MOLECULAR DYNAMIC SIMULATIONS OF LARGE RNA MOLECULES: THE YEAST tRNA^{Phe}

BARTOSZ PLISZKA AND STANISŁAW OŁDZIEJ

Faculty of Chemistry, University of Gdansk, Sobieskiego 18, 80–952 Gdansk, Poland

Abstract: Molecular dynamics trajectories (700 ps) of the solvated and neutralized 75–residue yeast tRNA^{Phe} were generated using the AMBER 5.0 molecular dynamics software package. The cut–off scheme was used to treat electrostatic interactions; consequently, all long–range interactions beyond 12 angstroms were neglected. The equilibration procedure and conditions during simulations led to a dynamically stable model of the tRNA molecule. During the simulations all base–base interactions (which determine the secondary and the tertiary structure of the molecule) were well preserved. Consequently, the global shape of the molecule was preserved well and the RMS deviation calculated between the starting x–ray structure and the final structure after 700 ps of simulations was 3.25 angstroms. The biggest deviation is observed in the region of the anticodon hairpin loop; this high mobility is associated with the presence of a very unusual Y–base and a binding site of a magnesium ion in this region.

Keywords: tRNA^{Phe}, molecular dynamics, modified bases, nucleic acids, electrostatic interaction

1. Introduction

Transfer RNAs belong to the first class of functional nucleic acid molecules for which three–dimensional structures were determined by x–ray crystallography [1, 2]. The difficulties in obtaining good quality crystals of nucleic acids and many technical problems did not allow to obtain three–dimensional structures of other RNAs and the tRNAs are still some of the biggest ribonucleic acids whose three–dimensional structures are known. In 1987 McCammon and Harvey undertook the first large–scale molecular dynamics studies of the yeast tRNA^{Phe} [3]. In their calculations, solvation was implicit and reducing the atomic charges on the phosphate group atoms approximated the effect of the presence of counterions. A very short (only 32 ps) simulation showed that most of the base–base interactions (which determine the secondary structure of the tRNA molecule) had been preserved; however several tertiary interactions had disappeared. This first attempt at molecular dynamics simulations of a large RNA molecule has demonstrated that proper treatment of long–range electrostatic interactions, as well as taking into account the presence of metal ions and water are necessary to correctly reproduce

the nucleic acid structure and dynamics. The introduction of new algorithms into molecular dynamic programs, which allows one to carry out accurate calculations of electrostatic interaction energy [for example the Partial Mesh Ewald (PME) method] [4, 5], as well as the development of force–field parameters and new water models has made it possible to obtain long time molecular dynamics trajectories of large and complex nucleic acid molecules. Recently Westhlof and coworkers published a series of papers on long time simulations (500 ps) of three RNA molecules: the anticodon hairpin loop from yeast tRNA(Asp) [6–8], hammerhead ribozyme [9] and yeast tRNA(Asp) [10, 11]. In all calculations the authors used identical equilibration protocol and identical simulation conditions: constant volume, constant temperature (300 K), and the PME method for the calculation of electrostatic interactions. The results of their work are very promising; for the first time the authors obtained long–time stable molecular dynamic trajectories for nucleic acids and, as opposed to the previous calculations, no important base–base interactions were lost. Nevertheless, many technical problems still remain unsolved. One of the biggest problems is that all these calculations were carried out with a too low density of the solvent, which is an unphysical condition. Automatic procedures implemented in molecular dynamics packages, such as AMBER or CHARMM add water molecules to the solute molecule and usually create a rectangular box that includes the solute and the solvent. The solvent density in such a box is usually in the range of 0.7–0.9 g/cm, which is lower than the density of pure water. This problem can be solved by running a short molecular dynamics simulation at a constant pressure, in order to allow the density to reach the expected value. The purpose of our work was to check the importance of the equilibration protocol, based on molecular dynamics simulations of yeast tRNA(Phe).

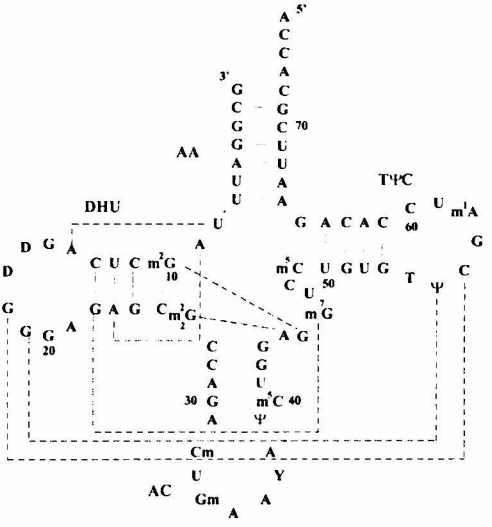

Figure 1. *The secondary structure of the yeast tRNA^{Phe}. Tertiary base-base interactions are indicated as dashed lines. DHU — dihydrouracil loop, TYC — pseudouracil loop, AC — anticodon loop AA — anticodon arm, A — adenine, C — cytosine, U — uridine, G — guanosine, abbreviation of modified nucleosides see Figure 2*

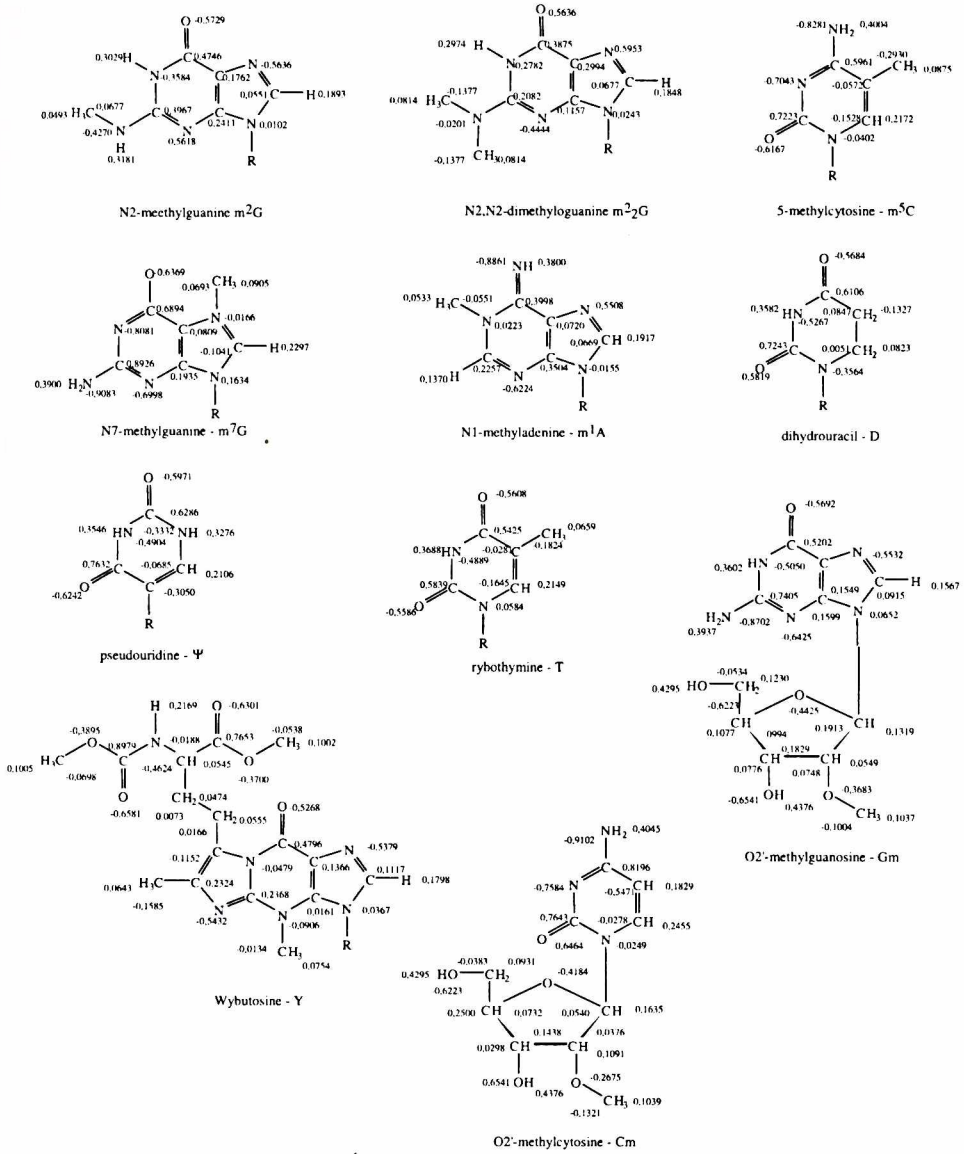

Figure 2. *Partial atomic charges and chemical structure of modified nucleosides presented in yeast tRNA^{Phe}*

The secondary structure of yeast tRNA(Phe) is shown in Figure 1. All tRNAs possess about 17% post–translationally modified nucleosides in their nucleoside sequences [12, 13]. Most of them are products of different types of mono– or dimethylation of the amino groups of cytosine, guanine or adenine nitrogen base, respectively. As shown in Figure 1, there are eleven different types of such nucleosides in the yeast tRNA^{Phe} sequence. The lack of fully relevant force–field parameters for modified nucleosides was one of the drawbacks of molecular dynamics simulations of such types of RNA. These modified bases were parameterized in the present work.

2. Methods

The starting coordinates were extracted from the yeast $tRNA^{Phe}$ x-ray structure (PDB code 6TNA) [1]. All metal ions and water molecules present in the original structure were removed. The tRNA molecule was surrounded by 74 Na^+ counterions and 19812 TIP3P water molecules [15], filling a 97.4×78.4×97.4 angstroms rectangular box. The counterions were placed based on the electrostatic potential of the solvated system. A 12-angstrom truncation distance was applied to electrostatic interactions. The electrostatic interaction list for atoms which were in the distance range between 8 and 12 angstroms was updated every 25 steps, and every step for atoms in the distance range between 0–8 angstrom. A 9-angstrom truncation distance was applied to Lennard-Jones interactions. The nonbonded pair list was updated every 10 steps. The SHAKE procedure [16] was applied to all bonds in the system. The trajectory was run at a constant temperature of 300 K. All calculations were carried out using the SANDER module of the AMBER [17] package, with the AMBER94 force-field parameters [18].

The modified nucleosides not present in the standard AMBER database were built according to the AMBER program manual [17]. The partial atomic charges were obtained by fitting the electrostatic potential calculated at the *ab initio* level of theory using the 6–31G* basis set, by applying the RESP [18] program from the AMBER package. The structures of modified bases (nucleosides) and partial atomic charges are shown in Figure 2.

All other force-field parameters (the Lennard-Jones parameters, force constants etc.) were adapted from the AMBER database and literature [19].

The equilibration protocol consisted of 50000 steps of steepest-decent minimization, followed by 100 ps of molecular dynamics simulations at a constant pressure with 1 fs time step. The production phase consisted of 150 ps of molecular dynamics at a constant pressure with 2 fs time step and a 450 ps molecular dynamics at a constant volume with 2 fs time step.

All calculations were carried out at the Interdisciplinary Center for Mathematical and Computer Modeling (ICM), Warsaw, Poland, using the T3E SGI-Cray supercomputer and at the Informatics Center of the Metropolitan Academic Network (IC MAN) at the Technical University of Gdansk, Poland, using the ONYX SGI machine. Every 10 ps of simulations required 9 wall-clock time hours on sixteen processors of the T3E SGI-Cray machine and 21 wall-clock time hours on four processors of the ONYX SGI machine, respectively.

3. Results

To get a rough estimate of the quality of the MD trajectory, the root mean square deviation (RMSD) from the starting crystal structure was calculated. We used two sets of atoms to calculate the RMSD. One set contained only the phosphorus atoms and the corresponding RMSD was used to determine the mobility of the phosphate backbone, while the second one contained the N1 or N9 nitrogen atoms (for pyrimidines or purines, respectively) and the corresponding

RMSD was used to determine the mobility of the bases (Figure 3).

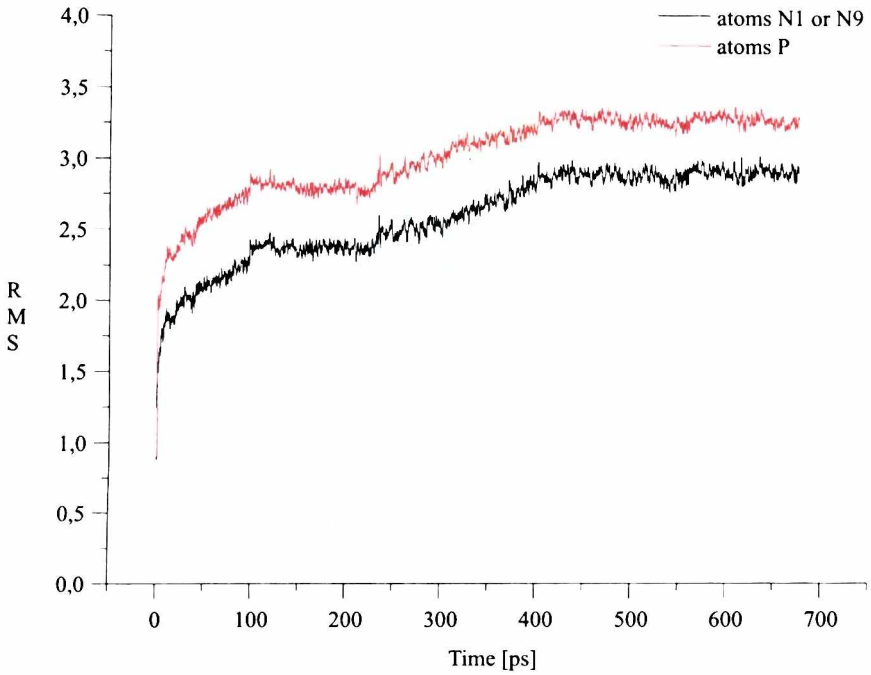

Figure 3. *Time course of the RMS deviation from the startin structure calculated by using P atoms (red) or by N1 and N9 atoms (black)*

The RMSD rises very fast during the first few picoseconds of simulation and is stabilized at 1.75 angstroms for the N1 or N9 atoms and 2.75 angstroms for the P atoms, respectively, after 100 ps. The change of simulation parameters (see Methods) after 250 ps led to another increase in RMSD, until the value of 2.75 angstrom (N1 and N9 atoms) and 3.25 angstrom (P atoms), respectively, was reached after 400 ps of simulations. As was expected, the RMSDs calculated for the phosphorus atoms are higher than the RMSDs calculated for the N1 and N9 atoms. The difference shows a higher mobility of the phosphate–sugar backbone, compared to that of the core of the molecule, which is composed of the hydrogen bonded bases. The RMSDs obtained from our calculations are similar with those calculated by Auflinger and coworkers for the yeast tRNA^{Asp} [11].

The initial structure of the yeast tRNA^{Phe} and the structure obtained after 700 ps of simulations are superposed in Figure 4 (the backbone atoms were used in superposition).

As shown in Figure 4, the global shape of molecule does not change. However, a significant deformation of the region of the anticodon hairpin loop can be observed. These changes are better visualized in Figure 5, which shows the time–evolution of the RMSD calculated over phosphorus atoms for every atom separately. It can be observed that the highest RMSDs occur for residues in the

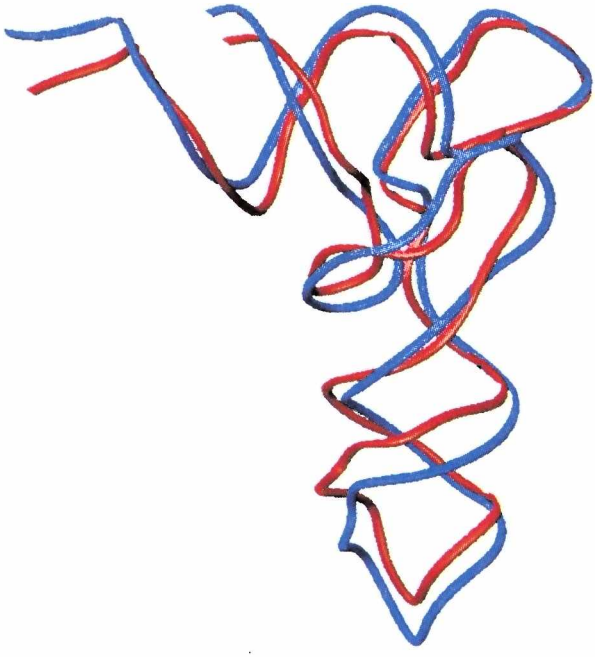

Figure 4. *The superposition of the x-ray structure of the yeast $tRNA^{Phe}$ (blue) and final structure after 700 ps of molecular dynamics simulations (red)*

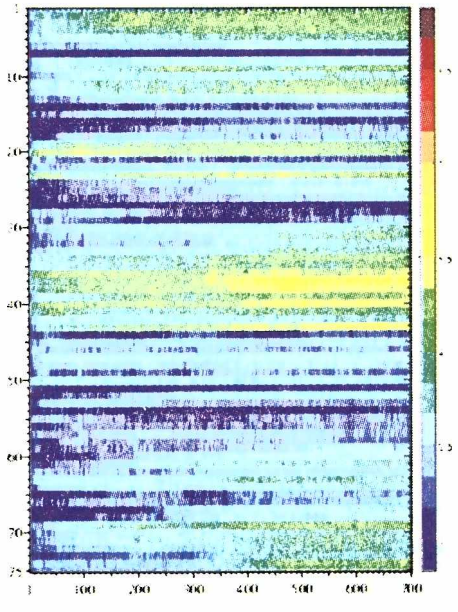

Figure 5. *The contour plot ilustrated the evolution of the geometry of the yeast $tRNA^{Phe}$, measured as deviation from the strating structure (in angstrom). Molecule is represented by P atoms sequential position along the vertical axis, over the time of simulation horizontal axis*

anticodon hairpin loop (the 3' and 5' ends) and in the region of dihydrouracil hairpin loop. The results shown in Figures 4 and 5 agree with the previously described observation that all structural elements of the tRNA moiety which are not involved in secondary or tertiary hydrogen bonds interactions exhibit high mobility [10]. All structural elements with well–defined secondary structure, such as helices, are clearly visible in Figure 5 as "cool" blue strips with very small deviation from the original crystal structure. The fact that the highest RMSDs are observed for the anticodon hairpin loop could also be associated with the absence of magnesium ions in our simulations. Four magnesium ions are present in the x–ray structure of yeast tRNA^{Phe}, one of which is deeply inserted into a small grove of the anticodon hairpin loop close to residues 39–37 [1, 2]. This region shows a considerable RMSD from the starting x–ray structure (Figure 5). Another magnesium ion binding site is localized close to residues 19–20 and higher mobility of this region is also observed in this case.

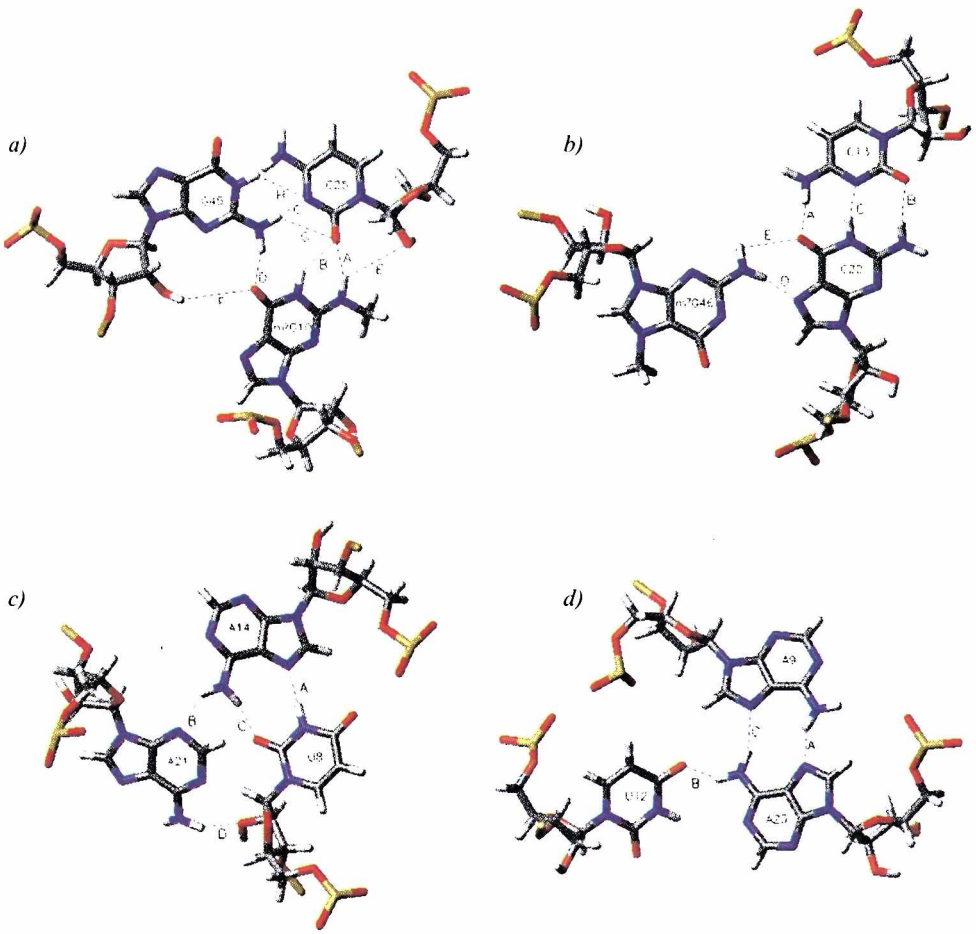

Figure 6. *The structure and hydrogen bond network of four base triplets presented in the structure of the yeast tRNA^{Phe}: a) G45–m^{2}G10–C25; b) m^{7}G46–C13–G22; c) U8–A14–A21; d) U12–A23–A9*

From the structural point of view the most important interactions observed in the three–dimensional structure of tRNA are base triplets. Four base triplets: G45–m^{2}G10–C25, m^{7}G46–C13–G22, U8–A14–A21, and U12–A23–A9 are present in the original yeast tRNA^{Phe} molecules. The hydrogen bond network and the structure of the triplets are shown in Figure 6. The dynamic stability of the triplets is the one of the best available criteria for estimating the quality of a MD trajectory. This stability has been evaluated by calculating hydrogen–bonding percentage (HB%). These HB% are defined as the time over which a hydrogen bond satisfies the following criteria: (i) the distance between acceptor and donor is smaller than 4 angstroms and (ii) the angle between acceptor–hydrogen–donor is in range of 120–180 degrees. As shown in Figure 6a, eight hydrogen bonds with HB% values: A = 100%, B = 100%, C = 92 %, D = 100%, E = 24%, F = 21%, G = 31% and H = 43% can be observed in the G45–m^{2}G10–C25 triplet. Because the bases in positions 10 and 25 are not coplanar and no standard Watson–Crick–like structure is formed, we observed only one strong hydrogen bond between them (C). A series of hydrogen bonds between the 10–25 pair and the guanosine in position 45 can be observed. The bases in positions 10 and 25 are almost coplanar and therefore the strongest hydrogen bonds are observed between those two bases. Figure 6b shows a very well defined m^{7}G46–C13–G22 triplet. In this case we can observe five hydrogen bonds with the following HB% values: A = 83%, B = 100%, C = 97%, D = 66%, E = 82%. This is an example of a perfect base triplet with all bases in one plane, two bases (of residues 13 and 22, respectively) forming a "classical" Watson–Crick base pair and very strong and stable hydrogen bonds. As shown in Figure 6c, the U8–A14–A21 triplet formed four hydrogen bonds with the following HB% values: A = 100 %, B = 99%, C = 100 %, D = 99%. This is an example of a very rare base triplet when two bases A14 and U8 form a very unusual base pair called "reversed Hoogsteen". All three bases in this triplet are coplanar and all observed hydrogen bonds are very strong. A similar very well defined triplet can be observed in Figure 6d. The U12–A23–A9 triplet is bound by three very strong hydrogen bonds with the following HB% values: A = 96%, B = 100% and C = 98%. Only the G45–m^{2}G10–C25 triplet shows significant distortion from coplanarity. For all triplets the hydrogen bond network observed in the original x–ray structure is preserved and in some cases some hydrogen bonds not observed in crystalline phase are formed. The results obtained in our simulations show that base triplets are very stable and all tertiary interactions are preserved. The molecular dynamics simulations of tRNAs reported by McCammon and Westhof [3, 6–11] groups show that the triplets are not very stable and some or all hydrogen bond interactions determining the triplet structure can be broken during simulations.

4. Conclusions

The present molecular dynamics simulations of the yeast tRNA^{Phe} constitute a step forward towards understanding the dynamic behavior of large RNA molecule. We found that (i) using the accurate, but expensive PME method for

long–range electrostatic interactions treatment is not necessary to obtain stable long–time molecular dynamic trajectory as was postulated in previous studies; (ii) the simulations at constant pressure conditions give a smaller deviation from the original x–ray structure compared to simulations at constant volume used widely so far; (iii) the presence of divalent metal ions is necessary to reduce some deformation in the tRNA structure during molecular dynamics simulations. The results presented in this paper show that it is now possible to carry out accurate molecular dynamics simulations of RNA molecules of the size of small proteins on nanosecond scale.

Acknowledgements

This wark was supported by grant DS.8245–4–0124–9 from the Polish State Committee for Scientific Research (KBN). Calculations were carried out with the use of the resources and software at the Informatics Center of the Metropolitan Academic Network (IC MAN) at the Technical University of Gdansk and Interdisciolinary Center for Mathematical and Computer Modeling (ICM), Warsaw, Poland. The authors thank to Dr Adam Liwo in his help in preparation of the manuscript.

References

[1] Sussman J.L., Holbrook S.R., Warrant R.W., Church G.M., Kim S.–H., *Crystal structure of yeast phenylalanine tRNA. I. Crystallographic refinement.* J.Mol.Biol., **123**: 607–628 (1978)

[2] Brown R.S., Dewan J.C., Klug A., *Crystallographic and biochemical investigation of the lead(II)-catalyzed hydrolizys of yeast phenylalanine tRNA.* Biochemistry, **24**: 4785–4801 (1985)

[3] McCammon J.A., Harvey S.C., *Dynamics of protein and nucleic acids.* Cambirdge University Press, New York, 1987

[4] Darden T., York D., Pedersen L.G., *Particle mesh Ewald: an Nlog(N) method for Ewald sum in large systems.* J.Chem.Phys., **98**: 10089–10092 (1993)

[5] Essmann U., Perera M., Berkowitz M.L., Darden T., Lee H., Pedersen L.G., *A smooth particle mesh Ewald method.* J.Chem.Phys., **103**: 8577–8593 (1995)

[6] Auffinger P., Louise–May S., Westhof E., *Molecular dynamics simulations of the anticodon hairpin of tRNA^{Asp}: structuring effects of C–H...O hydrogen bonds and long–range hydration force.* J.Am.Chem.Soc. **118**: 1181–1189 (1996)

[7] Auffinger P., Westhof *E., H–bond stability in the tRNA^{Asp} anticodon hairpin: 3ns of multiple molecular dynamics simulations.* Biophys. J. **71**: 940–954 (1996)

[8] Auffinger P., Westhof E., *RNA hydration: three nanoseconds of multiple molecular dynamics simulations of the solvated tRNA^{Asp} anticodon hairpin.* J.Mol.Biol. **269**: 326–341 (1997)

[9] Hermann T., Auffinger P., Westhof E., *Molecular dynamics investigations of the hammerhead ribozyme RNA.* Eur.J.Biophys. **27**: 153–165 (1998)

[10] Auffinger P., Westhof E., *Simulations of the molecular dynamics of nucleic acids.* Curr.Opin.Struct.Biol. **8**: 227–236 (1998)

[11] Auffinger P., Louise-May S., Westhof E., *Molecular dynamics simulations of solvated yeast $tRNA^{Asp}$*. **76**: 50–64 (1999)

[12] Sprinzl M., Horn C., Brown M., Ioudovich A., Steinberg S., *Compilation of tRNA sequences and sequences of tRNA genes*. Nucleic.Acid.Res., **26**, 148–153 (1998)

[13] Rozenski J., Crain P., McCloskey J., *The RNA modification database: 1999 update*. **27**: 196–197 (1999)

[14] Jorgensen W.L., Chandrasekhar J., Madura J.D., *Comparison of simple potential functions for simulating liquid water*. J.Chem.Phys. **79**: 926–935 (1983)

[15] Pearlman D.A., Case D.A., Caldwell J.W., Ros W.S., Cheatham III T.E., DeBolt S., Ferguson D., Seibel G., Kollman P.A., *AMBER, a package of computer program for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structure and energetics properties of molecules*. **91**: 1–41 (1995)

[16] Case D.A., Pearlman D.A., Caldwell J.W., Cheatham III T.E., Ross W.S., Simmerling C.L., Darden T.A., Metz K.M., Stanton R.V., Cheng A.L., Vincent J.J., Crowley M., Ferguson D.M., Radamer R.J., Seibel G.L., Singh U.C., Weiner P.K., Kollman P.A., AMBER 5.0, Uniwersity of California, San Francisco, 1997

[17] Weiner S.J., Kollman P.A., Nguyean D.T., Case D.A., *An all-atom force field for simulations of proteins and nucleic acid*. J.Comput.Chem., **7**: 230–252 (1996)

[18] Cieplak P., Cornell W.D., Bayly C., Kollman P.A., *Application of the multimolecule and multiconformational RESP methodology to biopolymers: charge derivation for DNA, RNA, and proteins*. J.Comput.Chem., **16**: 1357–1377 (1995)

[19] Czaplewski C., Pasenkiewicz-Gierula M., Ciarkowski J., *Molecular dynamics of vasoperssin V2 receptor in a phospholipid bilayer membrane*. **19**: 355–367 (1998)