First principles molecular dynamics study of catalytic reactions of biological macromolecular systems: toward analyses with QM/MM hybrid molecular simulations

2007 J. Phys.: Condens. Matter 19 365217

(http://iopscience.iop.org/0953-8984/19/36/365217)

IOP PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **19** (2007) 365217 (8pp) doi:10.1088/0953-8984/19/36/365217

First principles molecular dynamics study of catalytic reactions of biological macromolecular systems: toward analyses with QM/MM hybrid molecular simulations

Mauro Boero[1,2,3], **Jung Mee Park**[3,4], **Yohsuke Hagiwara**[3,5] **and Masaru Tateno**[1,2,3]

[1] Center for Computational Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan
[2] CREST, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan
[3] Graduate School of Pure and Applied Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8571, Japan
[4] Special Nanoscience Project, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8571, Japan
[5] Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Nagatsuta 4259, Midori-ku, Yokohama 226-8501, Japan

E-mail: tateno@ccs.tsukuba.ac.jp

Received 20 December 2006, in final form 22 February 2007
Published 24 August 2007
Online at stacks.iop.org/JPhysCM/19/365217

Abstract

First principles molecular dynamics simulations performed on a fully solvated RNA model structure allowed us to investigate the mechanism for enzymatic cleavage reactions, *in vitro*, of RNA enzymes (ribozymes). The concerted action of two metal catalysts turns out to be the most efficient way to promote, on the one hand, the proton abstraction from 2′-OH that triggers the nucleophilic attack and, on the other hand, the cleavage of the P–$O^{5'}$ bond. In fact, the elimination of one of the two metal cations leads to an increase in the activation energy of the reaction. The simulated pathway shows that an OH^- in the coordination shell of the Mg^{2+} close to $O^{2'}$ promotes the initial proton abstraction and prevents its transfer to the ribozyme. This suggests that, in a real ribozyme, the double-metal-ion reaction mechanism is preferred with respect to single-metal-ion mechanisms either in the presence or in absence of the OH^- anion. Finally, an insight into the importance of hybrid quantum mechanics/molecular mechanics (QM/MM) schemes is discussed in view of the modelling of a realistic system carrying all the features of a true ribozyme.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Biological macromolecular systems such as protein, RNA, and DNA have gained an enormous interest in molecular biology, in medical science, and also in material science, since their self-assembled character is suitable for possible application in nanodevices.

As far as RNA (figure 1) is concerned, its enzyme molecules (ribozymes) represent forefront research in cancer gene therapy [1–9]. In fact, ribozymes can be designed to cleave other target RNA molecules and show a remarkable ability to inhibit gene expression. The chemical reaction operated by all the self-cleaving ribozymes is the hydrolysis of the RNA phosphodiester, resulting in the cut of the RNA at a specific target site (transesterification). The main steps of the transesterification are summarized in figure 2 and, in most of such reactions *in vitro*, metal ions are known to play an important catalytic role, although still unsolved to various extents. In fact, the complexity of the various ribozyme systems and of the catalytic reaction involved poses severe obstacles to direct experimental probes. It is then not surprising that, despite extensive efforts, the intimate details of the reaction mechanism governing the transesterification have mainly been deduced on the basis of assumptions inferred from x-ray structural data, partly supported by static quantum calculations on simplified models.

Recent progress in computer hardware and software algorithms make it now possible to study in great detail such complicated chemical reactions. In this respect, the reaction pathway representing the main subject of the present paper and the related catalytic role of Mg^{2+} metal cations could be studied in great detail for the first time at a first principles molecular dynamics level, evidencing how divalent metal cations can play a central role in withdrawing electrons from the cleavage site and, thus, lowering the activation barrier [10, 11].

On the DNA front, research has focused mainly on the oxidative damage and electrical conductivity [12–15]. Again, a clue to interpreting the experimental results seems to be the role of metal counterions, which are expected to alter the electronic properties of pure DNA. In particular, although metallic cations are generally passivated by hydration, fluctuations of the solvation shell of counterions or partial dehydration processes can induce a doping of holes in DNA fibres [16], underscoring the general importance of metal ions in nucleic acids.

2. Computational details

We performed first principles molecular dynamics simulations [17], in the framework of the density functional theory (DFT), with gradient corrections on the exchange and correlation functional after Hamprecht, Cohen, Tozer and Handy (HCTH) [18]. Valence–core interactions are described by Troullier–Martins [19] norm-conserving pseudopotentials, and valence electrons are expanded in plane waves with an energy cut-off of 70 Ryd. The simulated systems consist of (N, V, T) ensembles containing one RNA *anion* unit. Then metal ions, with and without water, at the standard density value (1.0 g cm^{-3}) are added in a cubic supercell of side $L = 13.146$ Å with periodic boundary conditions. The temperature control ($T = 300$ K) was ensured by a Nosé–Hoover thermostat [20, 21]. An integration step of 4.0 au (0.0967 fs) with a fictitious electron mass of 400 au provided good control of the conserved quantities. The reaction path was sampled in two ways. (1) By using the Blue Moon ensemble method [22], assuming as a reaction coordinate the distance $\xi = |\text{P–O}^{2'}|$ and extending the Car–Parrinello Lagrangian L^{CP} as $L = L^{\text{CP}} + \lambda_\xi(\xi - \xi_0)$, where λ_ξ is a Lagrange multiplier. The free energy ΔF of the reaction from the initial $\xi(a)$ state to the final $\xi(b)$ state is given by

$$\Delta F = \int_{\xi(a)}^{\xi(b)} \langle \lambda_\xi \rangle \, \mathrm{d}\xi, \tag{1}$$

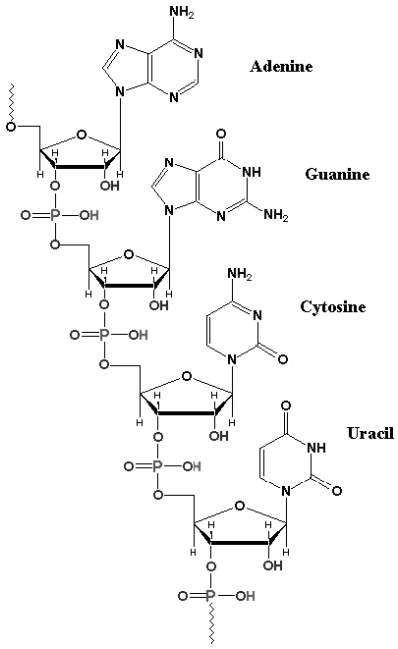

Figure 1. Scheme of a single strand of a typical RNA structure with the four bases.

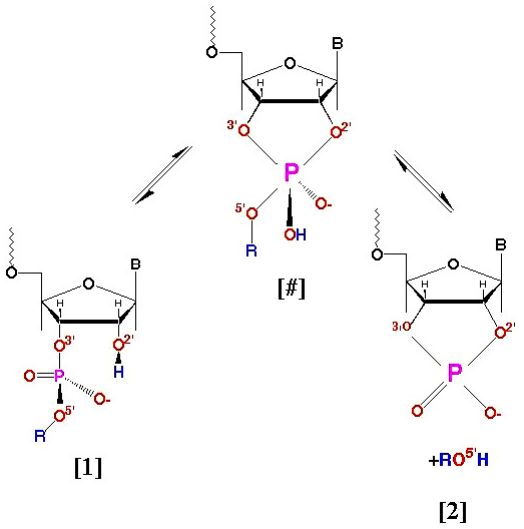

Figure 2. Main phases of the RNA self-cleavage reaction mechanisms [1]. The initial structure, [#] the trigonal-bipyramidal (TBP) transition state, and [2] the final cleaved product.

$\langle \cdots \rangle$ being the time average. (2) Via the metadynamics approach [23], by adding to L^{CP} the harmonic degrees of freedom of the collective variables s_α plus a history-dependent Gaussian potential $V(s_\alpha, t)$

$$L = L^{\mathrm{CP}} + \sum_\alpha \tfrac{1}{2} M_\alpha \dot{s}_\alpha^2(q) - \sum_\alpha \tfrac{1}{2} k_\alpha \left[s_\alpha(q) - s_\alpha^0 \right]^2 - V(s_\alpha, t) \tag{2}$$

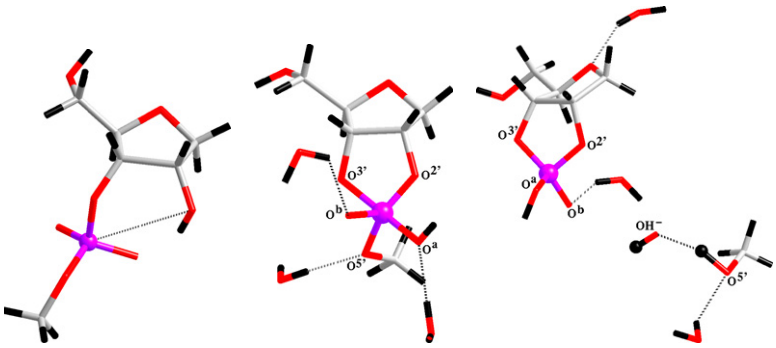

Figure 3. The main snapshots of the simulated transesterification reaction. (a) The initial equilibrated system, (b) the TBP transition state and (c) the cleaved final product. The colour code is black for H, red for O, grey for C and purple for P.

in which q represents any analytical function of the atomic coordinates. In our case, the collective variables s_α are the three coordination numbers that account for the formation of the bond between P and $O^{2'}$, the cleavage of the P–$O^{5'}$ bond and the 2′-OH proton abstraction. The free energy is obtained as

$$\lim_{t\to\infty} V(s_\alpha, t) = F(s_\alpha) + \text{constant} \tag{3}$$

where the limit has to be intended as a sufficiently long simulation time allowing for a complete exploration of both the reactant and the product local minima.

3. Results and discussion

We first equilibrated, by standard unconstrained first principles molecular dynamics simulations, the dry anionic system with a Mg^{2+} cation shown in panel (a) of figure 3. Although the lack of solvent water makes this simulation somehow artificial, this allows us to disentangle the effect of the metal from the influence of the solvent and to observe the Mg 3s energy level shift upon drying processes. When the equilibrium was reached, after about 2.0 ps, we sampled the reaction path that leads to the formation of the trigonal bipyramidal phosphorane (TBP) [1, 2] (panel (b) in figure 3) by approaching $O^{2'}$ to P via the Blue Moon sampling method. We could observe that, in the initial stage, the proton of the 2′-OH group is transferred from $O^{2'}$ to one of the two *dangling* O atoms bound to P (namely, $O^{(R)}$ and $O^{(S)}$ according to the standard notation of the literature) when the reaction coordinate P–$O^{2'}$ shrinks to ∼2.30 Å. As expected, a TBP is subsequently formed at $\xi = 2.00$ Å and the activation barriers computed for this process amount to $\Delta E = 27.6 \pm 2.0$ kcal mol^{-1}, $\Delta F = 23.9 \pm 1.9$ kcal mol^{-1} for total energy and free energy, respectively. The cleavage of the P–$O^{5'}$ bond occurs as shown in panel (c) in figure 3, completing the reaction and releasing the departing $O^{5'}$–R group.

An identical simulation on this same system performed in the absence of metal cations has shown that a much higher activation barrier ($\Delta E = 38.7 \pm 2.2$ kcal mol^{-1}, $\Delta F = 32.4 \pm 2.2$ kcal mol^{-1}) has to be overcome, and furthermore, that a 'wrong' reaction occurs, consisting of a pseudorotation [1, 2, 24] and a transfer of the phosphodiester from $O^{3'}$ to $O^{2'}$.

These results and a close inspection of the related Kohn–Sham electronic states (figure 4) suggest that divalent cations help in selecting the proper transesterification reaction path and in withdrawing electrons from the P–$O^{5'}$ bond. This bond is then weakened, and, hence, can be

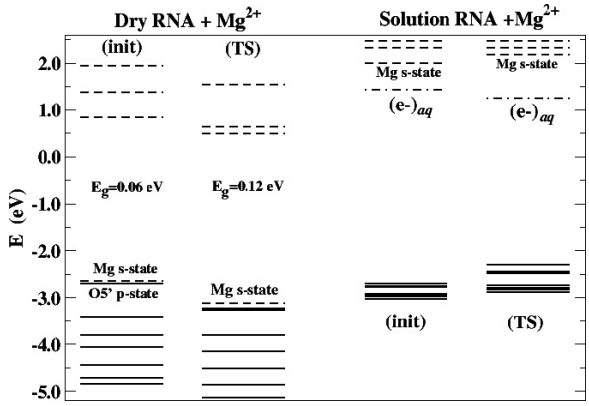

Figure 4. Kohn–Sham energy levels around the HOMO–LUMO gap. Solid lines indicate the occupied states, dashed lines the empty states. The labels (init) and (TS) stand for initial configuration and transition state, respectively. Details are discussed in the text.

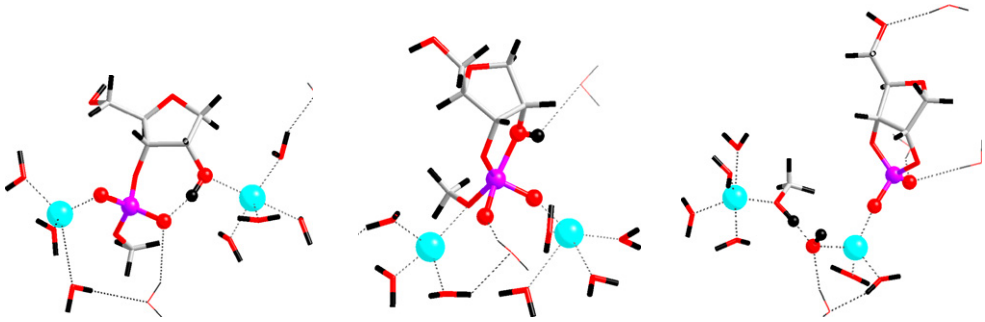

Figure 5. Main steps of the transesterification reaction in solution as obtained by metadynamics simulations. The initial state (a) passes across the TBP transition state (b) before reverting to the final product (c). The colour code is identical to figure 3 with the addition of Mg^{2+} ions shown as blue spheres.

more easily cleaved. We remark that, in the absence of solvent water molecules, the HOMO–LUMO gap is rather small, i.e. only 0.06 eV, in the initial configuration and enlarges to 0.12 eV at the transition state (TS), as shown in figure 4. The last occupied Kohn–Sham (KS) level is in both cases localized on the O atoms bound to P, while the first empty KS state is the s orbital of the Mg^{2+} that spreads over the neighbour O atoms of the RNA. Even in the pathological gap underestimation affecting all DFT-based calculations, these features and the general HOMO–LUMO trend seem to be a well assessed distinctive character of the system.

We then repeated the calculations in solution, in an attempt at inspecting the role of the solvent on both the reaction path and the electronic states. This second set of simulations has shown that the Mg^{2+} metal ion forms a solvation shell by coordinating H_2O molecules around it. Yet, the oxygen atoms of the RNA participate in the solvation shell of these cations and the average initial configuration is shown in panel (a) of figure 5. The reaction path was sampled via metadynamics, assuming as collective variables three coordination numbers representative of all the changes occurring in the chemical bonds. Namely, the coordination number between P and $O^{2'}$, accounting for the bond formation between these two sites, the coordination number

between P and $O^{5'}$, where the cleavage occurs, and the coordination number between $O^{2'}$ and its initially bound H atom, accounting for the initial deprotonation reaction. The most relevant (average) configurations are shown in figure 5. A proton is initially transferred from $O^{2'}$ to $O^{(S)}$ and subsequently $O^{2'}$ starts the nucleophilic attack to the phosphorous, leading to the formation of the TBP transition state (b). The last step of the reaction is represented by the cleavage of the P–$O^{5'}$ bond (figure 5(c)) in a way structurally not too different from the one evidenced in the dry system. The proton from $O^{2'}$ is eventually released to the $O^{5'}$ leaving group and this completes the reaction.

We verified that, also in solution, in the absence of metal cations, the overall barrier for the process is $\Delta E = 60.1$ kcal mol^{-1} and $\Delta F = 58.3$ kcal mol^{-1}, while with the help of Mg^{2+} this same barrier is lowered to $\Delta E = 46.5$ kcal mol^{-1} and $\Delta F = 44.7$ kcal mol^{-1}, closer to overall experimental barriers [1, 2, 8] (20–30 kcal mol^{-1}) which are, however, strongly dependent on the pH, Mg^{2+} binding and concentration, enzyme conformation, etc, and still rather controversial. In the presence of solvating water, the electronic structure close to the gap turned out to be very different: namely the HOMO level is still represented by the lone pairs of oxygen atoms bound to P, but occasionally it can switch position with the oxygen lone pairs of H_2O molecules of the solvent. On the other hand, the LUMO is always a diffuse orbital related to the solvated electron as described in [25], while LUMO + 1 and LUMO+ 2 are the (empty) levels due to the s orbital of the cation. In particular, the cation close to –$O^{2'}$–H withdraws electrons from the $O^{2'}$ oxygen, thus favouring the proton release. Analogously, the Mg^{2+} located in proximity of the P–$O^{5'}$ bond favours the cleavage reaction. The approaching of the LUMO orbital to the HOMO band is a phenomenon typical of drying processes: when the solvent is removed and the hydrated cation becomes dry, then the lowest unoccupied level is the empty orbital of the countercation. Indeed, a similar result was also found in a DNA system [16]. This seems to be a key issue in interpreting the experimental results: the degree of hydration and the amount of anhydrous metal countercations are expected to alter the electronic properties of nucleic acids. The resulting Kohn–Sham level distribution is shown in figure 4, where a direct comparison with the case of a dry system summarizes the above discussion graphically.

These analyses were performed on a simplified model system reduced to a prototype functional site in the whole biological macromolecule. However, for more realistic descriptions of the enzymatic reaction mechanisms, we should deal with entire structural models including all of the molecular structure necessary for the catalytic reactions. Of course, this would amount to an unfeasible computational cost if the whole systems were treated at a full quantum level. Such an effort, however, is not necessary. In the case of the ribozyme reaction, for instance, only a small portion of the system undergoes chemical reactions, while the rest of the system, although playing a role as source of hydrogen bonds and structural modifications, remains chemically unaltered [26]. This allows us to separate the two problems and to use hybrid quantum mechanics/molecular mechanics (QM/MM) approaches.

Beside the transesterification process, the role of an OH^- has also been shown to be essential in smoothing the Coulomb repulsion between two Mg^{2+} ion catalysts if they are too close to each other. Classical simulations have already attempted to address this issue [27]; however, difficulties in modelling an appropriate model potential for Mg^{2+} metal cations, able to take into account both polarization effects and their electron withdrawing properties, and OH^- anions make a pure force field approach a rather demanding task. In order to overcome these problems, we performed QM/MM simulations using a fully hydrated hammerhead ribozyme (figure 6). In the first simulation no OH^- was included and we simply solvated the (QM) Mg^{2+} ions with (QM) water molecules. A standard QM/MM molecular dynamics simulation performed to equilibrate the system showed that, in the presence of water only, one

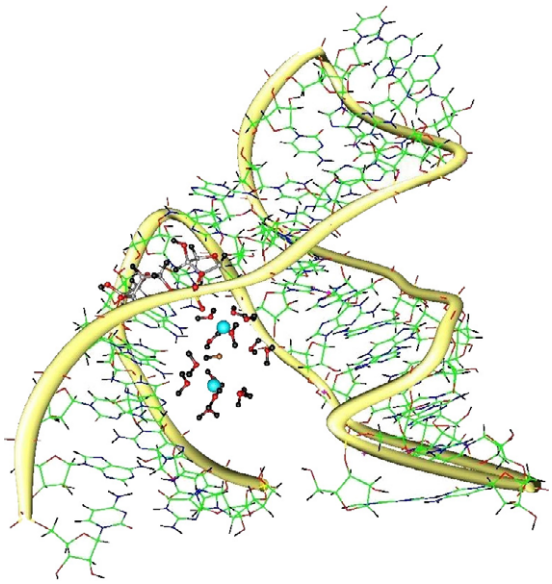

Figure 6. The hybrid QM/MM ribozyme system. Quantum atoms are shown as thick sticks and balls, classical atoms as thin sticks and the solid (yellow) ribbon represents the phosphate backbone. The colour code for atoms is the same as that of the previous figures.

of the two Mg^{2+} cations, namely the one closer to the phosphate, and having in its solvation shell the oxygen $O^{(R)}$, reaches a stable position at an average distance of about 3.5 Å from the P atom and 2.1 Å from $O^{(R)}$. At the same time, the second metal ion departs rather quickly from its initial position, deduced from x-ray data, and moves away from the other Mg^{2+}. This seems to indicate that, in the absence of an OH^- between the two Mg^{2+} ions, whose positions were deduced from the experimental crystal structure, only one of the two metal cations can remain close enough to the ribozyme to participate in the catalysis. Instead, if an OH^- is located between the two Mg^{2+} cations, the hydroxyl anion participates simultaneously in the solvation shell of both the cations. The net effect results in a screening of the two positively charged Mg^{2+} ions that reduces the Coulomb repulsion between them and keeps both the ions close to the catalytic centre. The conclusion that can be drawn from these simulations is that, besides the important catalytic properties described in the previous paragraph, a double-metal-ion mechanism also includes the participation of an OH^- anion as a fundamental ingredient.

Summarizing, we have shown here two examples of applications of first principles calculations to biological macromolecules. As for the RNA enzymatic reactions, first principles molecular dynamics calculations allow us to inspect in great detail both the atomic and electronic states of rather complicated reactions, such as the transesterification of RNA enzymes. The simulations show that metal catalysts, in either a dry or a wet environment, enhance the reaction rate with a high selectivity. The absence of cations not only gives rise to a remarkable increase of the activation energy barrier, but can also result in an incorrect reaction pathway. The study presented here, part of which is still in progress, can provide a clue as to the phosphate hydrolysis mechanisms of a large class of biological systems. From the point of view of the electronic state, the shift of the unoccupied s levels toward the valence band, occurring in drying conditions, can shed some light on both the bond cleavage processes and on the general role of metal counterions in nucleic acids.

Acknowledgments

This work is partly supported by grants-in-aid from the Japan Ministry of Education, Culture, Sports, Science and Technology (MEXT) under contract number 18036003. Computations were partly done by using the Computer Center for Agriculture, Forestry, and Fisheries Research (AFFRC) of the Japan Ministry of Agriculture, Forestry, and Fisheries.

References

[1] Perreault D M and Anslyn E V 1997 *Angew. Chem. Int. Edn Engl.* **36** 433
[2] Zhou D and Taira K 1998 *Chem. Rev.* **98** 991
[3] Uhlenbeck O C 1987 *Nature* **328** 596
[4] Haseloff J and Gerlach W L 1988 *Nature* **334** 585
[5] Fedor M J and Williamson J R 2005 *Nat. Rev. Mol. Cell. Biol.* **6** 399
[6] Pley H W, Flaherty K M and McKay D B 1994 *Nature* **372** 68
[7] Murray J B, Terwey D P, Maloney L, Karpeisky A, Usman N, Biegelman L and Scott W G 1998 *Cell* **92** 665
[8] Noodleman L, Lovell T, Han W G, Li J and Himo F 2004 *Chem. Rev.* **104** 459
[9] Pyle A M 2002 *J. Biol. Inorg. Chem.* **7** 679
[10] Boero M, Terakura K and Tateno M 2002 *J. Am. Chem. Soc.* **124** 8949
[11] Boero M, Tateno M, Terakura K and Oshiyama A 2005 *J. Chem. Theory Comput.* **1** 925
[12] Meggers E, Michiel-Beyerle M E and Giese B 1998 *J. Am. Chem. Soc.* **120** 12950
[13] Conwell E M and Rakhmanova S V 2000 *Proc. Natl. Acad. Sci. USA* **97** 4556
[14] Giese B, Amaudrut J, Köhler A K, Spormann M and Wesseley S 2001 *Nature* **412** 318
[15] Gervasio F L, Boero M and Parrinello M 2006 *Angew. Chem. Int. Edn Engl.* **45** 5606
[16] Gervasio F L, Laio A, Parrinello M and Boero M 2005 *Phys. Rev. Lett.* **89** 158103
[17] Car R and Parrinello M 1985 *Phys. Rev. Lett.* **55** 2471
[18] Hamprecht F A, Cohen A J, Tozer D J and Handy N C 1998 *J. Chem. Phys.* **109** 6264
[19] Troullier N and Martins J L 1991 *Phys. Rev.* B **41** 1993
[20] Nosé S 1984 *Mol. Phys.* **52** 255
[21] Hoover W G 1985 *Phys. Rev.* B **31** 1695
[22] Sprik M and Ciccotti G 1998 *J. Chem. Phys.* **109** 7737
[23] Iannuzzi M, Laio A and Parrinello M 2003 *Phys. Rev. Lett.* **90** 238302
[24] Draper D E 2004 *RNA* **10** 335
[25] Boero M, Parrinello M, Terakura K, Ikeshoji T and Liew C C 2003 *Phys. Rev. Lett.* **90** 226403
[26] Woodson S A 2005 *Curr. Opin. Chem. Biol.* **9** 104
[27] Hermann T, Auffinger P, Scott W G and Westhof E 1997 *Nucleic Acids Res.* **25** 3421