Communication

Molecular Dynamics Simulation Reveals Conformational Switching of Water-Mediated Uracil-Cytosine Base Pairs in an RNA Duplex

Christoph Schneider, Maria Brandl, Jürgen Sühnel*

Biocomputing, Institut für Molekulare Biotechnologie, Postfach 100813, D-07708 Jena / Germany

Corresponding author:

Jürgen Sühnel, Biocomputing, Institut für Molekulare Biotechnologie, Postfach 100813, D-07708 Jena / Germany.

phone: +49-3641-656200, fax: +49-3641-656210, e-mail: jsuehnel@imb-jena.de.

Running title:

Conformational switching of water-mediated base pairs

1

A 4 ns molecular dynamics simulation of an RNA duplex (r-GGACUUCGGUCC)$_2$ in solution with Na$^+$ and Cl$^-$ as counterions was performed. The X-ray structure of this duplex includes two water-mediated uracil-cytosine pairs. In contrast to the other base pairs in the duplex the water-mediated pairs switch between different conformations. One conformation corresponds to the geometry of the water-mediated UC pairs in the duplex X-ray structure with water acting both as hydrogen bond donor and acceptor. Another conformation is close to that of a water-mediated UC base pair found in the X-ray structure of the 23S rRNA sarcin/ricin domain. In this case the oxygen of the water molecule is linked to two base donor sites. For a very short time also a direct UC base pair and a further conformation that is similar to the one found in the RNA duplex structure but exhibits an increased H3(U)...N3(C) distance is observed. Water molecules with unusually long residence times are involved in the water-mediated conformations. These results indicate that the dynamic behaviour of the water-mediated UC base pairs differs from that of the duplex Watson-Crick and non-canonical guanine-uracil pairs with two or three direct hydrogen bonds. The conformational variability and increased flexibility has to be taken into account when considering these base pairs as RNA building blocks and as recognition motifs.

Water-mediated base pairs represent a new base pair motif that has recently been identified in RNA structures (Holbrook *et al*., 1991; Rould *et al*., 1991; Cruse *et al*., 1994; Arnez & Steitz, 1996; Correll *et al*., 1997; Rath *et al*., 1998; Correll *et al*., 1999; Tanaka *et al*., 1999). In these complexes the bases are connected by direct hydrogen bonds (H-bonds) and an additional water-mediated base-base H-bond interaction. One of the first examples was found by Holbrook *et al*. (1991). In the RNA duplex (r-GGACUUCGGUCC)$_2$ the central GU and UC mismatches do not form an internal loop, but rather a highly regular helix. In the UC base pair H42 of cytosine is H-bonded to O4 of uracil and H3 of uracil and N3 of cytosine are linked by H-bonds to a tightly associated water molecule (temperature factor: 11.7 Å^2). The secondary structure of the RNA duplex and a schematic drawing of the water-mediated UC base complex is shown in Figure 1. Due to the water molecule between the bases the C1'(U)...C1'(C) distance is even slightly larger than the corresponding distances of Watson-Crick pairs in an ideal helix. Hence, even though the UC base pair is of the pyrimidine-pyrimidine type, its incorporation into a helix leads only to minor distortions of the helix geometry.

Thus far, it is not clear, whether the geometry of these unusual base complexes is due to their intrinsic properties or enforced by backbone or stacking restraints exerted by the nucleic acid environment. In addition, it has to be analysed whether or not crystal forces are important. Finally, the dynamic behaviour of water-mediated pairs remains to be assessed.

We have recently shown that the optimised geometry of the water-mediated UC complex obtained from quantum-chemical *ab initio* calculations closely resembles the experimental geometry observed in the nucleic acid environment (Brandl *et al*., 1999; Brandl *et al.*, in press). Further, the total interaction energy of the water-mediated UC pair is substantially more negative than the sum of the three pairwise contributions. Thus, the interaction energy shows a high degree of cooperativity. From these facts we have concluded

that the base pair geometry is primarily governed by the base constituents alone and is almost not affected by stacking and backbone effects. Hence, this base pair can be regarded as a structurally autonomous building block of RNA.

The quantum-chemical results indicate that neither the crystal nor the nucleic acid environment are required for the formation of this base pair. They cannot prove, however, if the base pair is also stable in solution. For another base pair there is, however, an argument in favor of the occurrence of water-mediated pairs in solution. The loop E 5S rRNA structure has been resolved both by X-ray crystallography (Protein Data Bank (PDB) entry: 354d) and NMR spectroscopy (PDB entry: 1a4d) (Correll *et al*., 1997; Dallas & Moore, 1997). The loop E structure includes three water-mediated base pairs GG, GA and UG. It turns out that in the GA pair the distance between the base donor and acceptor sites involved in the water link (N1-H(G)...N1(A)) is with 3.88 and 3.92 Å almost the same in the NMR and X-ray structures even though in the NMR refinement water was not directly taken into account (in the X-ray structure the hydrogen atom positions were added with InsightII from Molecular Simulations, Inc). This clearly shows that at least this base pair occurs in solution with a geometry similar to the X-ray structure. This is not true for the other two water-mediated base pairs. In the GG case the O6...N7 distance is 1.34 Å longer in the X-ray structure, whereas for the UG complex the X-ray N3-H...H-N2 distance is shorter by 1.0 Å. In these cases the distances between the base donor/acceptor sites differ by approximately 1 Å between the X-ray and NMR structures. It should be noted, however, that in the NMR structure calculations various restraints were inferred from the spectra rather than measured directly (Dallas & Moore, 1997). A detailed comparison between the X-ray and NMR structures has shown that the water-mediated base pairs exhibit the same pairing pattern (Rife *et al.,* 1999). However, a quantitative comparison of single base-base distances was not done by these authors.

Molecular dynamics (MD) simulations can be performed for complete medium-sized biopolymers in solution. They yield information on both the average structure and its dynamical behaviour. In particular these simulations enable one to analyse short timescales that are usually not studied by standard NMR spectroscopy. Note, however, that more recent techniques extend the timescale accessible by NMR spectroscopy into the picosecond range (Kay, 1998; Phan et al., 1999). This opens up the possibility to compare directly simulation results with NMR data for movements occurring at this timescale. Finally, given the simulations are performed over sufficient long simulation times it becomes also possible to detect conformational changes. We have performed a 4 ns MD simulation of the RNA duplex (r-GGACUUCGGUCC)$_2$ in order to study the structure and dynamics of water-mediated UC base pairs.

Results and Discussion

We have first analysed the time course of the root-mean-square deviation (RMSd) of the simulated structure as compared to the starting geometry (results not shown). The RMSd value oscillates around a stable mean value of 2 Å. This indicates that the overall geometry of the complex remains stable during the simulation.

In order to study the structure and dynamics of base pairs we have analysed the time course of various inter-base distances. In Figure 2 results for the two water-mediated UC pairs (U6-C19, C7-U18) and for the flanking GU base pairs (U5-G20, G8-U17) are shown. For the GU pairs the H1(G)...O2(U) and the H3(U)...O6(G) distances of the direct inter-base H-bonds are displayed. For the water-mediated UC pairs the inter-base distances H42(C)...O4(U) and H3(U)...C(N3) are monitored. The GU pairs exhibit the stable behaviour typical for base pairs

with two or three direct H-bonds. The H-bond distances fluctuate around a value of about 2 Å which is in agreement with values found in X-ray and NMR structures. Occasionally, very short base pair opening events are found. The only longer opening event occurs for the G8-U17 pair in the time range between 1000 and 1200 ps.

By contrast, the water-mediated UC pairs show a completely different dynamics. They adopt different conformations that are stable over simulation times between 500 and 1500 ps and switch between these conformations from time to time. In the time range between 800 and 4000 ps altogether four conformations can be identified. This analysis is focussed on base pair geometries. Therefore, we adopt the term conformation for different relative orientations of the bases within a base pair. Of course, the bases are linked to the sugar-phosphate backbone and therefore different base pair conformations may be accompanied by different backbone torsion angles as well.

Conformation I is characterised by a direct H-bond between one hydrogen atom of the cytosine amino group and the O4 atom of uracil with an average distance of about 2 Å. The distance between the water donor and acceptor sites H3(U) and N3(C) is around 4 Å. In conformation II both of these distances are substantially increased. In addition to the dominating conformations I and II, we find a UC pair with two direct H-bonds (III) for a short time of 400 ps and a further conformation that is similar to I but exhibits an increased H3(U)...N3(C) distance (IV) (Figure 2).

In Figure 3a average structures of conformations I and II are compared to the experimental RNA duplex structure. The superposition was done for the P, O1P, O2P, O3', and O5' backbone atoms of the flanking GU base pairs. It can be seen that conformation I corresponds to the water-mediated UC pair found in the RNA duplex X-ray structure. On the other hand, in conformation II the relative orientation of U and C is completely different to I. The direct H-bond between the exocyclic amino group of cytosine and O4 of uracil is broken

6

and the shortest inter-base distances occur between the cytosine amino group and O2 of uracil. It is also obvious that uracil shows a much larger movement than cytosine in passing from I to II. We have performed a systematic analysis of backbone torsional angles in both conformations. It turned out that indeed for cytosine there are no major differences between I and II. On the other hand, for uracil significant changes occur for ε and ζ (U18: ε_I = - 153°, ε_{II} = - 170°, ζ_I = -66°, ζ_{II} = -82°).

Interestingly, conformation II bears resemblance to the geometry of a water-mediated UC pair found in the X-ray structure of the sarcin/ricin domain from *E. coli* 23 S rRNA (PDB code: 483d) (Correll *et al*., 1999). This can be seen from the superposition in Figure 3b. In this case the UC pairs of the two experimental structures and of the average structures of conformations I and II are shown by superimposing cytosine (C7). Note, however, that the UC pair in this structure is surrounded by a different nucleic environment as compared to the simulated structure. Not only the flanking base pairs are different. The UC pair also occurs in a flexible region of the RNA structure. Nevertheless, the relative orientation of U and C is similar in the experimental and simulated structures. Whereas, however, the both U18 bases of the RNA duplex and of conformation I are approximately located in one common plane, the base planes of U18 are slightly different for the experimental sarcin/ricin structure and for conformation II.

In order to get an impression on the flexibility of the four central base pairs of the duplex structure we have monitored their C1'...C1' distances (Figure 4). The fluctuations in this distance are much more pronounced for the water-mediated pairs than for the flanking GU pairs. The same result is obtained from the time course of H-bond base-base distances shown in Fig. 2. Within the various conformations of the water-mediated pairs the amplitudes of the H-bond distances are much larger than for the GU pair. We have also analysed the time course of base-base distances in the Watson-Crick GC pairs next to the GU pairs (results not shown).

The three H-bonds in the GC pairs exhibit a slightly different flexibility, the central N1-H1...N3 H-bond being the most rigid one. The other two H-bonds between exocyclic amino groups and carbonyl oxygens show a more marked flexibility that is similar to the pattern found in GU pairs. So, from the point of view of flexibility the non-Watson-Crick GU pairs bear resemblance to the Watson-Crick GC pair.

As expected, conformation III with two direct H-bonds leads to a marked reduction of the C1'...C1' distance. This geometrical constraint also affects the neighbouring base pairs and induces the long base pair opening of G8-U17 observed between 1000 and 1200 ps. In the X-ray structure the C1'...C1' distance for the GU pairs is 10.3 Å. This value is by about 0.4 Å smaller than in the conformation III average structure (10.7 Å). For the water-mediated UC pair the experimental distance is 11.7 Å which is in good agreement with the simulation results for conformation I. For example, the average C1'...C1' distance for conformations Ic,d is 11.9 Å. The C1'...C1' distance of the UC pair in the sarcin/ricin domain structure is 10.7 Å. This distance has to be compared to the mean C1'...C1' distance of conformations IIa and IIb. It is approximately 1 Å smaller than found in conformation II. It should be noted again, however, that the nucleic acid environment in this structure is different to the RNA duplex for which the simulations were performed.

Holbrook *et al.* (1997) have pointed out that the RNA dodecamer duplex exhibits only slight deviations from a canonical A-form RNA. To gain more detailed insight we have analysed the experimental duplex structure and two different average structures generated from the simulations using the CURVES algorithm (Lavery & Sklenar, 1988). In one simulated structure both UC pairs adopt conformation I (simulation time range: 1920 - 2230 ps) and in the other one the base pair U7-C18 is in conformation II whereas the C7-U18 pair is still in conformation I (simulation time range: 2230 – 2630 ps). In the experimental structure the most significant deviations in the base pair parameters are seen for the buckle (-11.59º;

+11.59º) and opening (-35.43º; -35.43º) values of the water-mediated UC pairs. These values have to be compared with 0º and –4.2º for a canonical A-RNA. With 2.36 Å the rise between the two water-mediated UC pairs is decreased as compared to the standard value of 2.81 Å. Finally, deviations from canonical A-RNA data are also seen for the roll and twist angle pattern. The simulated structure where both UC pairs adopt conformation I exhibits helical parameters that are similar to the corresponding values of the experimental RNA duplex structure. Finally, we have analysed the average duplex structure with U6-C19 in conformation I and C7-U18 in conformation II. Many of the geometrical parameters are similar to the value of the other simulated structure. Significant differences are seen for the rise between the two UC pairs that is further reduced (1.79 Å) and for the twist angle pattern of base pair steps 5-6, 6-7, and 7-8. The latter change is due to the different orientation of U and C in conformations I and II.

Thus far we have analysed the base pair geometries without taking water directly into account. In a next step, we have selected all water molecules that fulfil H-bond criteria (H-bond donor and acceptor heavy atom distance D...A < 3.5 Å, D-H...A angle > 120º) to both of the base donor or acceptor sites for more than 10 % of the lifetime of conformations I or II. The results are displayed in Table 1. It should be clarified that in conformation I water is both acceptor and donor and in conformation II it is a double acceptor. Within conformations Ic and Id only one water is included. By contrast, conformations Ia, Ib, IIa, and IIb involve two to four long-lived water molecules. In all cases, however, over the major part of the lifetime of a particular conformation water-mediated base-base H-bond contacts are observed. The most extreme case is found for conformation Ic. In this case the residence time of water 1255 is practically identical to the lifetime of the conformation (~ 1 ns). This is the longest water residence time we have ever seen in a molecular dynamics simulation. Note, that the total residence times of water molecules in a particular conformation are smaller than the ''from-

9

to'' residence times. The ''from-to'' times indicate the first and last occurrence of a water molecule fulfilling the H-bond criteria. Within this time range the H-bond criteria may be violated for a short time.

A detailed view of two representative examples of the water dynamics is shown in Figure 5. For conformation Ic the water remains in place over the whole lifetime of the conformation. On the other hand, in conformation IIb four water molecules are involved (Table 1). In Figure 5 the dynamics of the water exchange for three of them is shown. First, water 2146 is linked to the UC pair. It is then replaced by water 1461 which is in turn replaced by water 352. This water exchange occurs without major effects on the base pair geometry.

Figure 5 displays a further interesting feature of the water-base interaction. In conformation IIb both U and C donate H-bonds to the water oxygen and both H...O(water) H-bond distances exhibit only minor fluctuations around a value of 2 Å. A different situation is observed for conformation I where water both accepts and donates an H-bond. As for conformation II the H...O(water) distance is on the average 2 Å. Yet, the distance between the water hydrogen H1 and the N3 acceptor site of cytosine adopts two different average distances around 2 and 3-3.5 Å. A closer look showed that a short distance of water hydrogen 1 is usually accompanied by a large distance for hydrogen 2 and vice versa. This means that the water molecule is rotating from time to time thereby disrupting an H-bond and forming a new one with the other water hydrogen. During this process the H-bond distance between the base donor site and the water oxygen is almost unaffected.

Conformational switching is a general phenomenon in nucleic acids. For example, it has been recognised that switches between alternative conformations of rRNA must occur during translation. Direct evidence for this phenomenon has recently be found in *E. coli* 16S rRNA (Lodmell & Dahlberg, 1997). Base pair switching is also assumed to occur in homologous genetic recombination (Nishinaka *et al.*, 1998). In both cases the base pairs

10

first unpair and then form new pairs after the conformational change.

In addition to these specific examples a systematic base pair disruption is required for replication and transcription of double-stranded nucleic acids. In model systems the dynamics of base pair opening has been studied by the ^{1}H NMR exchange rates of the imino protons upon titration with the exchange catalyst ammonia, for example (Gueron & Leroy, 1995; Dornberger *et al*., 1999). The values of the base pair dissociation constants depend on the base pair type, on the location of the base pair in the helix centre or end and on possible effects exerted by proteins or other ligands. On the whole, this process occurs in the millisecond time range. There are, however, exceptions like the internal C.C+ pairs in the so-called i-motif with a lifetime of hundreds of seconds (Gueron & Leroy, 1995).

A conformational switch within one and the same base pair has been observed for a DNA fragment, which is capable of forming an intramolecular triple helix as well as a hairpin structure (Van Dongen *et al*., 1996). In both conformations base pairing occurs between the first cytosine and guanine of the CCCG loop. However, a Watson-Crick GC pair is formed at neutral pH, whereas a Hoogsteen C(+)-G pair is found at low pH. In this case the base pair remains intact but its geometry and H-bond pattern changes.

In the examples mentioned thus far the conformational change is in all cases induced by external factors like the translation process, genetic recombination or a pH change. In contrast, the conformational switching observed for the water-mediated UC pair is an intrinsic property of the system.

Conformational transitions in nucleic acids have also been studied by molecular dynamics simulations. Here the main focus has been on A/B conformational preferences in DNA (Yang & Pettitt, 1996; Cheatham & Kollman, 1996) or on the conversion between incorrect and correct loop geometries in RNA (Miller & Kollman, 1997; Williams & Hall, 1999). It is also known that extremely short base pair opening events are occasionally

11

observed in molecular dynamics simulations. Examples can be seen for the GU pairs in Figure 2. However, the conformational switching found for the water-mediated UC pairs is to our knowledge a hitherto unknown type of conformational transitions in nucleic acids. It adds a new feature to a better understanding of the dynamics of RNA structures.

Conclusions

A 4 ns MD simulation of the RNA duplex $(r\text{-GGACUUCGGUCC})_2$ reveals switching of the two water-mediated UC base pairs between different conformations. One of the two dominating conformations found in the simulation corresponds to the geometry of water-mediated UC pairs of the duplex X-ray structure with water acting both as H-bond donor and acceptor. The H-bond between the water oxygen and the N3-H3 donor site of uracil is retained during the lifetime of this conformation. However, the water molecule is rotating from time to time thereby disrupting the H-bond between one water OH group and the N3 acceptor site of cytosine and forming a new H-bond involving the other water OH group. Another conformation bears resemblance to a water-mediated UC base pair found in the X-ray structure of the 23S rRNA sarcin/ricin domain. In this case two base donor sites form H-bonds with water oxygen. For a very short time also a direct UC base pair and a further conformation that is similar to the one found in the RNA duplex structure but exhibits an increased H3(U)...N3(C) distance is found. These results show that the water-mediated UC base pairs are characterised by a different dynamics as compared to the Watson-Crick and non-canonical guanine-uracil pairs with two or three direct H-bonds.

The conformational variability and greater flexibility of the water-mediated UC base pairs as compared to conventional pairs is very likely a general property of these unusual pairs. This

12

should be taken into account when considering these structural elements as RNA building blocks and as recognition motifs.

Computational Procedure

MD simulations of the RNA duplex (r-GGACUUCGGUCC)$_2$ were performed with the AMBER program package (Case *et al.*, 1997) adopting the Cornell force field (Cornell *et al.*, 1995) and the starting coordinates were taken from the experimental RNA duplex X-ray structure (PDB entry: 255d; Holbrook *et al.*, 1991). The counterion placement is described elsewhere (Schneider & Sühnel, 1999). We have used 22 Na$^+$ counterions, additional 12 Na$^+$/Cl$^-$ pairs (0.1 mol/l) and 2962 TIP3P water molecules to solvate the RNA. The system density was 1.06 g/cm^3 with box dimensions of 42*57*42 Å^3. The total system consisted of 9696 atoms. It was heated up over 10 ps with a following equilibration run of 100 ps at 300 K. Then an unrestrained simulation was done over 4000 ps at constant temperature (300K) and constant pressure (1 atm) with anisotropic scaling and a time step of 2 fs. SHAKE was applied to all bonds including hydrogen. The nonbonded pair list was updated every 10 steps. The electrostatic interactions were calculated by particle-mesh Ewald summation with a grid spacing of approximately 1 Å. Coordinates were saved at time steps of 1 ps.

References

Arnez, J. G. & Steitz, T. A. (1996). Crystal structures of three misacylating mutants of Escherichia coli glutaminyl-tRNA synthetase complexed with tRNA(Gln) and ATP. *Biochemistry,* **35**, 14725-14733.

Beveridge, D. L. & McConnel, K. J. (2000). Nucleic acids: theory and computer simulation, Y2K. *Curr. Opinion Struct Biol.* **10**, 182-196.

Brandl, M., Meyer, M., Sühnel, J. (1999). Quantum-chemical study of a water-mediated uracil-cytosine base pair. *J. Am. Chem. Soc.* **121**, 2605-2606.

Brandl, M., Meyer, M., Sühnel, J. (in press). Water-mediated base pairs in RNA: A quantum-chemical study. *J. Phys. Chem. A.*

Case, D. A., Pearlman, D. A., Caldwell, J. W.; Cheatham III, T. E., Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., Cheng, A. L., Vincent, J. J., Crowley, M., Ferguson, D. M., Radmer, R. J., Seibel, G. L., Singh, U. C., Weiner, P. K. & Kollman, P.A. (1997). AMBER 5. University of California, San Franscisco.

Cheatham III, T. E. & Kollman, P. A. (1996). Observation of the A-DNA to B-DNA transition during unrestrained molecular dynamics in aqueous solution. *J. Mol. Biol.* **259**, 434-444.

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz Jr., K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **117**, 5179-5197.

Correll, C. C., Freeborn, B., Moore, P. B. & Steitz, T. A. (1997). Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. *Cell,* **91**, 705-712.

Correll, C. C, Wool, I. G. & Munishkin, A. (1999). The two faces of the Escherichia coli 23S rRNA sarcin/ricin domain: the structure at 1.11 Å resolution. *J. Mol. Biol.* **292**, 275-287.

Cruse, W. T. B., Saludjian, P., Biala, E., Strazewski, P., Prange, T. & Kennard, O. (1994). Structure of the mispaired RNA double helix at 1.6 Å resolution and implications for the prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. USA,* **91**, 4160-64.

Dallas, A. & Moore, P. B. (1997). The loop E-loop D region of Escherichia coli 5S rRNA: the solution structure reveals an unusual loop that may be important for binding ribosomal proteins. *Structure,* **5**, 1639-1653.

Van Dongen, M. J., Wijmenga, S. S., Van der Marel, G. A., Van Boom. J. H. & Hilbers, C. W. (1996). The transition from a neutral-pH double helix to a low-pH triple helix induces a conformational switch in the CCCG tetraloop closing Watson-Crick stem. *J. Mol. Biol.* **263**, 715-729.

Dornberger, U., Leijon, M. & Fritzsche, H. (1999). High base pair opening rates in tracts of GC pairs. *J. Biol. Chem.* **274**, 6957-6962.

Gueron, M. & Leroy, J.-L. (1995). Studies of base pair kinetics by NMR measurement of proton exchange. *Methods Enzymol.* **261**, 383-413.

Holbrook, S. R., Cheong, C., Tinoco, Jr., I. & Kim. S.-H. (1991). Crystal structure of an RNA double helix incorporating a track of non-Watson-Crick pairs. *Nature,* **353**, 579-581.

Kay, L. E. (1998). Protein dynamics from NMR. *Nature Struct. Biol.,* **5**, 513-517.

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946-950.

Lavery, R. & Sklenar, H. (1988). The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* **6**, 63-91.

Lodmell, J. S. & Dahlberg, A. E. (1997). A conformational switch in Escherichia coli 16S ribosomal RNA decoding of messenger RNA. *Science,* **277**, 1262-1267.

Miller, J. L. & Kollman, P. A. (1997). Theoretical studies of an exceptionally stable RNA tetraloop: observation of convergence from an incorrect NMR structure to the correct one using unrestrained molecular dynamics. *J. Mol. Biol.* **270**, 436-450.

Nishinaka, T., Shinohara, A., Ito, Y., Yokoyama, S. & Shibata, T. (1998*).* Base pair switching by interconversion of sugar puckers in DNA extended by proteins of RecA-family: a model for homology search in homologous genetic recombination. *Proc. Natl. Acad. Sci. USA,* **95**, 11071-11076.

Phan, A. T., Leroy, J.-L. & Gueron, M. (1999). Determination of the residence time of water molecules hydrating *B'*-DNA and B-DNA, by one-dimensional zero-enhancement Nuclear Overhauser effect spectroscopy. *J. Mol. Biol.* **286**, 505-519.

Rath, V. L., Silvian, L. F., Beijer, B., Sproat, B. S. & Steitz, T. A. (1998). How glutaminyl-tRNA synthetase selects glutamine. *Structure,* **6**, 439-449.

Rife, J. P., Stalling, S. C., Correll, C. C., Dallas, A., Steitz, T. A. & Moore, P. B. (1999). Comparison of the crystal and solution structures of two RNA oligonucleotides. *Biophys. J.,* **76**, 65-75.

Rould, M. A., Perona, J. J., & Steitz, T. A. (1991). Structural basis of anticodon loop recognition by glutaminyl-tRNA synthetase. *Nature,* **352**, 213-218.

Schneider, C. & Sühnel, J. (1999). A molecular dynamics simulation of the flavin mononucleotide-RNA aptamer complex. *Biopolymers,* **50**, 287-302.

Tanaka, Y., Fujii, S., Hiroaki, H., Sakata, T., Tanaka, T., Uesugi, S., Tomita, K. & Kyogoku, Y. (1999). A'-form RNA double helix in the single crystal structure of r(UGAGCUUCGGCUC). *Nucleic Acids Res.* **27**, 949-955.

Williams, D. J. & Hall, K. B. (1999). Unrestrained stochastic dynamic simulations of the UUCG loop using an implicit solvation model. *Biophys. J.* **76**, 3192-3205.

Yang, L. & Pettitt, M. (1996). B to A transition of DNA on the nanosecond time scale. *J. Phys. Chem.* **100**, 2564-2566.

Table 1. Lifetimes of conformations I and II and residence times of long-lived water molecules involved in these conformations (in ps) [a].

Conformation	Base pair	Water number	Conformation lifetime (from-to)	Residence time of long-lived water molecules (from-to)	Total residence time of long-lived water molecules within that conformation
Ia	U6-C19	88	710 (1920-2630)	211 (1925-2136)	632
		643		421 (2208-2629)	
Ib	U6-C19	2939	920 (3080-4000)	102 (3167-3269)	821
		2673		322 (3281-3602)	
		1964		397 (3604-4000)	
Ic	C7-U18	1255	1006 (1225-2231)	1006 (1225-2231)	1006
Id		1911	320 (3680-4000)	270 (3684-4000)	270
IIa	U6-C19	273	1150 (770-1920)	123 (774-931)	749
		1980		347	

				(1047-1474)	
		761		279	
				(1474-1837)	
IIb	C7-U18	426	1450	184	1227
			(2230-3680)	(2233-2463)	
		2146		409	
				(2471-2954)	
		1461		260	
				(2937-3231)	
		352		374	
				(3237-3652)	

[a] Only water molecules are taken into account that fulfil simultaneously H-bond criteria (H-bond donor and acceptor heavy atom distance D...A < 3.5 Å, D-H...A angle > 120º) to both bases and have a residence time exceeding 10% of the lifetime of a particular conformation. See Figure 2 for more information on the occurrence of conformations I and II. The from-to difference indicates the first and last frame for which the H-bond criteria with a water molecule are fulfilled. The residence times of long-lived water molecules amounts to 77-94% of the 'from-to' difference. This is due to the fact that the H-bond criteria may be violated and are therefore not fulfilled for a short time.

Figure Captions

Figure 1. Secondary structure of the simulated RNA duplex and structural formula of the water-mediated UC base pair occurring in the duplex X-ray structure (Holbrook *et al*., 1991; PDB code: 255d).

Figure 2. Time course of selected inter-base distances for the central UC and the flanking GU pairs. Roman numerals indicate different base pair conformations.
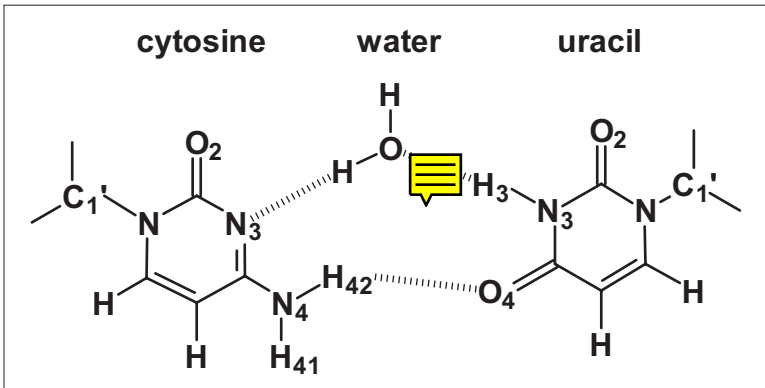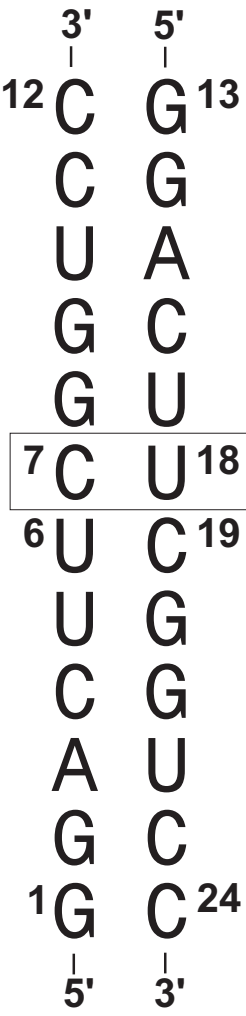
Figure 3.a. Superposition of the experimental RNA duplex structure (Holbrook *et al*., 1991; PDB code: 255d; dark grey) with two average structures from the simulations (light grey) where either both UC pairs adopt conformation I (simulation time 1920 – 2230 ps; thick sticks) or one UC pair switches to conformation II (simulation time 2230 – 2630 ns; thin sticks). The superposition was done for the backbone atoms P, O1P, O2P, O3' and O5' atoms of the flanking GU pairs. Only the base pair C7-U18 is shown. The Figure was generated with InsightII (Molecular Simulations, Inc.).
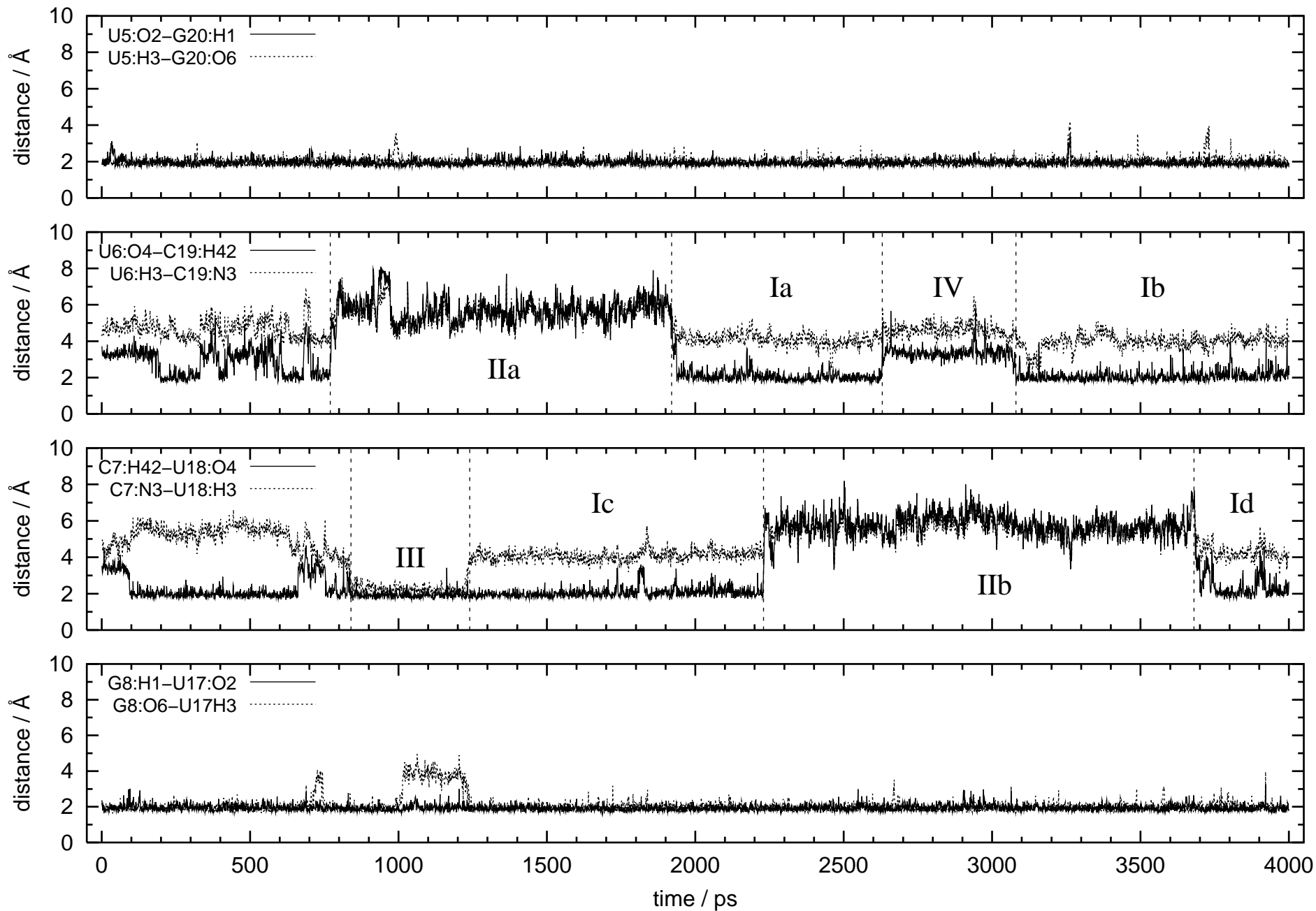
Figure 3.b. Superpositions of average structures of conformations I (simulation time 1920 – 2230 ps; light grey, thick sticks) and II (simulation time 2230 – 2630 ps; light grey, thin sticks) with the experimental geometries of the UC pairs in the A-RNA duplex (Holbrook *et al*., 1991; PDB code: 255d; dark grey, thick sticks) and the sarcin/ricin domain 23 S rRNA structure (Corell *et al*., 1999; PDB code: 483d; dark grey, thin sticks). In order to show the relative orientation of bases cytosine was superimposed. Due to the superposition only the thick sticks representation is visible for cytosine. Note, that in the simulation both cytosine and uracil are moving in passing from conformation I to II. The water molecules were

20

manually placed. Conformation I is similar to the UC base pair geometry in the A-RNA duplex and conformation II has a similar geometry as the water-mediated UC pair in the sarcin/ricin 23 S rRNA structure. The Figure was generated with Molscript (Kraulis, 1991).
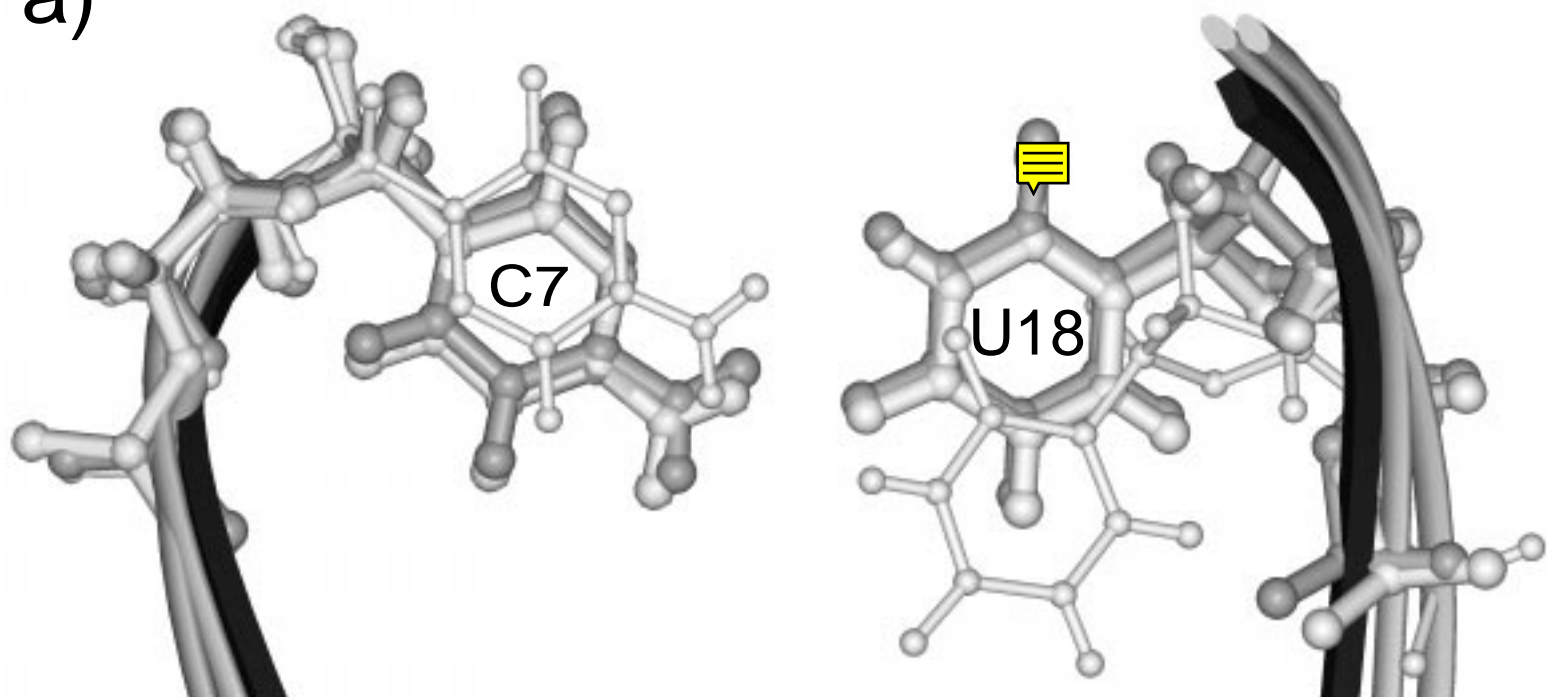
Figure 4.   Time course of C1'...C1' distances for the central UC and the flanking GU pairs of the simulated duplex structure.

Figure 5.   Time course of water-base H-bond distances for selected long-lived water molecules incorporated in conformations I and II of the water-mediated UC pair. In conformation Ic the water residence time is identical to the conformation lifetime. For conformation IIb the exchange of three out of the four water molecules involved is shown.