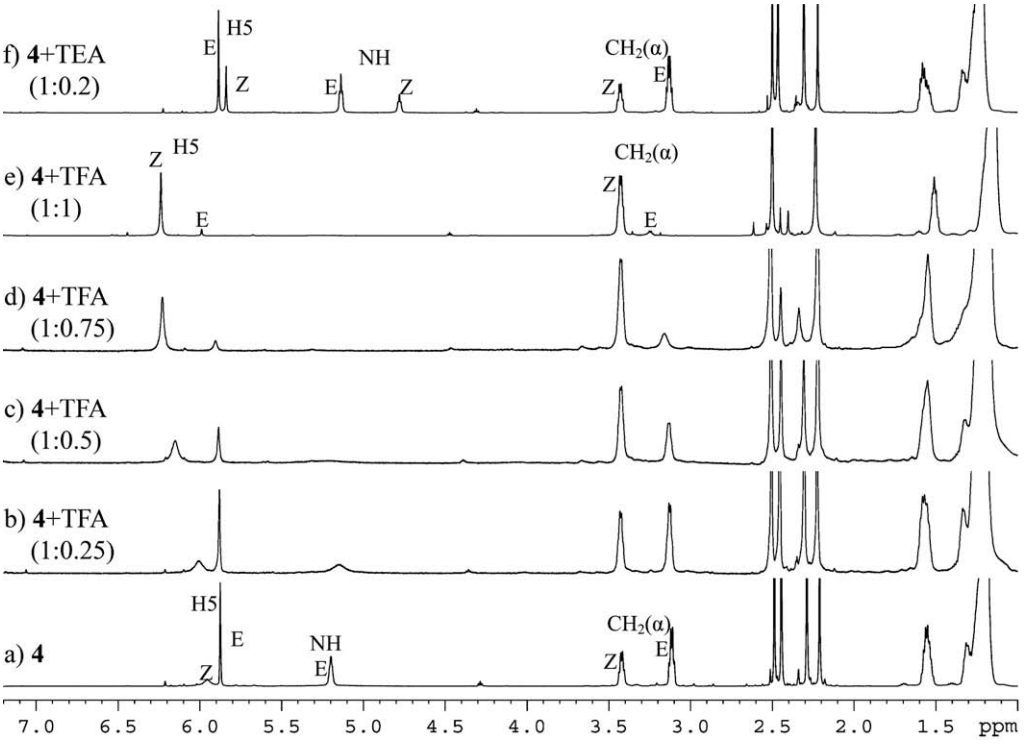


Figure 4. Dependence of the ^{1}H NMR spectra of **4** in $CDCl_3$ at low temperature on the concentration of TFA (b–e)/TEA (f) (*E* and *Z* are denoted according to Fig. 3). In 4 + TFA(1:1) (e) the NH(Z) proton is at 14 ppm (fragment not shown).

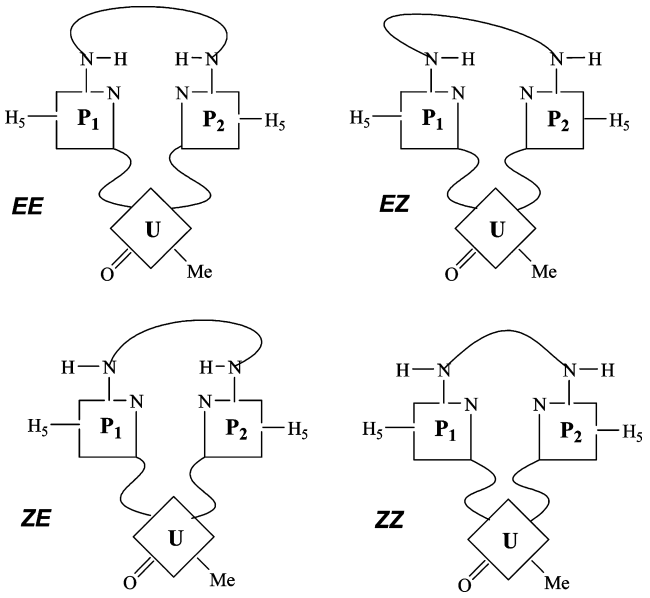


Figure 5. Schematic representation of the conformation around the C4–NH bond in the macrocycles.

the protonated form and as a result, the lines become sharp and shift to high field as they should for the neutral form (Fig. 4f). Thus, we conclude that it is the presence of the protonated form and its fast exchange (on the NMR time scale) with the neutral form which leads to such line shape evolution in the ^{1}H NMR spectra of **4** at low temperatures. In addition, it was found that protonation has a great impact on the conformational structure and associative properties of the molecules. It was established that protonation occurs at N1, and slows down the rotation around the C4–NH bond ($\Delta H^{\#}$ = 25.8 kcal/mol), leads to bias conformational equilibrium in favour of the *Z* form (*Z*/*E*—93/7) and leads to aggregation (*V* increases by 16 times).[19] According to DOSY data and GIAO chemical shift calculations, the acid counterion participates as a bridge between protonated monomers. Important results derived for this model are spectral 'finger prints' for the *Z* and *E* conformations in the neutral and protonated forms of **4**, which were used for analysis of more complicated macrocycles **1–3** (vide infra).

Having indicators of different forms for **4,** we derived conclusions on the conformational structure of the macrocycles (at least, of some fragments). First of all, the first collapse at ca. 263 K corresponds to a slowing down of the rotation around the C4–NH bonds, which leads to exchange of the forms due to different orientations of the NH protons in the two thiopyrimidine moieties. In general, this can lead to equilibrium of the four combinations (*EE*, *EZ*, *ZE* and *ZZ*, Fig. 5). As in the case of **4**, for the macrocycles there is a marked impact of the protonated form in $CDCl_3$ solution, particularly for **1**. Deprotonation by TEA diminishes this process and well resolved signals of several of the NH and H5 protons are observed in the spectra at low temperature (Fig. 6b). In contrast, protonation produces extensive broadening and downfield shifts of these lines (Fig. 6c).

In general, there are no structurally specific shielding/deshielding effects in the ^{1}H NMR spectra ($CDCl_3$) that could be correlated with mutual orientations of the pyrimidine moieties. Only in **2** does one NH proton resonate at a remarkably low field (6.7 ppm) while another NH in **2** and in other macrocycles are ca. 5.05–5.35 ppm (Fig. 1e) which is close to the NH chemical shift in **4** where there is no hydrogen bond (Fig. 2). Thus according to ^{1}H NMR spectra, only for compound **2,** and only for one NH bond can moderate hydrogen bonding be identified at 213 K.

Moreover, attempts to obtain 3D structural information through NOE measurements were also of little success for these macrocycles at T = 213 K. The only 'non-trivial' NOE is between the 'hydrogen bonded' NH(*E*) and H7 protons in this conformation (Fig. 1f).

It is significant that in the low temperature ^{1}H spectra of **2** in $CDCl_3$, there are several well resolved lines for H5(U). The dominant signal resonates at lower field than the other and corresponds to a conformation in which the NH(*E*) group is hydrogen bonded. Thus, if one takes into account that there are no such hydrogen bonds and no such H5(U) signals in **1** and **3**, and that there is an NOE between $NH(E)_{HB}$ and H7 it can be supposed that it is the hydrogen bonding of the NH with the C=O group of the uracil moiety which leads to such a low field shift of the $H5(U)_{HB}$ proton.

To check this hypothesis, the conformational space was scanned by the molecular dynamics method in order to search for stable forms. The energies of suitable conformers were minimized by molecular mechanics, and those forms with energies within 7 kcal/mol respective to the corresponding equilibrium form were selected. Next, these structures were optimized by ab initio methods (HF/6-31G). According to the calculations, in the most stable structure (Fig. 7): (1) the NH proton is in *E* orientation with respect to the endocyclic N3 of the thiopyrimidine which is in agreement with the experiment; (2) in this conformation, a hydrogen bond can occur between this NH and O=C of the uracil (N···O and H···O distances are 3.1 and 2.11 Å, respectively) which is also in accordance with the low field shift of this NH proton. Moreover, the experimentally observed low field shift for $H5(U)_{HB}$ is also predicted in this structure according to ^{1}H chemical shift calculations (GIAO B3LYP/6-31G(d)//HF/6-31G).

The intramolecular hydrogen bonding energy in **2** can be estimated from the NH_{HB} chemical shift taking into account the NH

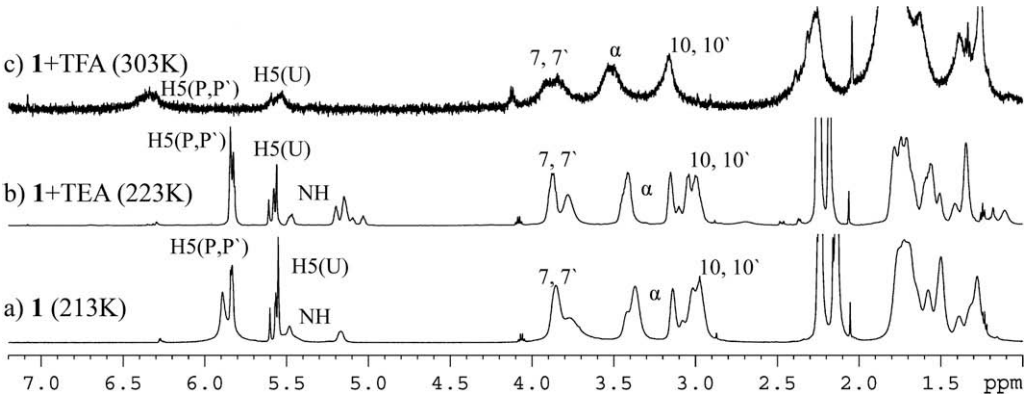


Figure 6. Dependence of the ^{1}H NMR spectra of **1** in $CDCl_3$ on the addition of base/acid. In **1** + TFA (c) the NH proton is at 9 ppm (fragment not shown).

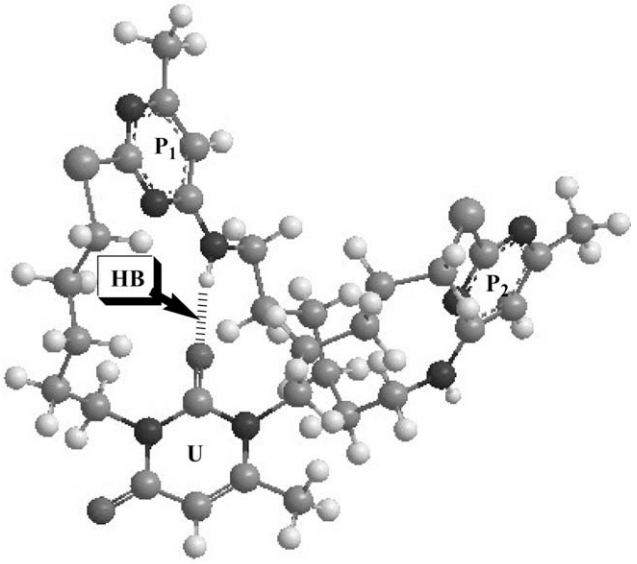


Figure 7. HF/6-31G optimized geometry of **2**.

chemical shifts of **4** in $CDCl_3$ and acetone. According to the ^1H NMR data of **4** there is no hydrogen bond in $CDCl_3$, while in acetone there is an intermolecular hydrogen bond between the NH and the O═C group. Thus, there are NH chemical shifts with and without hydrogen bonds. The DFT level of theory (B3LYP/6-31G(d,p)// B3LYP/6-31G(d,p)) gives a quite reliable evaluation of the hydrogen bond energy for a variety of hydrogen bonded pyrimidine pairs. Our calculations of the hydrogen bonding energy in the complex of **5** with acetone (NH–O═C, **5** is a simpler analogue of **4**) is ca. 6.7 kcal/mol. Thus, if one takes into account that the NH chemical shifts for **4** and **2** in acetone are close (6.9 and 7.05 ppm) and the NH_{HB} chemical shift in the hydrogen bonded conformation of **2** is 6.7 ppm, the hydrogen bond intramolecular interaction can be estimated as ca. 6–7 kcal/mol.

At the same time, the pH effect is crucial both for intramolecular packing and for the supramolecular structure. Protonation leads to: (1) stabilization of the *Z*-conformation which is not suited to hydrogen bonding of NH; (2) creates an additional strong proton donor centre ($N1H^+$); (3) produces a much stronger acceptor group—for the acid counterion group which begins to compete with the O═C group of uracil. All these factors lead to disruption of the intra- and to the creation of intermolecular hydrogen bonds in which the counterion serves as a non-covalent bridge between the nucleic acids moieties. This results in strong association in acid media both for model **4** (according to diffusivity measurements, a 16-fold increase of volume is observed) and in much more extended macrocycles (a volume increase of about 6 orders). According to diffusivity data in polar solvents, these aggregates are disrupted.

In conclusion, the conformational structures of the title macrocycles are determined by a variety of factors: chemical structure (length of spacers), weak non-covalent interactions, pH and polarity of the solvent. In neutral non-polar solvents ($CDCl_3$), hydrogen bonding can impact on the overall stabilization of some conformations but a spatial fit to allow such a hydrogen bond seems to be a prerequisite (spacers should allow such hydrogen bonding or should be flexible enough). However, an increase of spacer flexibility may result in an increase of the negative entropic contributions and in the loss of conformational preference. Protonation disrupts intramolecular hydrogen bonding, destabilizes the folded conformation, and leads to strong counterion-assisted (mediated) self-aggregation of macrocycles which can be partly destroyed only in polar solvents.

The above results demonstrate how complicated the conformational and supramolecular behaviour of such macrocycles is in solution and how it depends on internal/external factors. In fact, these results represent the very first step in studies of such biologically relevant compounds in order to proceed to water solutions which are more complicated due to a number of effects (hydrogen bonding, weak proton donor/acceptor, etc.) and will be the subject of future investigations.

Acknowledgements

One of us (Sh.L.) acknowledges the Russian Science Support Foundation for a doctoral grant. This investigation was carried out in the NMR department of the Federal collective spectral analysis centre for physical and chemical investigations of structure, properties and composition of matter and materials (CKP SAC), and Federal CKP for physical and chemical investigations of matter and materials (FCKP PCI) (state contracts of the Russian Federation Ministry of education and science Nos. 02.451.11.7036 and 02.451.11.7019).

References and notes

1. Gloe, K. *Macrocyclic Chemistry. Current Trends and Future Perspectives*; Springer: Netherlands, 2005.
2. Lehn, J. M.; Atwood, J. L.; Davies, J. E. D.; MacNicol, D. D.; Vögtle, F. *Comprehensive Supramolecular Chemistry. Molecular Recognition: Receptors for Molecular Guests*; Pergamon Press: Oxford, 1996.
3. Lehn, J. M. *Supramolecular Chemistry Concepts and Perspectives*; VCH: Weinheim, 1995.
4. Murakami, Y.; Kikuchi, J.-I.; Hisaeda, Y.; Hayashida, O. *Chem. Rev.* **1996**, *96*, 721.
5. Sivakova, S.; Rowan, S. J. *Chem. Soc. Rev.* **2005**, *34*, 9.
6. Jason, R.; Thomas, J.; Paul, J. *Chem. Rev.* **2008**, *108*, 1171.
7. Muller-Dethlefs, K.; Hobza, P. *Chem. Rev.* **2000**, *100*, 143.
8. Hobza, P. *Chem. Rev.* **1999**, *99*, 3247.
9. Zendlova, L.; Hobza, P.; Kabelac, M. *J. Phys. Chem. B* **2007**, *111*, 2591.
10. Marushkevich, K.; Khriachtchev, L.; Rasanen, M. *J. Phys. Chem. A* **2007**, *111*, 2040.
11. Schlucker, S.; Koster, J.; Singh, R. K.; Asthana, B. P. *J. Phys. Chem. A* **2007**, *111*, 5185.
12. Mizutani, T.; Wada, K.; Kitagawa, S. *J. Am. Chem. Soc.* **1999**, *121*, 11425.
13. Stephan, H.; Kruger-Rambuscn, T.; Gloe, K.; Hasse, W.; Ahlers, B.; Cammann, K.; Rissanen, K.; Brodesser, G.; Vogtle, F. *Eur. Chem. J.* **1998**, *4*, 434.
14. Shi, Y.; Schneider, H.-J. *J. Chem. Soc., Perkin Trans. 2* **1999**, 1797.
15. Itahara, T. *Bull. Chem. Soc. Jpn.* **1996**, *69*, 3239.
16. Galiullina, L.; Nikolaev, A.; Semenov, V.; Reznik, V.; Latypov, Sh. *Tetrahedron* **2006**, *62*, 7021.
17. Semenov, V.; Galiullina, L.; Lodochnikova, O.; Kataeva, O.; Gubaidullin, A.; Chernova, A.; Efremov, Yu.; Latypov, Sh.; Reznik, V. *Eur. J. Org. Chem.* **2007**, 4578.
18. Shagidullin, R. R.; Chernova, A. V.; Bazhanova, Z. G.; Lii, J. H.; Kataev, V. E.; Katsyuba, S. A.; Reznik, V. S. *J. Mol. Struct.* **2004**, *707*, 1.
19. Kozlov, A. V.; Semenov, V. E.; Mikhailov, A. S.; Aganov, A. V.; Smith, M. B.; Reznik, V. S.; Latypov, Sh. K. *J. Phys. Chem. B* **2008**, *112*, 3259.
20. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. GAUSSIAN *98, Revision A.3*; Gaussian, Inc.: Pittsburgh, PA, 1998.
21. According to our preliminary data the same protonation effects on ^1H NMR line shapes are observed with several acids ($ClCH_2COOH$, $4\text{-}CH_3C_6H_4SO_2OH$, HCl).