ARTICLE

OBC
www.rsc.org/obc

The base sequence dependent flexibility of linear single-stranded oligoribonucleotides correlates with the reactivity of the phosphodiester bond

Ulla Kaukinen,*[a,b] Tuomas Venäläinen,[b] Harri Lönnberg[a] and Mikael Peräkylä[b]
[a] *Department of Chemistry, University of Turku, FIN-20014, Turku, Finland*
[b] *Department of Chemistry, University of Kuopio, FIN-70211, Kuopio, Finland*

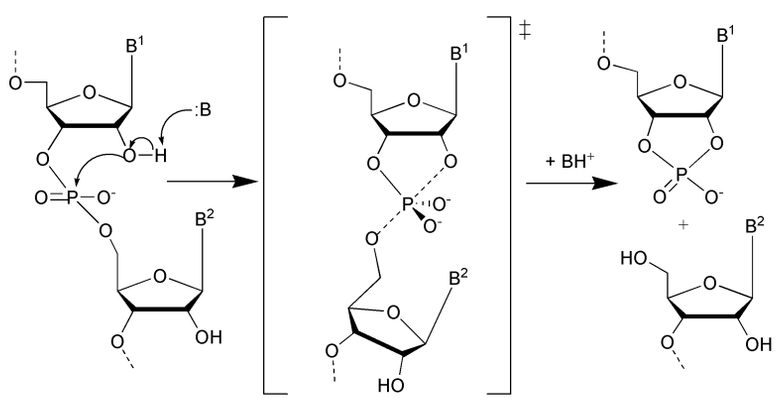
Received 13th March 2003, Accepted 27th May 2003
First published as an Advance Article on the web 9th June 2003

The effect of base sequence on the structure and flexibility of linear single-stranded RNA molecules and the influence of the base sequence on phosphodiester bond reactivity have been studied. Molecular dynamics simulations of 2.1 ns were carried out for nine chimeric oligonucleotides containing only one unsubstituted ribo unit, all the rest of sugars being 2′-*O*-methylated. The base sequence has recently been reported to make a big contribution to the reactivity of these compounds. A detailed examination of the interaction energies between the base moieties shows that base stacking is strongly context-dependent and cooperative. The strength of stacking at the site susceptible to chain cleavage by intramolecular transesterification was observed to be dependent on both the flanking bases of the cleavage site and those further apart in the molecule. The interaction energies between the bases in the vicinity of the scissile linkage were found to correlate well with the experimental phosphodiester bond cleavage rates: the stronger the bases close to the cleavage site are stacked, the slower the cleavage rate is.

Introduction

Since the discovery of ribozymes two decades ago,[1,2] the mechanism of the hydrolysis of RNA phosphodiester bonds (Scheme 1) has been intensively studied.[3–7] Detailed knowledge of the factors that govern the hydrolytic stability of internucleosidic bonds is seen as an inevitable step towards better understanding of the action of catalytic nucleic acids, and hence a step towards rational design of artificial RNases,[8] which, it is hoped, will offer new methods to treat, for example, cancer, viral infections and hereditary diseases. Furthermore, the central role that RNA possibly played during the origin of life[9] has also increased the interest.

Under physiological conditions the hydrolytic stability of a phosphodiester bond is determined by its molecular environment.[4,10,11] Various secondary structure motifs, such as double-stranded helices, hairpin and internal loops as well as bulges, have been observed both to accelerate and retard the inherent cleavage of internucleosidic linkages compared to an isolated phosphodiester bond.[10,12,13] In addition, the base sequence also has a considerable impact on the transesterification rate. In particular, the internucleosidic linkages between 5′-UpA-3′ and 5′-CpA-3′, in the context of natural RNA polymers, have frequently been reported as less stable than those flanked by other nucleosides.[14–18] Efficient and selective cleavage of the 5′-UpA-3′ bond within synthetic oligoribonucleotides in the presence of organic cofactors, has also been documented.[19–21] Recent extensive studies with single-stranded oligoribonucleotides, however, show that the reactivity of an individual phosphodiester bond is strongly dependent on the base sequence of the whole oligomer.[11] Among 20 different sequences that did not exhibit any tendency to form a defined secondary structure, more than 200-fold reactivity differences, in the absence of any cofactors, were observed. Interestingly, not only the neighbouring bases of the scissile phosphodiester bond but also those further apart were found to contribute to the reactivity. Compared to a fully flexible reference compound, both rate-accelerations and -retardations of more than one order of magnitude were reported. The results obtained did not lend support to the view that the 5′-UpA-3′ or 5′-CpA-3′ bonds were inherently more labile than the others, since within dodecameric oligonucleotides the reactivity of both of these bonds has been observed to be, depending on the base sequence, either retarded or enhanced compared to that of the respective isolated phosphodiester bond. Within tetramers, the cleavage rates of the 5′-UpA-3′ or 5′-CpA-3′ bonds have, in turn, been observed to be independent of the nature of the 5′-linked nucleoside. The marked sensitivity of the cleavage of the phosphodiester bonds to the base sequence has been assumed to result from the fact that the oligoribonucleotide chain may adopt a structure that either facilitates or retards the transesterification.[11] This assumption is supported by the fact that the observed differences in the



Scheme 1

DOI: 10.1039/b302751a

cleavage rates at 35 °C were lost on going to elevated temperatures, where the structure is less ordered.[11]

The stacking interactions between nucleic acid bases are the driving forces responsible for the structure of a single-stranded RNA molecule. Stacking preferences are known to be strongly context-dependent, and the magnitude of stacking of two adjacent base moieties within a single-strand is influenced by the flanking bases,[22] and most obviously also by the rest of the compound, since stacking geometries observed in nucleic acids do not coincide with optimal structures of isolated stacked dimers.[23] Solvent interactions also play a crucial but complex role in the stability and structure of nucleic acids; solvent interactions are known to affect the conformational state of nucleic acids[24] but the magnitude of the contribution of solvation to the stacking stability is substantially dependent on the three-dimensional architecture of nucleic acids.[25] In addition, the different base moieties seem to be solvated in a different way.[25–28] Knowledge of the effect of the base sequence on the structure of single-stranded RNA oligomers is, however, scarce.

The stability of phosphodiester bonds within single-stranded RNA molecules has frequently been attributed to strong intrastrand stacking interactions between bases.[11,14,20,29] The rigid structure has been thought to prevent the scissile phosphodiester bond from obtaining the in-line conformation of the 2′-oxygen, the phosphorus and the 5′-oxygen, which is needed for the cleavage reaction.[11,29] However, no quantitative data for the correlation between strong stacking and slow cleavage rates are available. The rate accelerations compared to unconstrained linkage have been more difficult to explain. The favourable in-line conformation in the initial state has been speculated to be necessary for efficient cleavage.[29] In addition, both the weak stacking tendency between the bases across the scissile linkage and the changes in the hydration pattern in the vicinity of the scissile bond have been offered as possible explanations for the enhanced cleavage.[11,19,20] The hydrogen bonding network around the scissile bond has been speculated to accelerate the transesterification by deprotonating the attacking 2′-OH or by donating a proton to the negatively charged phosphodiester and/or the 5′-leaving oxyanion. In particular, a hydrogen bond, either direct or water mediated, between the phosphate and the base of the 5′-linked nucleoside has been suggested to enhance the electrophilicity of the phosphate.[19]

The aim of the present work is to give an insight into the effect of the base sequence on the structure and flexibility of linear single-stranded RNA molecules as well as the relationship between the RNA base sequence and the phosphodiester bond reactivity in order to better understand the structural factors that govern the hydrolytic stability of internucleosidic linkages. Nine linear single-stranded oligoribonucleotides, whose inherent reactivities were recently reported, were selected.[11] The selected compounds experience both rate-accelerations and -retardations compared to a fully flexible reference compound. For these oligoribonucleotides, molecular dynamics (MD) simulations were performed and the interaction energies between adjacent nucleobases along the trajectories obtained were calculated. To our knowledge, no data concerning the influence of different base sequences on the flexibility of single-stranded RNA oligomers are available to date. The impact of the base-sequence on the structure of the oligoribonucleotides and the origin of the significant reactivity differences reported for various sequences are discussed.

Results

The oligonucleotides studied

The compounds studied (**1**–**9**), 12- and 13-mer chimeric ribo/2′-*O*-methylribo oligonucleotides, are depicted in Table 1. All compounds contain only one ribo unit, the rest of the nucleosides being 2′-*O*-methylated. Accordingly, in such compounds only one of the phosphodiester bonds is able to react by the intramolecular transesterification reaction, which has allowed an accurate determination of the cleavage rate of one particular bond in different molecular environments.[11] Chimeric ribo/2′-*O*-methylribo oligoribonucleotides were chosen as model compounds because 2′-*O*-methyl oligoribonucleotides are known to resemble natural RNA reasonably well.[30,31] Evidence that compounds **1**–**9** really exist in a linear single-stranded form has been presented previously.[11]

Table 1 Pseudo first order rate constants for the cleavage of compounds **1**–**11** in CHES buffer at 35 °C[11]

Compound	Sequence[a]	$k/10^{-7}$ s^{-1} [b]
1	5′GGGUA**U**|AAGUGC3′	14.6 ± 0.2
2	5′GGGUA**A**|AAGUGC3′	< 0.2[c]
3	5′GGGUA**C**|AAGUGC3′	0.3 ± 0.1
4	5′GGGUU**U**|AAGUGC3′	31.7 ± 0.6
5	5′GGGUA**U**|AUGUGC3′	43.8 ± 0.2
6	5′GGGUA**U**|AAGUUC3′	30.1 ± 0.1
7	5′GUGUA**U**|AAGUGC3′	1.2 ± 0.1
8	5′CCCCAA**U**|AACCCC3′	< 0.2[c]
9	5′UCUCAA**U**|AACUCU3′	2.5 ± 0.1
10	5′UUUUUU|UUUUUUU3′	0.9 ± 0.1
11	5′UpU3′	2.2 ± 0.1

[a] Bold letters refer to ribonucleosides, the other nucleosides being 2′-*O*-methylated. The position of the strand scission is indicated with a vertical line. [b] In 0.1 M CHES buffer, pH 8.5 (I = 0.1 M with $NaNO_3$). [c] No reaction in three months.

The cleavage of compounds **1**–**9**, in the absence of any co-factors, has recently been reported (Table 1).[11] The reactivities have been compared to those of uridylyl-3′,5′-uridine (UpU, **11**) and the 13-mer uracil homooligomer **10** (Table 1), which are assumed to be models of a fully flexible phosphodiester bond[11] because in these molecules the stabilising stacking interactions between uracil bases are known to be of minor importance.[32–34] Compounds **1** and **4**–**6** exhibit faster cleavage rates than the reference compounds, compounds **7** and **9** are as equally reactive as the reference molecules and the rest of the compounds are clearly cleaved more slowly.

General features of the structures

Molecular dynamics simulations for a duration of 2.1 ns were performed for compounds **1**–**9**. The starting structures were generated as a helical conformation, which is known to be more stable than a random coil conformation.[35] The average structures of compounds **1** and **7** over the last 100 ps are shown as an example in Fig. 1. The stability of the structures during the simulations was examined by measuring the time evolution of the root-mean-squared deviations (RMSd) of the phosphodiester backbone and by calculating the free energies using the molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) method.[36] As clearly seen from Fig. 2, where the time evolutions of the sugar–phosphate backbone RMSd for structures **2** and **4** are shown as an example, the stable average structures are adopted during the last 1.5 ns of the total 2.1 ns simulation time, and the convergence to these average structures is rather good. The RMSd curves are similar for the other compounds. In Table 2 the calculated free energies (G), the standard errors of the mean and the differences in the free energies between 450–1000 ps and 1000–2000 ps (ΔG) for all the compounds studied are shown. Both the standard errors of the mean and ΔG values are very small indicating that the structures are energetically stable during the MD simulation.

Influence of the base sequence on the structure

The conformational stability of single-stranded RNA molecules is mainly determined by the interactions between bases. In order to evaluate how the base sequence influences the structure

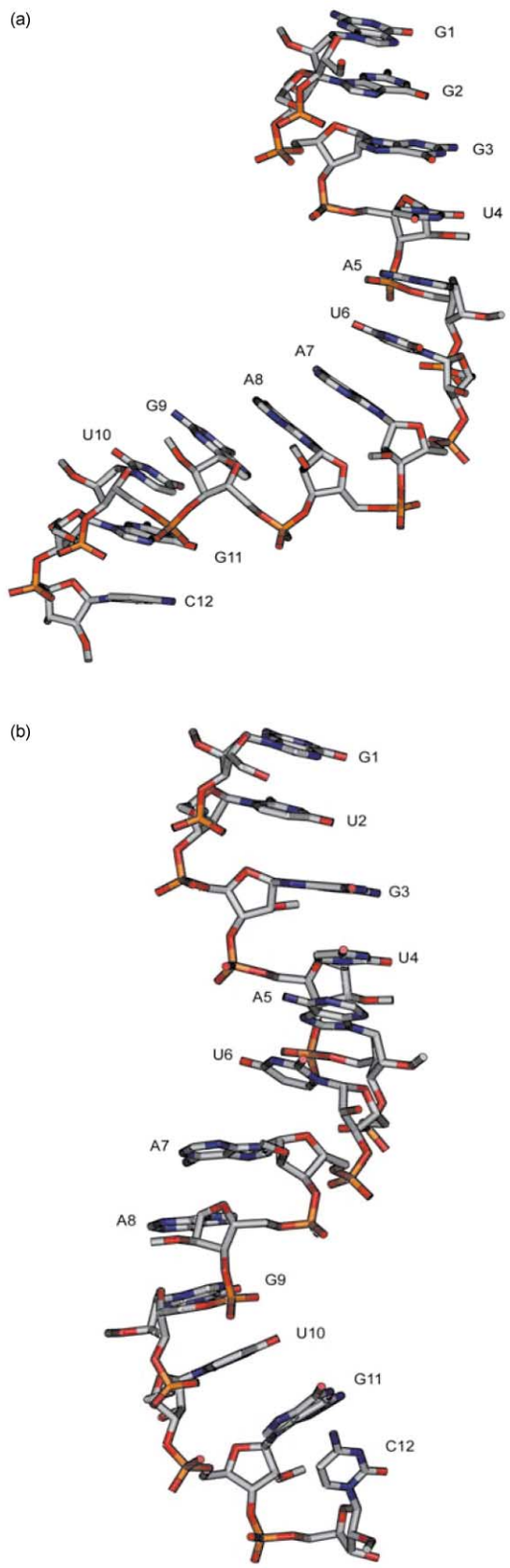


Fig. 1 The average conformations of compound **1** (a) and compound **7** (b) during the last 100 ps of molecular dynamics simulations.

and its stability, molecular mechanical (MM) interaction energies of each base with both the neighbouring bases were calculated along the MD trajectories. MM interaction energies, which consist of van der Waals and electrostatic terms, are known to account for base stacking phenomena.[23] The quality

Table 2 The free energies G, the standard errors of the mean and the differences in the free energies between 1000–2000 ns and 450–1000 ns for each compound studied

Compound	G(mean)/kJ mol^{-1} [a]	$G(\rho)$/kJ mol^{-1} [b]	ΔG/kJ mol^{-1} [c]
1	−10222.4	3.0	−3.7
2	−10291.1	2.9	−2.5
3	−10642.0	2.9	−4.6
4	−10148.6	2.8	2.9
5	−10150.0	2.9	10.8
6	−9923.4	3.2	19.6
7	−9918.1	2.9	−12.9
8	−12477.9	3.3	−2.0
9	−10804.0	2.9	−10.8

[a] Average energies from 416 structures (450–2000 ps). [b] The standard errors of the mean energies. [c] G(450–1000 ps) − G(1000–2000 ps).

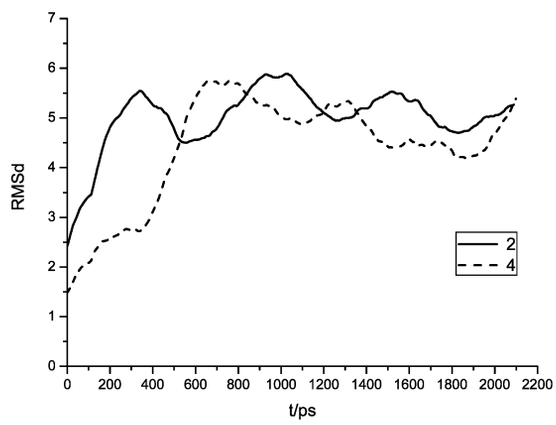


Fig. 2 The time evolution of the root-mean-squared deviations of the phosphodiester backbone for structures **2** and **4**.

of the molecular mechanical interaction energies depends considerably on the quality of the empirical potential used. Therefore, in order to assess the quality of the molecular mechanical interaction energies calculated in this study, the quantum mechanical (QM) interaction energies between the base moieties were determined for one compound (compound **4**). For the six randomly chosen structures obtained from the MD simulations, which resulted in a total of 66 nucleic acid base dimers, *ab initio* single-point calculations at the MP2/6-31G*(0.25) level[23] were performed. Fig. 3 shows the correlation between the MM stacking energies (*i.e.* interaction energies between the bases), calculated using the parm99 parameter set[37] of AMBER 7.0,[38] and the BSSE-corrected *ab initio* QM stacking energies. The correlation coefficient (r) of 0.89 indicates that the empirical potential used provides a correct description of the base–base interactions, consistent with earlier studies.[23,28,39] The force field seems slightly to overestimate the stacking stabilisation, but this falls, as discussed previously,[23,28] within the accuracy limits of the *ab initio* procedure. The largest deviations (~10 kJ mol^{-1}) between the molecular mechanical and the quantum mechanical energies were observed in G–G dimers, which might be due to the neglect of explicit polarization effects.[23,28]

Stacking energies were calculated *in vacuo*, giving the intrinsic stacking of bases within compounds without the influence of solvation. The total MM stacking energies shown in Fig. 4 are the sums of the van der Waals and coulombic energies. The van der Waals energy (Fig. 5) is overlap-dependent and includes the dominating dispersion attractions as well as steric effects, whereas the electrostatic contribution refers to the interactions of molecular electrostatic potentials. The electrostatic contribution of the total gas-phase energy is, however, known to be almost completely compensated by solvation energies.[25,28,40]

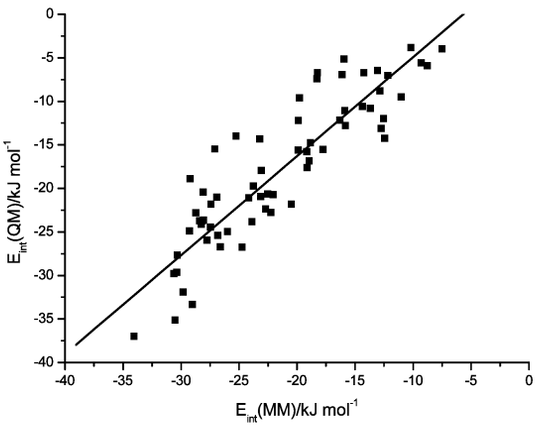


Fig. 3 The correlation between the molecular mechanical (MM) and the BSSE-corrected quantum mechanical (QM) stacking energies. The correlation coefficient (r) is 0.893.

As can be seen from the curves shown in Figs. 4 and 5, the base sequence has a significant effect on the stability of the strand. Oligonucleotides **1**–**3** all have the same base sequence 5′-GGGUAN|AAUGC-3′ with the exception of the 5′-ribonucleoside of the scissile bond, which is either U (**1**), A (**2**) or C (**3**). The interaction energies of the bases around the scissile phosphodiester bonds with their neighbours, *i.e.* the interaction of N6 with A7 and A5, and A7 with N6 and A8, depends on the nature of N6: within **1** the interaction is weaker than within **2** and **3**, with **2** being the most strongly stacked (Fig. 4a). When A5 of **1** is changed to U5 (**4**), keeping the rest of the sequence untouched, base stacking at this position is weakened (Fig. 4b). In addition, this modification (A5⟶U5) has interesting effects on the interaction energies along the rest of the sequence. As can be observed from the curves shown in Fig. 4b, in both sequences U6 has similar interaction energies with N5 and A7, but A7 is stacked significantly more strongly in **4** than in **1**. Compounds **1** and **5** differ due to the replacement of A8 in **1** by U in **5**. This modification does not have such a dramatic effect on the stacking of the rest of the bases, as observed in the case of **4**; only the interaction energies of U8 and G9 in **5** with their neighbours are weaker than those in **1**. Interestingly, the replacement of G2 of **1** with U2 (**7**) also has a long range effect on the interaction energies: U2 and G3 in **7** interact more weakly with their neighbours than those at the same positions in **1**, whereas A5, U6 and A7 in **7** are, in turn, more strongly stacked than within compound **1**. The modification of G11 (**1**) to U11 (**6**) destabilises the 3′-end. This modification slightly increases the stability in the central part of the sequence and destabilises the 5′-end of the sequence compared to **1**, but the effect is smaller than that in **7** (Fig. 4c). Compounds **8** and **9** have the same 5′-AAUAA-3′ sequence in the middle but the terminal sequences are different being CCCC in **8** and UCUC in **9**. The 3′-terminal end of **9** is more stable than that of **8** (Fig. 4d).

Comparison of the total energy curves in Fig. 4 and the intra-strand Lennard-Jones curves in Fig. 5 shows that the coulombic term, in general, makes only a small contribution to the total gas-phase stacking energy. The electrostatic interactions exhibit, however, some interesting features. The electrostatic interactions oppose the stacking of G2 between its neighbours in **1**–**6** by approximately 10 kJ mol^{-1} but have no effect on the interaction propensity of uracil at the same position in **7**. The influence of the electrostatic term on G11 in **1**–**5** and **7** is opposite to that on G2: it favours stacking by approximately 7 kJ mol^{-1}, but uracil at the same position in **6** is destabilised by the coulombic interactions. A similar unfavourable contribution of the electrostatic term is seen at U5 in **4**, whereas its influence on the stacking of adenosines at this position is negligible. Comparison of the total gas-phase interaction energy curves and the van der Waals interaction energy curves of compounds **8** and **9** shows that the electrostatic term has a reverse impact on the 3′-end of the sequences: the coulombic interactions clearly destabilise the terminal 5′-AACCCC-3′-sequence of **8** whereas they slightly stabilise the terminal 5′-UAAUCUC-3′- sequence of **9**.

The geometry of the scissile phosphodiester bond

The structural factors of the scissile phosphodiester bonds, which are assumed [12,13,29] to be important for the cleavage reaction, were examined in more detail. In all cases, the 2′O–P–5′O angle remains between 60–80° over the entire MD simulation. The conformation of the sugar moieties at the 5′-side of the

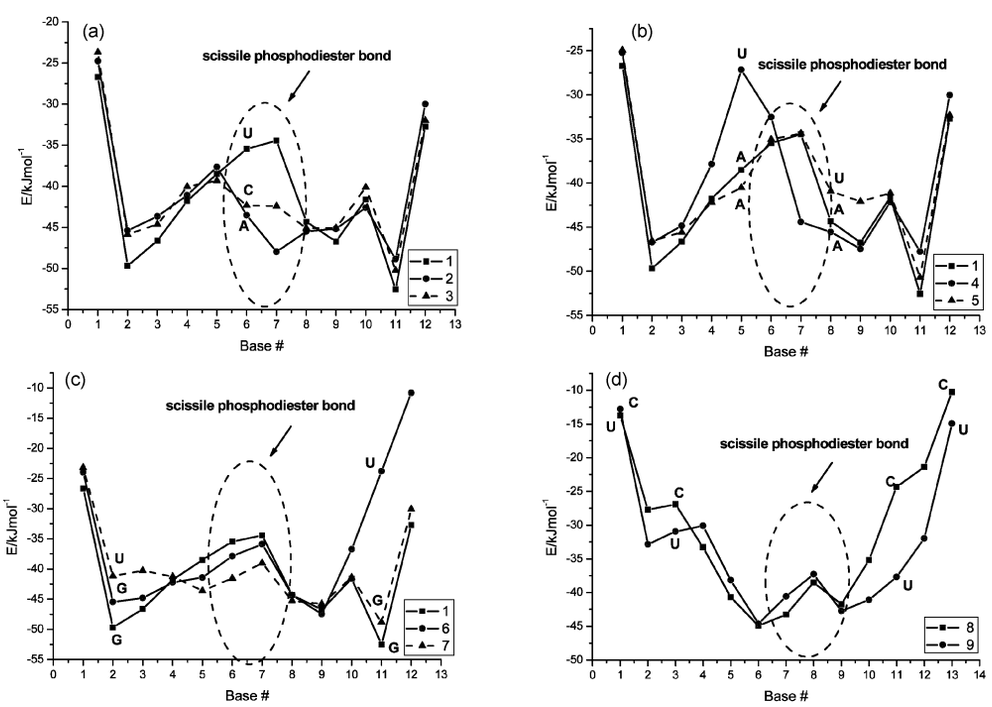


Fig. 4 The total interaction energies (average of 36 structures) of each base with its neighbouring bases. The bases, which are different in the sequences shown in the same picture, are marked.

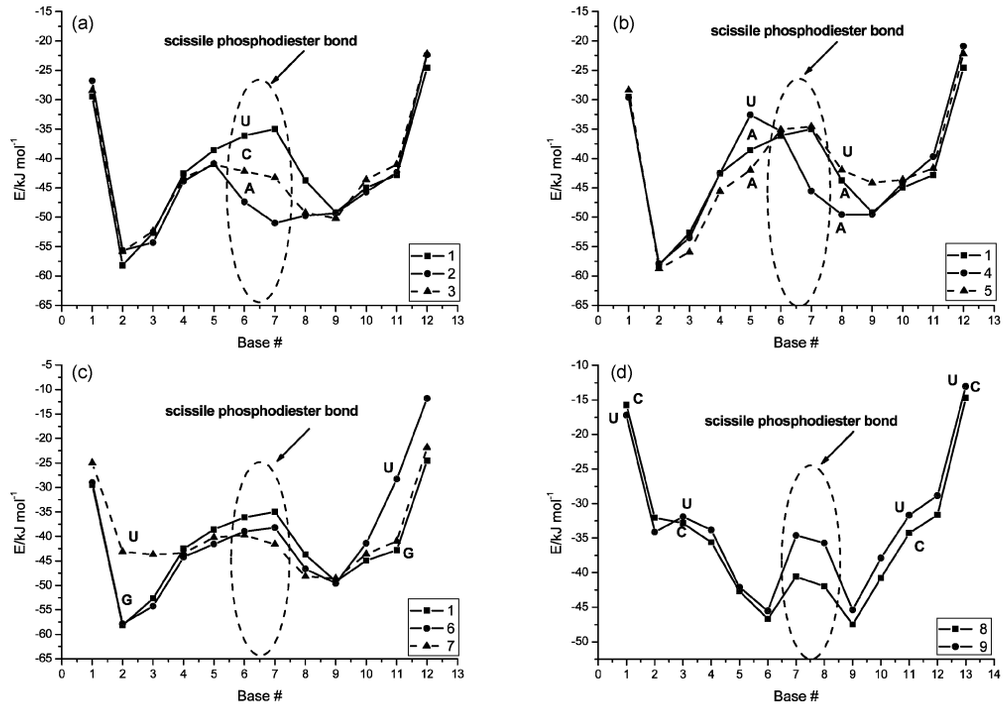


Fig. 5 The van der Waals interaction energies (average of 36 structures) of each base with its neighbouring bases. The bases, which are different in the sequences shown in the same picture, are marked.

scissile linkage, within all the compounds, is C3′-*endo* throughout the simulation, whereas the puckering of the terminal sugar moieties changes between the C2′-*endo* and C3′- *endo* forms. The distance between the nucleophilic 2′-oxygen and the electrophilic phosphorus atoms ranges randomly among the compounds between 3.5 Å and 4.0 Å.

Discussion

The structures of the oligoribonucleotides

The interaction energy of any one particular base with both its neighbours seems to be strongly dependent on the overall base composition so that not only the nature of bases in the nearest neighbourhood but also those further apart have an influence on the interaction energies. Among the compounds studied, the stacking propensity of uracil with its neighbours, which is measured by calculating the interaction energies between bases, is generally weaker than that of any other of the three bases at the same position within otherwise similar sequences. This is consistent with the known negligible interaction energy of uracil with other bases within dimers[34] and poly(U).[32,33] Although the replacement of purine with uracil weakens the stacking at the modified position, this modification might stabilise the strand a few bases away. For example, the interaction energy of A7 with U6 and A8 in **4** is approximately 10 kJ mol^{-1} stronger than in **1** (Fig. 4b). A similar effect is also seen when compound **7** is compared with **1** and to a lesser extent also in the case of the pair of **1** and **6** (Fig. 4c). This phenomenon cannot, however, be generalised as seen with compounds **1** and **5** (Fig. 4b), where the replacement of A8 (**1**) with U8 (**5**) influences the energies only in the vicinity of the modification. Evidently, a long distance effect of one base moiety manifests itself in the cooperative nature of the stacking process. According to the results obtained, it seems obvious that weak base stacking within one part of the oligonucleotide is compensated by strong stacking elsewhere in the molecule. Interestingly, weak stacking in the middle of the chain influences the shape of the oligonucleotide by making the helix curvature more shallow (**1**) compared to a situation where the structure is more compact in the middle of the molecule (**7**) (Fig. 1). As represented earlier with short DNA duplexes,[41,42] the structural transformations caused by the stacking preferences between the bases in the oligomer can easily be mediated through the rather flexible sugar–phosphate backbone.

The nature of the base, the composition of the base sequence around it, as well as the position of the base within an oligonucleotide all seem to influence how the electrostatic and van der Waals terms contribute to the total molecular mechanical interaction energy of one particular base. Generally the net effect of electrostatic interactions is negligible or it opposes stacking, although for some pairs of bases a favourable coulombic interaction in the stacked conformations is found, consistent with an earlier study.[27] At a certain position within the otherwise similar sequences, the coulombic term is different for uracils than for purines. These differences between the total gas-phase energies and the van der Waals energies may result from the differences in the solvation of the bases because the presence of efficient solvent screening is known to reduce the role of the electrostatic term in stacking.[25,28,40] Solvation has earlier been observed to be base-dependent.[25–28,40] For example, smaller purines have been found to more easily become independently solvated than larger pyrimidines in dinucleotide monophosphates.[26] In addition, it has been speculated that the solvation of cytosines is different in dimers and in long polymers.[43,44]

The influence of the base sequence on the reactivity of phosphodiester bonds

As discussed above, the intra-strand interaction energies between bases seems to be strongly sequence dependent. Accordingly, it seems obvious that the interaction energies, especially in the vicinity of the scissile phosphodiester bond, have some impact on the reactivity differences observed for the different sequences. In Fig. 6, the interaction energies between the bases around the scissile linkages are shown as a function of the logarithm of the rate constant for the intramolecular transesterification reaction.[11] In Fig. 6a, the interaction energy between the two flanking bases of the scissile linkage has been calculated, and in Fig. 6b the interactions between four flanking bases of the scissile bond (two bases on both sides) are taken

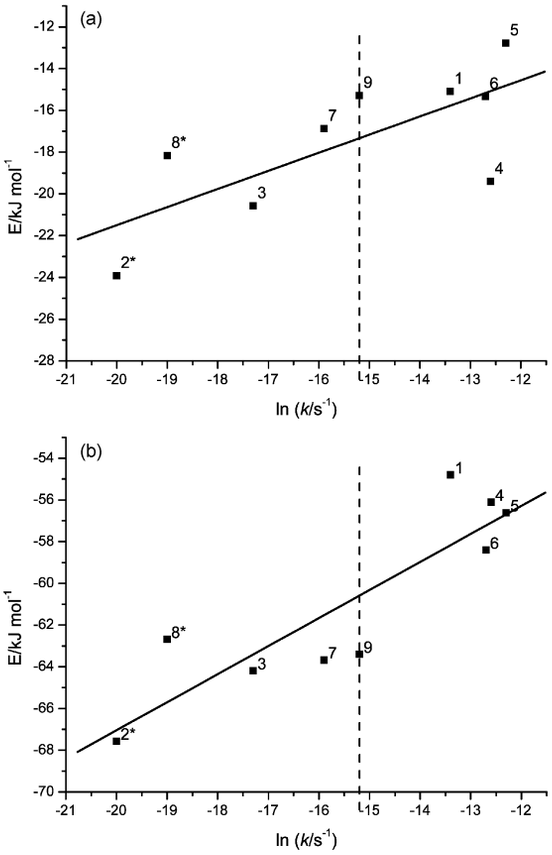


Fig. 6 The total interaction energies between bases around the scissile linkage shown as a function of the natural logarithm of the rate constant.[11] (a) The interaction energy of the 5′-base of the scissile phosphodiester bond with the base at the 3′-side. The correlation coefficient (r) is 0.735. (b) The interaction energy between four flanking bases of the scissile bond. The correlation coefficient (r) is 0877. Compounds **2** and **8** are marked with an asterisk because only the upper bounds of the rate constants are known and the cleavage rate of the model compounds of the fully flexible phosphodiester bond is marked with a dashed line (see Table 1).

into account. In both cases a rather good correlation between the stacking energy around the scissile linkage and the experimental cleavage rate is observed: with two flanking bases the correlation coefficient is 0.735 and with the four flanking bases it is 0.877. These correlations indicate that the stronger the bases close to the scissile phosphodiester bond are stacked, the slower the cleavage rate is. Better correlation observed in Fig. 6b than in 6a shows that not only the interaction energy between the bases across the scissile phosphodiester bond but also the interactions of these bases with the next bases at both 3′- and 5′-sides contribute to the reactivity significantly. It is, however, difficult to determine exactly over how many bases the stabilising effect extends and whether this longer distance stabilisation is fully additive or not. However, within less reactive compounds the stacking interactions between the bases close to cleavage site are stronger than within the more reactive compounds (Fig. 6b). The difference of 7 kJ mol^{-1} observed between these two groups is quite expected compared to *e.g.* the difference of 20 kJ mol^{-1} seen between the most and the least stable base dimers having the optimum stacking geometry.[23] Comparison of the magnitudes of the interaction energies obtained in this study with those published earlier is difficult, because earlier investigations are for isolated nucleobase dimers without the sugar–phosphate backbone. The observation that strong stacking interactions between bases in the initial state retards the cleavage compared to the fully flexible phosphodiester bond is consistent with the mechanism of the intramolecular transesterification reaction (Scheme 1): the reaction is initiated by nucleophilic attack by the 2′-hydroxy group on the adjacent phosphorus, which results in the formation of a pentacoordinated phosphorane species, and the cleavage of the exocyclic P–O bond leads to the RNA strand scission.[3] According to Westheimer's rules,[45] the nucleophile may enter and leave only *via* apical positions, *i.e.* the exocyclic P–O should be positioned in-line with the endocyclic 2′O–P bond. Within helical single-stranded RNA polymers, in-line attack of the 2′-oxygen on the phosphorus is not possible for conformational reasons. A co-linear rearrangement between the 2′-oxygen, the phosphorus, and the 5′-oxygen destroys the stacking interactions between bases. Evidently, the stronger the interactions between bases across the linkage in the initial state are, the more difficult is the rotation of the backbone, which is manifested as a slower reaction rate. In other words, the situation resembles that of double-helical RNA molecules, where the phosphodiester bond cleavage is known to be severely retarded.[10,13,46]

Compounds **1** and **4**–**6** are stacked more strongly in the initial state than in the fully flexible reference compounds having negligible stacking. The reactivity of these compounds is, however, enhanced by one order of magnitude compared to the reference molecules. The rate accelerations might be due to the transition state stabilisation caused by the hydrogen bonding, as suggested earlier.[11] A hydrogen bond network, either direct or water mediated, between the 5′-linked base, especially the adenine base, the phosphorane oxyligands and the attacking and leaving oxygen atoms is believed to facilitate proton transfer from the attacking nucleophile to the leaving group, hence accelerating the cleavage.[3,11,19,47] Enhanced proton transfer to the developing 5′-oxyanion might be especially important. This intramolecular proton transfer from the phosphorane hydroxy ligand to the departing 5′-oxygen has been shown to markedly facilitate cleavage of the exocyclic P–O bond for a pH-independent reaction operating through a monoanionic phosphorane.[48,49] This mechanism is most probably a concurrent reaction with the predominant hydroxide ion catalysed reaction under the experimental conditions used in the cleavage studies.[11] The feasibility of transition state stabilisation by hydrogen bonding cannot be assessed by the approach used in the present study, since the simulations are performed for the initial state structures. The observed cooperative phenomenon of base stacking, where weak stacking within one part of the oligonucleotide is compensated by strong stacking elsewhere in the molecule, offers another explanation for rate accelerations: the complete loss of stacking interactions between bases across the cleavage site in the transition state might enhance stacking within the flanking sequences in heterooligomers **1**–**9** whereas within the reference compounds uracil homooligomer (**10**) and UpU (**11**) transition state stabilisation is not accomplished due to negligible stacking.

The reactivity of phosphodiester bonds within linear single-stranded oligonucleotides is as a whole a complex problem. The experimental rate constants observed are related to the free energy difference between the transition state and the initial state and, accordingly, the stabilisation of both states affects the reactivity observed. The good correlation between the cleavage rate and the stacking of bases in the initial state shown in Fig. 6b, where stacking interactions between bases are distinctly weaker within more reactive compounds than in less reactive compounds, indicates that stacking close to the cleavage site plays an important role in the reactivity. However, there are undoubtedly also other factors, especially transition state stabilisation by hydrogen bonding and/or enhanced stacking as discussed above, which affect the cleavage rate. In addition, as mentioned earlier, it is difficult to know over how many bases the transition state stabilisation extends in the initial state and whether it is fully additive or not.

The co-linear orientation of the attacking 2′-nucleophile, the phosphorus and the 5′-leaving group in the initial state, as well as the short (≤3.25 Å [50]) interatomic distances between the

2′-oxygen nucleophile and the phosphorus electrophile are believed to be important for efficient cleavage.[12] Within secondary structure motifs a rough correlation between this "in-line fitness" and the cleavage rates has been observed[12] and this is also thought to play a role within single-stranded oligonucleotides.[29] For both sets of compounds studied here, those which are cleaved faster and those which are cleaved slower than the reference compounds, the "in-line fitness" is unfavourable. The initial C2′-*endo* conformation of the 3′-linked sugar moiety, which has earlier been speculated to enhance the reaction,[13] does not seem to be necessary for efficient cleavage, since in all cases, independent of the reactivity, the 3′-sugar moiety prefers a C3′-*endo* conformation. Although, this point has not been extensively investigated here, the conclusion that the favourable initial state conformation is not important for the reactivity is consistent with the Curtin–Hammett principle, since as long as the conformational changes take place in a rapid pre-equilibrium state they should not influence the reaction rate, as discussed previously.[11]

Conclusions

The intra-strand energies between bases are significantly dependent on the base sequence. Not only the neighbouring nucleic acid bases, but also those further apart in the molecule, as well as the position of the base within the sequence contribute to the stacking tendency of the bases. Accordingly, base stacking within linear single-stranded oligoribonucleotides is a cooperative process. The interaction energies across the scissile phosphodiester bond in the initial state structures were found to correlate with the intramolecular transesterification rate constants previously reported. The stacking across the scissile linkage within compounds exhibiting slower cleavage rates than the fully flexible reference compound is stronger than within compounds which are cleaved faster than the reference one. Evidently, strong stacking retards the transesterification by preventing the 2′-oxygen, the phosphorus and the 5′-oxygen from adopting the co-linear orientation required for the reaction to take place. The initial conformation of the phosphodiester bond does not seem to be important for the cleavage reaction.

Experimental

Starting structures

The starting structures were created using the BIOPOLYMER module of the SYBYL program.[51] Compounds were initially built up as A-form double-helices and the complementary strand was then deleted giving the single-stranded oligoribonucleotide with the desired base sequence. The 2′-*O*-methylribonucleoside units were generated using the LEAP module[52] of AMBER 7.0[38] from the corresponding ribonucleoside units. The charges and atom types of the 2′-*O*-methylribosyl moieties, which differ from those of unmodified residues, were taken from the work of Kaukinen *et al.*[10] (Fig. 7). The force field parameters for the corresponding groups were taken from the parm99 parameter set[37] of AMBER 7.0. The

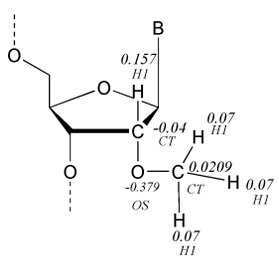


Fig. 7 The charges and the atom types of the 2′-*O*-methylribosyl moiety. Only those charges and atom types which differ from those in the corresponding ribonucleoside unit are shown.

chimeric ribo/2′-*O*-methylribo oligonucleotides were solvated with a box of TIP3P waters extending 13.0 Å in all dimensions around the solute using the LEAP module. This led to box sizes of ~51 × ~53 × ~69 Å. The number of water molecules ranged from 4519 to 4938 depending on the size of the oligonucleotide. The system was then neutralized by adding the appropriate number of sodium counter ions.

Molecular dynamics simulations

Energy minimisations and molecular dynamics simulations were performed using the Sander module of AMBER 7.0 with the parm99 parameter set.[37] The equilibration was performed using the following procedure. In the first step, the hydrogen atoms in the system were minimised using the steepest descent algorithm (1000 steps) keeping the rest of the system fixed followed by a similar relaxation of the water molecules and a short (30 ps) MD run for the water molecules. In the MD run the system was heated up from the initial temperature of 100 K to 300 K in 3 ps, and thereafter the temperature was maintained at 300 K. After the relaxation of the hydrogens and the water molecules, the whole system was energy minimised for 1000 steps. Finally, the unrestrained molecular dynamics simulation of the whole system was started by heating up the whole system as done in the equilibration MD simulation of the water molecules. After that production simulations of 2.1 ns at constant temperature (300 K) and pressure (1 atm) were started. All simulations were run using a 1.5 fs step time. A cutoff of 8.0 Å was used for van der Waals interactions and the long-range electrostatic interactions were treated using the particle mesh Ewald summation method[53–56] with a charge grid spacing of ~1.0 Å. The bonds containing hydrogen were constrained using the SHAKE algorithm.[57]

Trajectory analyses

The resulting MD trajectories were analysed using the ptraj, Carnal, ANAL and MM-BPSA modules of AMBER 7.0.[38] The calculations of the root-mean-squared deviations (RMSd) and geometrical parameters were performed using the ptraj module. Sugar pucker pseudorotation values were determined based on the Altona and Sundaralingam conventions.[58]

In order to define the total interaction energies between bases, only the coordinates of the base moieties were recorded from the trajectories analysed using the Carnal module. For each compound, 36 snapshot structures were collected at even intervals (45 ps) between 450–2025 ps simulation time. To the base residues of each snapshot hydrogen was added at the position of the *N*-glycosylic bond using the LEAP module. The charge of this hydrogen was adjusted to produce a residue with a net zero charge. The energy decomposition between the base moieties within the same compound was calculated with ANAL using the same force field parameters as in the MD simulation. The interaction energies between the bases reported in Figs. 4 and 5 are averages of 36 structures. The standard errors of the mean of the calculated interaction energies are on average 0.85 kJ mol^{-1}.

The absolute free energies of the oligonucleotides studied were calculated using the molecular mechanics Poisson–Boltzman surface area (MM-PBSA) method.[36] The free energy (G) was calculated here as the sum of the molecular mechanical energy and solvation free energy ($G = E_{MM} + \Delta G_{solv}$) for snapshot structures taken from the MD trajectory. The free energies reported are averages of 416 structures. The molecular mechanical energy was determined using the Sander module of AMBER 7.0. The electrostatic contribution (ΔG_{PB}) of the solvation free energy was computed by the finite difference Poisson–Boltzman method[59] as implemented in the Delphi II program.[60] This was approximated as the reaction field energy of taking a solute from a vacuum dielectric ($\varepsilon = 1$) to an aqueous dielectric ($\varepsilon = 80$). A probe radius of 1.4 Å and atomic radii

adapted from the PARSE parameter set[61] were used to define the molecular surface. Atomic charges were taken from the Amber parm99 parameter set[37] in order to have a consistent set of charges for calculating the total electrostatic energies. For each computed structure, an 80% boxfill cubic lattice with a grid resolution of 0.5 Å per grid point was applied in the Delphi calculations. The non-polar contribution to the solvation free energy was estimated according to $\Delta G_{\text{non-polar}} = \gamma \times \text{SASA} + \beta$, where $\gamma = 0.00542$ kcal Å^{-1} and $\beta = 0.92$ kcal mol^{-1}[62] and the solvent accessible surface area (SASA) was calculated with MSMS software.[63]

Ab initio calculations

Six snapshot structures of compound **4**, containing only the coordinates of the base moieties (see above), were randomly chosen for quantum mechanical calculations. For each of the resulting 66 base dimers, having the geometry obtained from the MD simulation, *ab initio* quantum mechanical single-point energy calculations were performed. The stacking energies of the complexes were calculated at the second order Møller–Plesset (MP2) level. A standard 6-31G basis set was used with a set of d-polarisation functions having an exponent of 0.25 instead of a standard 0.8 (abbreviated: 6-31G*(0.25)).[23] The use of diffuse polarisation functions instead of the standard ones are required in order to include a sufficient amount of intermolecular electron correlation effects.[64,65] The stacking energy was corrected for the basis set superposition error (BSSE) using the function counterpoise method.[66] The stacking energy of a dimer was obtained by subtracting the sum of the BSSE-corrected energies of the monomers (*i.e.* calculated within the dimers where one of the monomers was set as a "ghost" molecule) from that of the dimer. The BBSE-corrected MP2/6-31G*(0.25) energies have been proved, based on high-level *ab initio* quantum mechanical calculations, to produce accurate stacking energies.[23,67] The *ab initio* calculations were carried out using the GAUSSIAN98 program.[68]

Acknowledgements

The Centre for Scientific Computing (CSC, Espoo, Finland) is gratefully acknowledged for the access to its computational resources. This work was supported by the Academy of Finland (grants 68528 and 74097 to M.P.).

References

1 T. R. Cech, A. J. Zaug and P. J. Grabowski, *Cell*, 1981, **29**, 487–496.
2 C. Guerrier-Takada, K. Gardiner, T. Marsh, N. Pace and S. Altman, *Cell*, 1983, **35**, 849–857.
3 M. Oivanen, S. Kuusela and H. Lönnberg, *Chem. Rev.*, 1998, **98**, 961–990.
4 S. Mikkola, U. Kaukinen and H. Lönnberg, *Cell Biochem. Biophys.*, 2001, **34**, 95–119.
5 S. Mikkola, M. Kosonen and H. Lönnberg, *Curr. Org. Chem.*, 2002, **6**, 523–538.
6 P. Perrault and E. V. Anslyn, *Angew. Chem., Int. Ed. Engl.*, 1997, **36**, 432–450.
7 D. M. Zhou and K. Taira, *Chem. Rev.*, 1998, **98**, 991–1026.
8 B. N. Trawick, A. Daniher and J. K. Bashkin, *Chem. Rev.*, 1998, **98**, 939–960.
9 G. F. Joyce, *Nature*, 2002, **418**, 214–221.
10 U. Kaukinen, L. Bielecki, S. Mikkola, R. W. Adamiak and H. Lönnberg, *J. Chem. Soc., Perkin Trans. 2.*, 2001, 1024–1031.
11 U. Kaukinen, S. Lyytikäinen, S. Mikkola and H. Lönnberg, *Nucleic Acids Res.*, 2002, **30**, 468–474.
12 G. A. Soukup and R. R. Breaker, *RNA*, 1999, **5**, 1308–1325.
13 I. Zágorowska, S. Mikkola and H. Lönnberg, *Helv. Chim. Acta*, 1999, **82**, 2105–2111.
14 A. Bibillo, M. Figlerowicz and R. Kierzik, *Nucleic Acids Res.*, 1999, **27**, 3931–3937.
15 V. J. Cannistraro, M. N. Subbarao and D. Kennel, *J. Mol. Biol.*, 1986, **192**, 257–274.
16 A. C. Dock-Bergeon and D. Moras, in *Cold Spring Harbor Symposia on Quantitative Biology*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1987, Vol. LII, pp. 113–121.
17 H. Hosaka, I. Sakabe, K. Sakamato, S. Yokoama and H. Takaku, *J. Mol. Chem.*, 1994, **269**, 20090–20094.
18 N. S. Zhdan, I. L. Kuznetsova, A. V. Vlassov, V. N. Silnikov, M. A. Zenkova and V. V. Vlassov, *Bioorg. Khim.*, 1999, **25**, 723–732.
19 R. Kierzek, *Nucleic Acids Res.*, 1992, **20**, 5073–5077.
20 R. Kierzek, *Nucleic Acids Res.*, 1992, **20**, 5079–5084.
21 A. Bibillo, M. Figlerowicz, K. Ziomek and R. Kierzek, *Nucleosides, Nucleotides & Nucleic Acids*, 2000, **19**, 977–994.
22 J. Nordberg and L. Nilsson, *Biopolymers*, 1996, **39**, 765–768.
23 P. Hobza and J. Šponer, *Chem. Rev.*, 1999, **99**, 3247–3276.
24 M. J. Lowe and J. A. Schellman, *J. Mol. Biol.*, 1972, **65**, 91–109.
25 F. Lankaš, T. E. Cheatham III, N. Špackova, P. Hobza, J. Langowski and J. Šponer, *Biophys. J.*, 2002, **82**, 2592–2609.
26 J. Nordberg and L. Nilsson, *Biophys. J.*, 1998, **74**, 394–402.
27 R. A. Friedman and B. Honig, *Biophys. J.*, 1995, **69**, 1528–1535.
28 J. Šponer, J. Florián, H.-L. Ng, J. E. Šponer and N. Špacková, *Nucleic Acids Res.*, 2000, **28**, 4893–4902.
29 Y. Li and R. R. Breaker, *J. Am. Chem. Soc.*, 1999, **121**, 5364–5372.
30 T. M. Popenda, E. Biala, J. Milecki and R. W. Adamiak, *Nucleic Acids Res.*, 1997, **25**, 4589–4598.
31 P. Auffinger and E. Westhof, *Angew. Chem., Int. Ed.*, 2001, **40**, 4648–4650.
32 E. G. Richards, C. P. Flessel and J. R. Fresco, *Biopolymers*, 1963, **1**, 431–446.
33 L. D. Inners and G. Felsenfeld, *J. Mol. Biol.*, 1970, **87**, 817–833.
34 J. Nordberg and L. Nilsson, *J. Am. Chem. Soc.*, 1995, **117**, 10832–10840.
35 S. Sen and L. Nilsson, *J. Am. Chem. Soc.*, 2001, **123**, 7414–7422.
36 P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. Cheatham III, *Acc. Chem. Res.*, 2000, **33**, 889–897.
37 J. Wang, P. Cieplak and P. A. Kollman, *J. Comput. Chem.*, 2000, **21**, 1049–1074.
38 D. A. Case, D. A. Pearlman, J. W. Caldwell, T. E. Cheatham III, J. Wang, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, V. Tsui, H. Gohlke, R. J. Radmer, Y. Duan, J. Pitera, I. Massova, G. L. Seibel, U. C. Singh, P. K. Weiner, P. A. Kollman, AMBER 7.0, University of California, San Francisco, USA, 2002.
39 F. L. Gervasio, R. Chelli, P. Procacci and V. Schettino, *Proteins: Struct., Funct., Genet.*, 2002, **48**, 117–125.
40 J. Florián, J. Šponer and A. Warshel, *J. Phys. Chem. B*, 1999, **103**, 884–892.
41 C. Hunter, *J. Mol. Biol.*, 1993, **230**, 1025–1054.
42 M. J. Packer and A. C. Hunter, *J. Mol. Biol.*, 1998, **280**, 407–420.
43 D. Pörschke, *Biopolymers*, 1978, **17**, 315–323.
44 S. M. Freier, K. O. Hill, T. G. Dewey, L. A. Marky, K. J. Breslauer and D. H. Turner, *Biochemistry*, 1981, **20**, 6198–6206.
45 F. H. Westheimer, *Acc. Chem. Res.*, 1968, **1**, 70–78.
46 D. A. Usher and A. H. McHale, *Proc. Natl. Acad. Sci. U. S. A.*, 1976, **73**, 1149–1153.
47 Y. Takagi, M. Warashina, W. J. Stec, K. Yoshinari and K. Taira, *Nucleic Acids Res.*, 2001, **29**, 11815–1834.
48 M. Kosonen, K. Hakala and H. Lönnberg, *J. Chem. Soc., Perkin Trans, 2.*, 1998, 663–670.
49 M. Kosonen, R. Seppänen, O. Wichmann and H. Lönnberg, *J. Chem. Soc., Perkin Trans. 2*, 1999, 2433–2439.
50 R. A. Torres and T. C. Bruice, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 11077–11082.
51 SYBYL, Tripos Associates, St. Louis, Missouri, USA, 1998.
52 C. E. A. F. Schafmeister, W. S. RossV. Romanovski, LEAP, University of California, San Francisco, 1995.
53 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
54 U. Essmann, L. Perera, M. L Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577–8593.
55 C. Sagui and T. A. Darden, in *Simulation and Theory of Electrostatic Interactions in Solution*, ed. L. R. Pratt and G. Hummer, American Institute of Physics, Melville, New York, 1999, pp. 104–113.
56 A. Toukmaji, C. Sagui, J. Board and T. Darden, *J. Chem. Phys.*, 2000, **113**, 10913–10927.
57 J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.
58 C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, 1972, **94**, 8205–8212.
59 K. A. Sharp and B. Honig, *J. Phys. Chem.*, 1990, **94**, 7684–7692.

60 A. Nicholls, K. A. Sharp and B. Honig, *Delphi*, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, USA, 1990.
61 I. Massova and P. A. Kollman, *Perspect. Drug Discovery Des.*, 2000, **18**, 1–23.
62 T. E. Cheatham, III, J. Srinivasa, D. A. Case and P. A. Kollman, *J. Biomol. Struct. Dynam.*, 1998, **16**, 265–280.
63 M. F. Sanner, A. J. Olson and J. C. Spehner, *Biopolymers*, 1996, **38**, 305–320.
64 J. Šponer, H. A. Gabb, J. Leszczynski and P. Hobza., *Biophys. J.*, 1997, **73**, 76–87.
65 J. Šponer, J. Leszczynski and P. Hobza, *J. Phys. Chem.*, 1996, **100**, 5590–5596.
66 S. F. Boys and F. Bernardi, *Mol. Phys.*, 1970, **19**, 553–556.
67 P. Hobza and J. Šponer, *J. Am. Chem. Soc.*, 2002, **124**, 11802–11808.
68 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle and J. A. Pople, Gaussian 98, Revision A.1., 1998, Gaussian, Inc., Pittsburgh PA, USA.