CrossMark
click for updates

Stability of 2′,3′ and 3′,5′ cyclic nucleotides in formamide and in water: a theoretical insight into the factors controlling the accumulation of nucleic acid building blocks in a prebiotic pool†

Giuseppe Cassone,*[a] Jiří Šponer,[ab] Franz Saija,[c] Ernesto Di Mauro,[d] A. Marco Saitta[e] and Judit E. Šponer*[a]

Synthesis of the first RNAs represents one of the cornerstones of the emergence of life. Recent studies demonstrated powerful scenarios of prebiotic synthesis of cyclic nucleotides in aqueous and formamide environments. This raised a question about their thermodynamic stability, a decisive factor determining their accumulation in a prebiotic pool. Here we performed *ab initio* molecular dynamics simulations at various temperatures in formamide and water to study the relative stabilities of the 2′,3′ and 3′,5′ isomers of cyclic nucleotides. The computations show that in an aqueous environment 2′,3′ cyclic nucleotides are more stable than their 3′,5′ counterparts at all temperatures up to the boiling point. In contrast, in formamide higher temperatures favor the accumulation of the 3′,5′ cyclic form, whereas below about 400 K the 2′,3′ cyclic form becomes more stable. The latter observation is consistent with a formamide-based origin scenario, suggesting that 3′,5′ cyclic nucleotides accumulated at higher temperatures subsequently allowed oligomerization reactions after fast cooling to lower temperatures. A statistical analysis of the geometrical parameters of the solutes indicates that thermodynamics of cyclic nucleotides in aqueous and formamide environments are dictated by the floppiness of the molecules rather than by the ring strain of the cyclic phosphodiester linkages.

Introduction

Cyclic nucleotides have been long considered as plausible building blocks for the synthesis of the first RNA oligomers.[1,2] Their strained intramolecular phosphodiester linkage accumulates chemical energy which can be utilized to drive transphosphory-lation reactions leading to oligonucleotides.[3] Recently, various plausible prebiotic pathways have been elaborated for their synthesis,[4–6] among which the one based on formamide[5,7,8] has been found to yield a mixture of both 2′,3′ and 3′,5′-linked

cyclic forms. Based on these experimental results it has been suggested that higher temperatures of the prebiotic environment favor the accumulation of the 3′,5′ form over the 2′,3′ variant.[5] This observation fostered the efforts aimed at the oligomerization of 3′,5′ cyclic nucleotides, which has been demonstrated on the example of 3′,5′ cyclic guanosine monophosphate (GMP) under a variety of conditions.[9–12] Since the anionic ring opening polymerization of 3′,5′ cyclic GMPs is so far the only known chemistry that leads to selectively 3′,5′-linked oligomers (all other known methods lead to a mixture of 2′,5′- and 3′,5′-linkages),[12] it provides a plausible answer to the question about the evolutionary selection pressure which led to the 3′,5′-linkage selectivity in modern RNA-molecules.

The phosphodiester bond of 2′,3′ cyclic nucleotides is confined in a 5-membered ring in contrast to that of 3′,5′ cyclic nucleotides, where it is embedded in a less strained 6-membered ring.[13,14] This allowed us to assume that 2′,3′ cyclic nucleotides are thermo-dynamically less stable than their 3′,5′-linked counterparts. Some-what surprisingly early calorimetric measurements systematically showed the opposite trend: the hydrolysis heat of 3′,5′ cyclic nucleotides being markedly more exothermic than that of the 2′,3′ form.[3,15] This suggested that solvation effects could be decisively important in addressing the thermodynamic stability of 3′,5′ cyclic nucleotides.[16]

[a] *Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, CZ-612 65 Brno, Czech Republic. E-mail: judit@ncbr.muni.cz, Giuseppe.Cassone@impmc.upmc.fr*

[b] *Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacky University, 17. Listopadu, 1192/12, 771 46 Olomouc, Czech Republic*

[c] *CNR-IPCF, Viale Ferdinando Stagno d'Alcontres 37, 98158 Messina, Italy*

[d] *Dipartimento di Scienze Ecologiche e Biologiche, Università della Tuscia, Via San Camillo De Lellis, 01100 Viterbo, Italy*

[e] *Sorbonne Universités, Université Pierre et Marie Curie Paris 6, CNRS, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, Muséum National d'Histoire Naturelle, Institut de Recherche pour le Développement, UMR 7590, F-75005 Paris, France*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6cp07993h

This journal is © the Owner Societies 2017

Phys. Chem. Chem. Phys., 2017, **19**, 1817–1825 | 1817

In the current paper we use state-of-the-art Born–Oppenheimer Molecular Dynamics simulations to address the relative stability of the 2′,3′ and 3′,5′ cyclic nucleotides as a function of the temperature in two prebiotically relevant environments. This quantum-based Molecular Dynamics method enables a unique description of the electronic effects stabilizing cyclic nucleotides in solution, which includes a complete statistical treatment of the solvation shell by explicitly treated solvent molecules. In addition, it allows us to take into account temperature effects.

Due to the general interest in cyclic nucleotides as substrates for non-enzymatic, template-free oligomerization scenarios relevant to the emergence of the most ancient oligonucleotides from the prebiotic pool, we decided to investigate their thermodynamic stability in water and formamide. These are the two solvents most frequently considered in the recent origin of life studies.[17–22] We will show that while aqueous solutions and lower temperatures could favor the accumulation of 2′,3′ cyclic nucleotides over that of the 3′,5′ cyclic form, at higher temperatures and in formamide solvent 3′,5′ cyclic nucleotides are the thermodynamically preferred species. Thus, after a temperature drop,[23] initial cycles of polymerization leading to the first RNA molecules could indeed occur in a 3′,5′-cyclic-nucleotide rich environment.

Models

In order to explain systematic stability differences caused by different binding modes of the phosphate moiety in 2′,3′ and 3′,5′ cyclic nucleotides, we have simplified the nucleotides to their corresponding phosphorylated sugar part: *i.e.* the nucleobase part has been substituted with a hydrogen atom (see Scheme 1). We think that this simplification is justified since the calorimetric studies by Rudolph *et al.*[3] suggest that the systematic differences in the experimentally measured hydrolysis heats originate from the different constitution of the phosphorylated sugar of the 2′,3′ and 3′,5′ cyclic nucleotides. A total charge of −1 (equivalent to a non-protonated phosphate form) has been assumed throughout the calculations. This reduction of the model system is important to make the computations feasible.
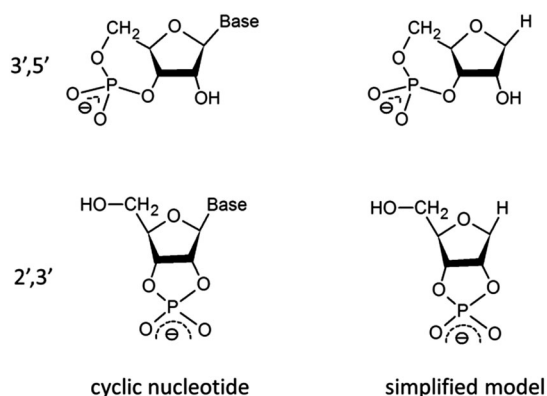
Scheme 1 Schematic drawings of cyclic nucleotides and their simplified models used in the computations.

Methods

Using the Born–Oppenheimer (BO) and Car–Parrinello[24] (CP) approaches we performed a series of *ab initio* Molecular Dynamics (AIMD) simulations on simplified models of 2′,3′ and 3′,5′ cyclic nucleotides both in liquid formamide and water at several temperatures. All results were obtained using the software suite Quantum ESPRESSO[25] and a plane-wave/pseudopotential scheme.

Simulations in formamide

The phosphorylated sugar models described above along with the solvent molecules were placed in a cubic simulation box. Two different simulation box sizes (with sides of 11.0 and 12.7 Å, respectively) were chosen in order to assess the effect of the simulation box size on the computed results. In the smaller model each cyclic nucleotide has been surrounded by 16 formamide molecules while in the box at lower concentration (*i.e.* in the wider simulation cell) 26 formamide molecules have solvated each cyclic form. These two choices led us to deal with samples composed of 116 and 176 atoms, respectively. Let us note that computational treatment of systems of this size is rather challenging for state-of-the-art Born–Oppenheimer Molecular Dynamics (BOMD) simulations because of the presence of several heavy atoms. All these settings reproduce the typical formamide density of 1.13 g cm^{-3}. Periodic boundary conditions have been used in all simulations. Since both cyclic nucleotides models carry a charge of −1, a compensating jellium background[26] has been added in order to avoid the divergences due to the periodic infinite replica of the (charged) simulation box.

Before starting the AIMD, all boxes were set up and equilibrated at first *via* a classical force field. Both for the formamide and water (see below) samples we employed a common force field based on the typical Optimized Potential for Liquid Simulations (OPLSs),[27] containing Lennard-Jones and Coulombic terms. The cross-interactions were calculated using the standard Lorentz–Berthelot rule. The classical MD equilibration runs were executed for time-lengths of the order from at least several hundreds of ps to several ns.

The BOMD simulations were performed at four fixed temperatures (*i.e.* at 363 K, 403 K, 453 K, and 483 K) by exploiting the velocity rescaling method. As usual, the dynamics of the nuclei was simulated classically within a constant number, volume, and temperature (*NVT*) ensemble using the Verlet algorithm and acquiring a total time-length for each simulation box of 20 ps. A timestep of 0.48 fs has been used. A cutoff energy of 40 and 320 Rydberg (Ry) has been set for the wavefunctions and the charge density, respectively. With such cutoff values, the simulated system is described in a realistic way since the core electronic structure interaction is treated with ultrasoft pseudopotentials (USPPs). The Perdew–Burke–Ernzerhof (PBE)[28] exchange and correlation (XC) functional, belonging to the generalized gradient approximation class, has been used throughout the whole study. This choice is justified not only by its known reliability for H-bonded systems[29] (because of a proper description of polarization)[30] but also because of the adherence of previous

simulations carried out in neat liquid formamide[31] with the available experimental data.[32,33] Moreover, Grimme's DFT-D2 semiempirical corrections[34] were added to the PBE functional in order to properly handle possible dispersion interactions. This approximation ensures an adequate treatment of all four components (covalent, electrostatic, dispersion and polarization) of the H-bonding interactions.[35]

In order to ensure the most converged collection of data, the total energy averages have been extracted from the last 5 ps of each simulation and an estimate of the errors has been carried out (see the ESI†).

By exploiting the same BOMD settings (except, of course, for the timestep), additional Car–Parrinello Molecular Dynamics[24] (CPMD) simulations have been performed at 363 K and at 483 K in order to accumulate a wider range of structural data for statistics. The timestep in these "numerical experiments" has been set to 0.12 fs in order to obtain a reliable electronic description by assuming a fictitious electronic mass of 300 a.u. Finally, the sampled temperatures were kept fixed through the coupling of the system with a Nosé–Hoover thermostat, the frequency of which was set to 13.5 THz. These longer CPMD simulations showed the same trend of the computed time-averaged number of H-bonds between the solute and solvent molecules (see the Results and Discussion part) as that revealed by BOMD calculations. Since CPMD can be considered as a surrogate of the more exact BOMD,[36] we will describe here only the more reliable BOMD results.

Simulations in water

Simulations in liquid water were executed using similar DFT technical settings to those described for formamide. Thus, our phosphorylated sugar models along with 40 solvating water molecules were placed into a cubic simulation box with a length of 11.5 Å. This model was composed of 140 atoms and reproduced the nominal water density of 1.00 g cm^{-3}. BOMD simulations were conducted at 298 and 363 K. All the remaining technical settings (*i.e.* thermodynamic ensemble, timestep, XC functional, *etc.*) are the same as those provided in the previous subsection for the BOMD calculations in liquid formamide. The PBE XC functional was also chosen in this case because of its reliability in describing water, at least for the purposes of the present work. It is indeed one of the most tested XC functionals for aqueous systems and its limitations are well documented in the literature.[37–39] This functional adequately describes polarization[27] and hence it has been employed in many studies of neat water[40,41] (leading to results comparable to other widely employed XC functionals[42]) and of other simple H-bonded systems,[43] including a study dealing with the *in silico* reconstruction of the historical Miller experiment.[44]

Results and discussion

Stability of cyclic nucleotides in water

Table 1 compares the relative stability of 2′,3′ and 3′,5′ cyclic nucleotides derived from BOMD simulations performed in water at room temperature as well as at the temperature where

Table 1 Computed relative stabilities (in kcal mol^{-1}) and the average number of H-bonds between the studied 2′,3′ and 3′,5′ cyclic nucleotide models and the solvent molecules in water. Relative stabilities are listed at 298 and 363 K, for structural criteria defining an H-bond see the text

| Temperature (K)	Relative stability (kcal mol^{-1})		Average number of H-bonds	
	2′,3′	3′,5′	2′,3′	3′,5′
298	0	10 ± 2	6.6	6.0
363	0	6 ± 2	6.2	5.6

most oligomerization experiments reported in the literature have been conducted.[1,9–12] Albeit the energy data derived from BOMD are approximate (our estimation for the error bar referred to the energetically converged part of the simulations is about 2–3 kcal mol^{-1}, see the ESI†), the computations show that in an aqueous environment, the 2′,3′ cyclic nucleotides are more stable than their 3′,5′ cyclic counterpart. Thus, in this environment the relative stability of cyclic nucleotides follows the trend which one can derive from hydrolysis heats[3] rather than from the extent of the ring strain of the cyclic phospho-diester linkages.[13]

To resolve the controversy[15] between the trend of theoretically expected intrinsic ring-stabilities and that of the experimentally measured[3] (or our computed) thermodynamic data, we have analyzed the differences in the solvation of the two cyclic nucleotide forms. In particular, we have looked at the network of H-bonds formed between the solute and solvent molecules. To define an H-bond we have adopted the widely used (though obviously arbitrary) geometrical criteria by Desiraju.[45]

Four types of H-bonds can be found in the simulated models. The most common are strong N–H⋯O and O–H⋯O bonds as well as the markedly weaker C–H⋯O and C–H⋯N type of H-bonds. Since the latter two types of H-bonds can be assumed to contribute to the overall stabilization of the cyclic nucleotides rather negligibly, we have excluded them from our structural analyses. Hence, we performed a structural analysis of the H-bonding network that solvates the phosphorylated sugars and identified H-bonds by using the following criteria: $r(X\cdots Y) \leq 3.1$ Å and X–H⋯Y $\geq 145°$, where X = O, N is the H-bond donor and Y = O, N is the H-bond acceptor (for illustration of the parameter choice, see a typical simulation snapshot in Fig. 1).[45]

In addition, we checked whether our results were sensitive to a slight change in the choice of the above geometric criteria. We have analyzed the radial distribution function for the distances between the anionic oxygens of the phosphate group and the oxygen of the solvating water molecules. As shown in Fig. 2, the first maximum of the function is at 2.7 Å, the typical value found in pure liquid water, followed by a minimum in the range of 3.3–3.4 Å. Even by adopting the latter values as the cutoff lengths for the definition of the H-bonds, the results of the H-bonding analysis were identical to those obtained with the criteria taken from ref. 45.

Columns 4 and 5 of Table 1 list the average number of H-bonds formed between the solute and water molecules computed based on the above definition at 298 K and 363 K.
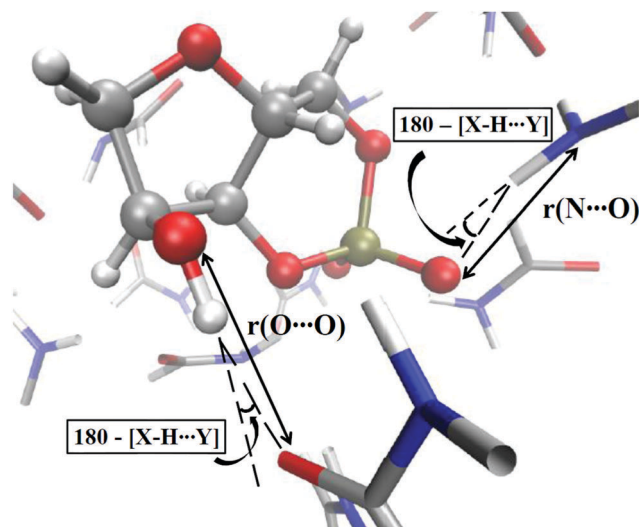
Fig. 1 Snapshot of a trajectory computed for the 3′,5′ cyclic nucleotide model in formamide. The relevant geometrical parameters defining an H-bond are highlighted. In order to better visualize the angle formed between the H-bond donor and acceptor atoms, the quantity $180 - [X-H\cdots Y]$ has been depicted. An H-bond is defined as follows: $r(X\cdots Y) \leq 3.1$ Å and $X-H\cdots Y \geq 145°$, where X = O, N is the H-bond donor and Y = O, N is the H-bond acceptor.[45]

From a comparison of the corresponding structural and energy data of Table 1, it is obvious that there is a correlation between the stabilization of cyclic phosphorylated sugars and the number of H-bonds formed with solvent water molecules. At both studied temperatures the more stable 2′,3′ cyclic form is stabilized by a noticeably higher number (on average 0.6) of H-bonds.
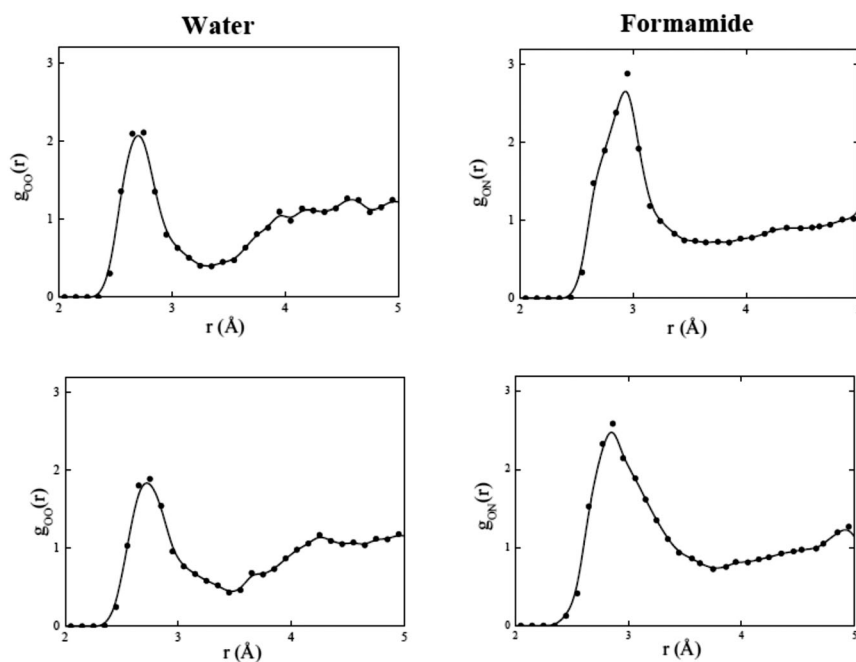
The experimentally determined hydrolysis heats of 3′,5′ cyclic nucleotides vary from 10.5 to 14.1 kcal mol^{-1} (in absolute value) and the corresponding data for the 2′,3′ cyclic forms fall in the range of 7.8–9.5 kcal mol^{-1}.[3] It means that the stability difference of the 2′,3′ and 3′,5′ forms is on average about 4 kcal mol^{-1},[3] which is roughly equivalent with the enthalpic contribution of 0.8 H-bonds.[46] Our theoretical prediction, *i.e.* a difference of about 0.6 in the number of H-bonds stabilizing the two cyclic forms, is thus in qualitative agreement with the experimental observation.

Stability of cyclic nucleotides in formamide

A great advantage of formamide as a solvent in organic chemistry lies in its considerably higher boiling point (483 K *vs.* 373 K of water) that enables chemical synthesis in a much broader temperature range. Indeed, the formamide-based synthesis experiments of prebiotic building blocks have commonly been conducted at 433 K.[7] More detailed explanation of the potential

Table 2 Temperature dependence of the computed relative stabilities (in kcal mol^{-1}) and the average number of H-bonds between the studied 2′,3′ and 3′,5′ cyclic nucleotide models and the solvent molecules in formamide. For structural criteria defining an H-bond, see the text and Fig. 1

	Relative stability (kcal mol^{-1})		Average number of H-bonds	
Temperature (K)	2′,3′	3′,5′	2′,3′	3′,5′
363	0	7 ± 3	4.2	3.8
403	4 ± 3	0	4.1	4.2
453	8 ± 3	0	3.7	4.2
483	9 ± 2	0	2.4	2.9

Fig. 2 Left panel: anionic oxygen–water oxygen radial distribution functions of the 2′,3′ cyclic nucleotide model (upper panel) and its 3′,5′ counterpart (lower panel) in water computed at 363 K. Right panel: anionic oxygen–formamide nitrogen radial distribution functions of the 2′,3′ cyclic nucleotide model (upper panel) and its 3′,5′ counterpart (lower panel) in formamide at 363 K.

prebiotic relevance of the temperature range can be found, *e.g.* in our preceding review paper (also see below).[23] Note that such conditions are considered as a relevant context for prebiotic processes on early Earth not only by us[23] but also, *e.g.*, by the so-called "discontinuous synthesis model" of the origin of life advocated by the Benner group;[17,18] for another independent latest research highlight, see ref. 21. This was the reason why we have considered a much broader temperature range (from 363 up to 483 K) when studying the solvation properties of our simplified cyclic nucleotide models in formamide.
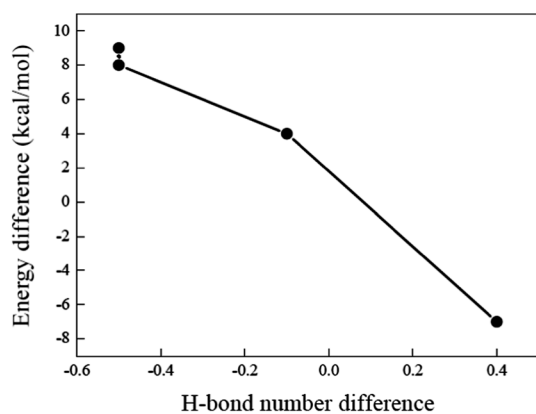

Fig. 3 Correlation between the computed energy differences of the 2′,3′ and 3′,5′ cyclic nucleotide models in formamide and the difference in the number of stabilizing H-bonds in the two cyclic forms.

In the second and third columns of Table 2, we compare the relative stabilities derived for our simplified 2′,3′ and 3′,5′ cyclic nucleotide models from BOMD simulations. The data reveal a striking inversion of the stability order at around 400 K: whereas below this temperature the 2′,3′ cyclic form is more stable than the 3′,5′ one, above this temperature the 2′,3′ cyclic model becomes destabilized with respect to its 3′,5′ counterpart. Again, scrutiny of the H-bonding network formed between the solute and the formamide solvent molecules provides a straightforward explanation for this interesting behavior of the relative stabilities. Whereas at lower temperatures the 2′,3′ cyclic model is able to accommodate more solvent molecules in its solvation shell, above 403 K this trend is clearly reversed in favor of the 3′,5′ cyclic form. The almost monotonic correlation between the differences in the number of stabilizing H-bonds in the 2′,3′ and 3′,5′ cyclic forms and the corresponding relative stabilities is shown in Fig. 3. The curve suggests that solvent–solute interactions are decisive at determining the computed stability orders.

To address the robustness of our H-bond analysis in formamide, parallel evaluations have been carried out by employing slightly different cutoff lengths. As shown in Fig. 2, in formamide the first maximum of the radial distribution function for the distance between the anionic oxygens of the phosphate and the nitrogen of formamide is at 2.8 Å, whereas the first minimum is located in the range of 3.6–3.7 Å. In spite of the higher density of formamide as compared to water, these values are larger than those found in water. This is most likely because the formamide molecules are
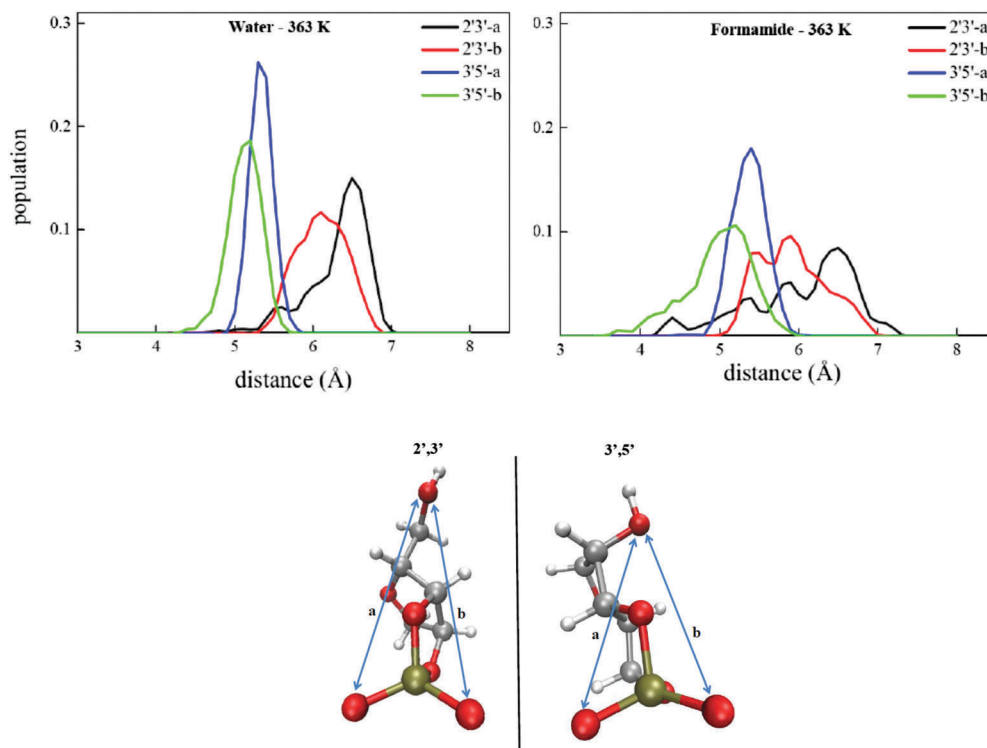

Fig. 4 Probability distribution of the distances between the anionic oxygens of the phosphate and the O5′ of the sugar moiety (labeled as **a** and **b**, for definition see the models on the bottom) adopted during the BOMD simulations of the studied simplified models of 2′,3′ and 3′,5′ cyclic nucleotides in water and formamide environments at 363 K.

somewhat larger than water and thus the steric hindrance between the solute and solvent molecules is more pronounced in this case. Nevertheless, we have found that the results of our H-bonding analysis also hold when defining the H-bonding cutoffs based on the radial distribution function minima rather than the crystallographic definition taken from ref. 45.

Floppiness: the molecular property that enables a better optimization of the solvation shell

In general, our computations suggest that in a given solvent at a given temperature in time average, the more stable cyclic form establishes a greater number of H-bonds with the solvent. Thus, we need to clarify the molecular properties that affect

the number of solvent molecules in the solvation shell of the cyclic nucleotides.

To answer this question we have analyzed a series of intramolecular structural parameters (*i.e.* bond angles, bond lengths, dihedrals, *etc.*). Among them the time-dependent behavior of the distances between the anionic oxygens of the phosphate and the O5' of the sugar moiety provides a putative explanation. These two intramolecular distances, being the longest, characterize large-amplitude motions of the molecule. In Fig. 4 we compare their distributions in water and formamide at 363 K. In formamide a general broadening of the peaks can be observed as compared to water, suggesting that the formamide environment ensures a greater freedom for intramolecular motions. Thus, it seems that in formamide
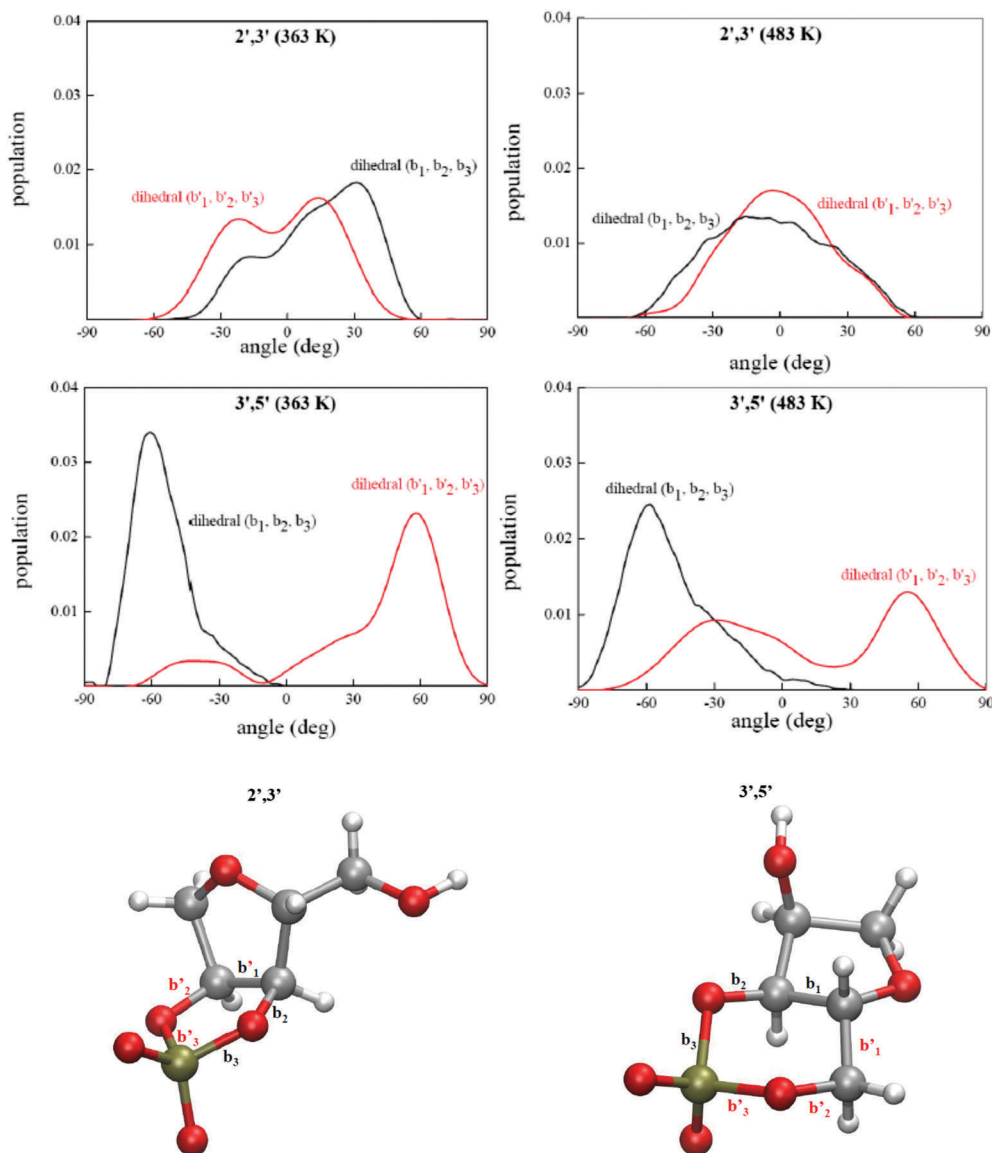

Fig. 5 Probability distribution of the P–O–C_α–C_β dihedrals (for definition see the basis vectors **b**$_1$, **b**$_2$, **b**$_3$ and **b**'$_1$, **b**'$_2$, **b**'$_3$ – *i.e.* the bonds – on the bottom) adopted during the BOMD simulations of simplified 2',3' and 3',5' cyclic nucleotide models at 363 K and 483 K in formamide. Note that for the 5-membered phosphate ring of the 2',3' form, **b**$_1$ and **b**'$_1$ refer to the same bond (*i.e.* it is the C–C bond shared by the sugar ring and the phosphate ring).

the studied cyclic nucleotide molecules are more flexible and are able to optimize their solvation shell more easily than in water. This finding is in line with the larger solvent-inaccessible volume around the cyclic nucleotides in formamide mentioned above.

A subsequent analysis of the two dihedral angles characterizing the rigidity of the 5- and 6-membered rings accommodating the intramolecular phosphodiester linkages sheds light on the differences in the temperature dependence of the dynamic behavior of the studied cyclic nucleotide models in formamide. Fig. 5 helps to envisage how the studied dihedrals have been defined: in both the 2′,3′ and 3′,5′ forms we consider those three bonds which connect the phosphorus atom of the phosphate with the carbon atoms of the sugar part of the simplified model. For the 3′,5′ cyclic model, the distribution of dihedrals is noticeably broadened when increasing the temperature from 363 K to 483 K. In contrast, the distributions computed for the 2′,3′ cyclic form are almost the same at both temperatures. In other words elevated temperatures enhance the flexibility (i.e. the range of thermal fluctuations) of the 6-membered ring of the 3′,5′ cyclic form, whereas they do not much influence that of the 5-membered ring in the 2′,3′ analogue. This might be the reason for the inversion of the relative stabilities of the 2′,3′ and 3′,5′ isomers upon increasing the temperature observed in formamide.

Prebiotic implications

In the above paragraphs, we have used state-of-the art QM simulations to explain the basic physical chemistry reason for variations in the relative stability of 2′,3′ and 3′,5′ cyclic nucleotides depending on the type of solvent and on the temperature. Let us now explain the potential link of this finding to prebiotic processes.

Accumulation of 3′,5′ cyclic nucleotides is relevant in a formamide-based scenario.[20,23] The simple formamide-based origin scenario we outlined in ref. 23 suggests a non-equilibrium solution for the origin of life where different chemical steps take place consecutively under different conditions (see Fig. 6).

In this model we propose that formamide was formed in a Miller–Urey chemistry[48,49] as the thermal dissociation product of ammonium–formate.[50] Note that formic acid is the most abundant compound synthesized in the Miller–Urey experiment, and ammonium–formate spontaneously dissociates to formamide at ca. 453 K,[50] which roughly coincides with the boiling point of the latter (483 K). At this temperature water is still vapor and thus it cannot accumulate in substantial amounts in condensed form. Thus, the accumulation of liquid formamide is a plausible scenario. Let us note that a recent study by Wiegand et al. suggests that formamide could accumulate even at lower temperatures in hydrothermal pores up to concentrations of 85 percent due to thermophoresis and convection.[51]

At around 430 K in liquid formamide, most of the prebiotic building blocks (i.e. nucleobases, nucleosides) can be synthesized,[7,8,52–54] and after phosphorylation (ref. 5), they can accumulate in the form of nucleotides. Note that formamide is considered to be the medium where phosphorylation chemistries could take place on the early Earth also in the "discontinuous synthesis model" elaborated by the Benner group.[17,18] As our computations suggest, at around 430 K in formamide solution, the 3′,5′ cyclic form is the thermodynamically favored one and thus it is reasonable to assume that it out-competes the 2′,3′ cyclic form on an extended time scale. However, upon an environmental stimulus, such as a sudden drop of the temperature (e.g. due to heavy raining), the chemical system is pulled far from equilibrium. Thus, under these conditions the 3′,5′ cyclic nucleotides may become unstable and, therefore, more reactive in oligomerizations. Actually at the typical temperatures of the polymerization reactions (around 360 K),[1,2,12] the absolute
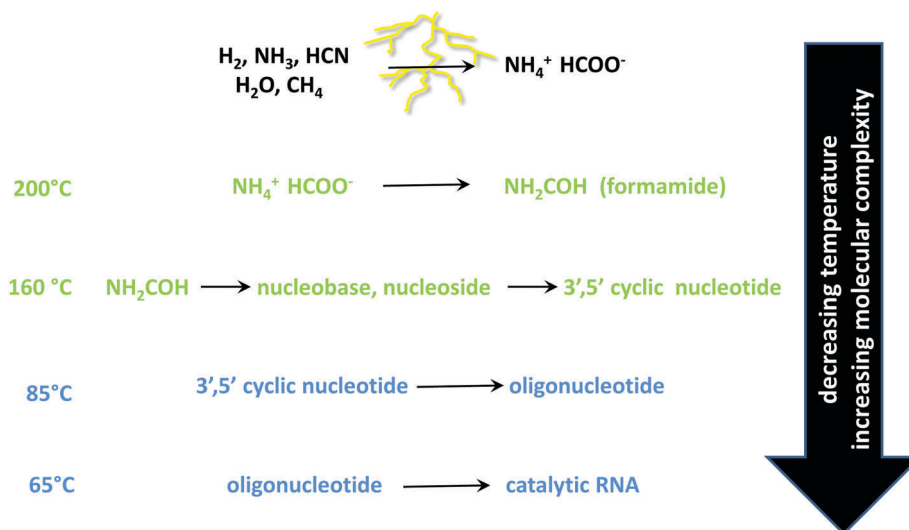

Fig. 6 Temperature-driven chemical evolution.[23] Prebiotic synthesis could occur under appropriate sequences of temperature drops in the primordial environment. These could push through those chemical transformations which could lead to more and more complex molecular entities. In the green part of the pathway (473 to 373 K) relatively slow cooling provided with nearly equilibrium conditions, ideal for the accumulation of prebiotic building blocks. In the blue part, fast cooling to temperatures below 373 K facilitated non-equilibrium chemistries, like oligomerization of cyclic nucleotides precursors, and eventually led to the spontaneous emergence of the simplest catalytically active RNA-molecules.[47]

thermodynamic stability of 3′,5′ cyclic nucleotides is noticeably lower than that of the 2′,3′ cyclic form. This is reflected by the different activity of 3′,5′ and 2′,3′ cyclic nucleotides in polymerization reactions: whereas 3′,5′ cyclic nucleotides readily polymerize spontaneously (see ref. 9, 11 and 12) their 2′,3′ cyclic counterparts require activation by amines to trigger the oligomerization.[1,2] Note that numerous independent recent studies demonstrate that non-equilibrium chemical processes could play an important role in the origin of terrestrial life.[55–57]

Present computations and former experiments[3] show that in water, 2′,3′ cyclic nucleotides are more stable over the whole studied temperature range, *i.e.* from room temperature up to 363 K. This adds to the credibility of aqueous prebiotic scenarios, which consider 2′,3′ cyclic nucleotides as the predecessors of the most ancient oligonucleotides.[19] In other words, if nucleotides were synthesized on the early Earth in a close-to-equilibrium chemistry below the boiling point of water, or if the cooling of the formamide environment was sufficiently slow, accumulation of 2′,3′ nucleotides could overrun that of the 3′,5′ cyclic form.

Conclusions

In the current paper we have employed *ab initio* simulations to address the relative stability of 2′,3′ and 3′,5′ cyclic nucleotides in the presence of solvating water and formamide molecules, respectively. Our computations have shown that in water 2′,3′ cyclic nucleotides are thermodynamically clearly favored over their 3′,5′ cyclic counterpart, practically over the whole temperature range relevant to aqueous chemistry. While the same stability order holds also at lower temperatures in formamide, above 400 K the 3′,5′ cyclic form becomes more stable than its 2′,3′ analogue. This is consistent with prebiotic scenarios in which (i) either 2′,3′ cyclic nucleotides form and accumulate in water and their oligomerization proceeds at nearly the same temperature after activation by amines or amino alcohols, or (ii) 3′,5′ cyclic nucleotides formed above 400 K in formamide spontaneously oligomerize after fast cooling to 358 K (*i.e.* essentially in a non-equilibrium chemistry).

A statistical analysis of the distribution of the geometrical parameters has pointed out two main effects which control the relative stability of the 2′,3′ and 3′,5′ isomers of cyclic nucleotides in formamide and water. Firstly, formamide enables a more intensive intramolecular motion of both solutes as compared to water. In addition, intramolecular motions of 2′,3′ cyclic nucleotides solvated by formamide are not much affected by the temperature, whereas 3′,5′ cyclic nucleotides become markedly more floppy in this environment at higher temperatures. The latter effect might be the reason for the stability inversion of the 2′,3′ and 3′,5′ forms observed at elevated temperatures (>400 K) in formamide.

Our computer experiments have provided an answer for a 40 year-old enigma,[15] which lies in the anomaly between the experimentally measured hydrolysis heats and the ring strain of 2′,3′ and 3′,5′ cyclic nucleotides.[14] It seems that the experimentally measured hydrolysis heats are dictated by a dynamic property of the cyclic nucleotides, *i.e.* their floppiness, rather than by static properties, like ring strain. Let us note that solvation effects have long been deemed to control the experimentally measured trends in the hydrolysis heats of cyclic nucleotides.[16] Our study, in addition to giving credit to this proposal, shows that these solvation effects are strongly dependent on the intrinsic flexibility of the cyclic nucleotides.

From a prebiotic chemistry perspective, our computations have shown that the accumulation of 2′,3′ cyclic nucleotides is more compatible with an aqueous origin scenario, whilst their 3′,5′ cyclic counterparts are clearly favored in a formamide context.[20,23]

References

1 M. S. Verlander, R. Lohrmann and L. E. Orgel, *J. Mol. Evol.*, 1973, **2**, 303.
2 D. A. Usher and D. Yee, *J. Mol. Evol.*, 1979, **13**, 287.
3 S. A. Rudolph, E. M. Johnson and P. Greengard, *J. Biol. Chem.*, 1971, **246**, 1271.
4 M. W. Powner, B. Gerland and J. D. Sutherland, *Nature*, 2009, **459**, 239.
5 G. Costanzo, R. Saladino, C. Crestini, F. Ciciriello and E. Di Mauro, *J. Biol. Chem.*, 2007, **282**, 16729.
6 M. W. Powner, J. D. Sutherland and J. W. Szostak, *J. Am. Chem. Soc.*, 2010, **132**, 16677.
7 R. Saladino, G. Botta, S. Pino, G. Costanzo and E. Di Mauro, *Chem. Soc. Rev.*, 2012, **41**, 5526.
8 R. Saladino, E. Carota, G. Botta, M. Kapralov, G. N. Timoshenko, A. Y. Rozanov, E. Krasavin and E. Di Mauro, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, E2746.
9 G. Costanzo, S. Pino, F. Ciciriello and E. Di Mauro, *J. Biol. Chem.*, 2009, **284**, 33206.
10 M. Morasch, C. B. Mast, J. K. Langer, P. Schilcher and D. Braun, *ChemBioChem*, 2014, **15**, 879.
11 G. Costanzo, R. Saladino, G. Botta, A. Giorgi, A. Scipioni, S. Pino and E. Di Mauro, *ChemBioChem*, 2012, **13**, 999.
12 J. E. Sponer, J. Sponer, A. Giorgi, E. Di Mauro, S. Pino and G. Costanzo, *J. Phys. Chem. B*, 2015, **119**, 2979.
13 T. Dudev and C. Lim, *J. Am. Chem. Soc.*, 1998, **120**, 4450.
14 F. J. Marsh, P. Weiner, J. E. Douglas, P. A. Kollman, G. L. Kenyon and J. A. Gerlt, *J. Am. Chem. Soc.*, 1980, **102**, 1660.
15 J. M. Sturtevan, J. A. Gerlt and F. H. Westheimer, *J. Am. Chem. Soc.*, 1973, **95**, 8168.
16 J. A. Gerlt, N. I. Gutterson, P. Datta, B. Belleau and C. L. Penney, *J. Am. Chem. Soc.*, 1980, **102**, 1655.

1824 | *Phys. Chem. Chem. Phys.*, 2017, **19**, 1817–1825

This journal is © the Owner Societies 2017

17 S. A. Benner, H.-J. Kim and M. A. Carrigan, *Acc. Chem. Res.*, 2012, **45**, 2025.

18 M. Neveu, H.-J. Kim and S. A. Benner, *Astrobiology*, 2013, **13**, 510.

19 J. D. Sutherland, *Angew. Chem., Int. Ed.*, 2016, **55**, 104.

20 J. E. Šponer, J. Šponer and E. Di Mauro, *Wiley Interdiscip. Rev.: RNA*, 2016, DOI: 10.1002/wrna.1400.

21 M. Fiore and P. Strazewski, *Angew. Chem., Int. Ed.*, 2016, **55**, 13930–13933.

22 D. Niether, D. Afanasenkau, J. K. G. Dhont and S. Wiegand, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 4272.

23 J. E. Sponer, J. Sponer, O. Novakova, V. Brabec, O. Sedo, Z. Zdrahal, G. Costanzo, S. Pino, R. Saladino and E. Di Mauro, *Chem. – Eur. J.*, 2016, **22**, 3572.

24 R. Car and M. Parrinello, *Phys. Rev. Lett.*, 1985, **55**, 2471.

25 P. Giannozzi, *et al.*, *J. Phys.: Condens. Matter*, 2009, **21**, 395502.

26 G. Galli and A. Pasquarello, in *Computer Simulation in Chemical Physics*, ed. M. P. Allen and D. J. Tildesley, Kluwer, Dordrecht, 1993, pp. 261–313.

27 W. L. Jorgensen and C. J. Swenson, *J. Am. Chem. Soc.*, 1985, **107**, 569.

28 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.

29 C. Fiolhais, F. Nogueira and M. A. Marques, *Primer in Density Functional Theory*, Springer Verlag, Berlin, 2010.

30 N. Marom, A. Tkatchenko, M. Rossi, V. V. Gobre, O. Hod, M. Scheffler and L. Kronik, *J. Chem. Theory Comput.*, 2011, **7**, 3944.

31 E. Tsuchida, *J. Chem. Phys.*, 2004, **121**, 4740.

32 E. Kálmán, I. Serke, G. Pálinkás, M. D. Zeidler, F. K. Wiesmann, H. Bertagnolli and P. Chieux, *Z. Naturforsch., A: Phys., Phys. Chem., Kosmophys.*, 1983, **38**, 231.

33 R. Ludwig, F. Weinhold and T. C. Farrar, *J. Chem. Phys.*, 1995, **102**, 5118.

34 S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787.

35 G. R. Desiraju, *Acc. Chem. Res.*, 2002, **35**, 565.

36 D. Marx and J. Hutter, *Ab Initio Molecular Dynamics – Basic theory and advanced methods*, Cambridge University Press, Cambridge, 2009.

37 J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi and G. Galli, *J. Chem. Phys.*, 2004, **120**, 300.

38 P. H.-L. Sit and N. Marzari, *J. Chem. Phys.*, 2005, **122**, 204510.

39 S. B. Rempe, T. R. Mattsson and K. Leung, *Phys. Chem. Chem. Phys.*, 2008, **10**, 4685.

40 D. R. Hamann, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1997, **55**, R10157.

41 A. M. Saitta, F. Saija and P. V. Giaquinta, *Phys. Rev. Lett.*, 2012, **108**, 207801.

42 J. VandeVondele, F. Mohamed, M. Krack, J. Hutter, M. Sprik and M. Parrinello, *J. Chem. Phys.*, 2005, **122**, 014515.

43 G. Cassone, P. V. Giaquinta, F. Saija and A. M. Saitta, *J. Chem. Phys.*, 2015, **142**, 054502.

44 A. M. Saitta and F. Saija, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 13768.

45 G. R. Desiraju, J. Vittal and A. Ramanan, *Crystal Engineering – a Textbook*, World Scientific, New Jersey, 2011.

46 J. D. Dunitz, *Chem. Biol.*, 1995, **2**, 709.

47 A plausible geological context for our formamide-model described herein and in ref. 23 is a pool on top of a stable, geologically active region, which undergoes rapid cooling upon heavy rain.

48 S. L. Miller, *Science*, 1953, **117**, 528.

49 S. L. Miller and H. C. Urey, *Science*, 1959, **130**, 245.

50 M. Lorin, *Comptes Rendus*, 1864, **T59**, 788.

51 D. Niether, D. Afanasenkau, J. K. G. Dhont and S. Wiegand, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**(16), 4272.

52 R. Saladino, C. Crestini, F. Ciciriello, G. Costanzo and E. Di Mauro, *Chem. Biodiversity*, 2007, **4**, 694.

53 R. Saladino, C. Crestini, G. Costanzo, R. Negri and E. Di Mauro, *Bioorg. Med. Chem.*, 2001, **9**, 1249.

54 R. Saladino, C. Crestini, F. Ciciriello, S. Pino, G. Costanzo and E. Di Mauro, *Res. Microbiol.*, 2009, **160**, 441.

55 C. B. Mast, S. Schink, U. Gerland and D. Braun, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 8030.

56 T. Göppel, V. V. Palyulina and U. Gerland, *Phys. Chem. Chem. Phys.*, 2016, **18**, 20135.

57 R. Saladino, G. Botta, B. M. Bizzarri, E. Di Mauro and J. M. Garcia Ruiz, *Biochemistry*, 2016, **55**, 2806.

This journal is © the Owner Societies 2017

Phys. Chem. Chem. Phys., 2017, **19**, 1817–1825 | 1825