Multiple Molecular Dynamics Simulations of the Anticodon Loop of $tRNA^{Asp}$ in Aqueous Solution with Counterions

Pascal Auffinger, Shirley Louise-May, and Eric Westhof*

Contribution from the Institut de Biologie Moléculaire et Cellulaire du CNRS, Modélisations et Simulations des Acides Nucléiques, UPR 9002, 15 rue René Descartes, 67084 Strasbourg Cedex, France

Received January 9, 1995[⊗]

Abstract: In a systematic search for a stable protocol with which to extend our dynamical investigations, a nanosecond of molecular dynamics simulations of the solvated anticodon loop of $tRNA^{Asp}$ consisting of ten unique trajectories was obtained by slight modifications to the starting conditions. These changes produced divergent trajectories which varied widely in structural and dynamical characteristics. However, the properties of these trajectories could not be directly correlated to the slight modifications introduced in the system, and thus, questions were raised regarding the probity of the standard protocol we utilized. Instead of a detailed analysis of the results, the multiple molecular dynamics (MD) approach was used as a diagnostic for estimating the reliability of the set of trajectories generated and the extent of relevant biochemical information which can be extracted from it. We address here issues concerning critical evaluation of molecular dynamics methodology and detection of protocol instabilities. We infer that an ensemble of initial uncorrelated trajectories should be generated in order to investigate the constancy of structural and dynamical properties of the system under study.

Introduction

Molecular dynamics simulations are becoming a powerful tool in studying a great number of complex physicochemical and biochemical systems[1−4] because they allow the calculation of many microscopic structural and dynamical properties for a plausible model of a real system, properties which may only be inferred with difficulty or not at all by classical experimental means. The molecular dynamics method utilizes an assumed force potential, and a protocol for preparation of the system from a rational starting configuration to generate a Newtonian trajectory through the accessible theoretical phase space at a given temperature. An analysis and interpretation of the properties of the system trajectory sampled during an equilibrium state can then be made. However, the molecular dynamics method is inherently a stochastic process, and it is not straightforward to design a system that will achieve a desired equilibrium state on a reasonable MD time scale. Further, the chaotic nature of such simulations, showing extreme dependency on slight modifications of the starting conditions, as shown by some authors,[5−9] should necessitate sampling over multiple trajectories in order to afford a statistical understanding of the system dynamics. Thus far, systematic testing of protocols is not a general practice. Additionally, there is no rigorous method for determining the relevance of the generated results with respect to experimental data, nor a reliable standard for defining an equilibrated system.

Presently, because of computational limitations and the complexity of physicochemical and biochemical systems of interest, usually only one of the many possible trajectories in the theoretically accessible phase space for a given system and protocol are generated and analyzed in detail. Early criteria used to demonstrate the stability of the system were related to the conservation of the total energy of the system.[10] While this is essential, structural equilibrium was not neccessarily sought. Calculations of the one dimensional root mean square deviation (RMSD) from the starting configuration of the system as a function of time has become a standard measure of the "structural equilibrium" of MD studies; a plateau in the time dependent plot of this property signaling an acceptable equilibrium state. In more sophisticated treatments, two dimensional plots of RMS deviations have been used to investigate the structural relatedness of differing parts of an MD trajectory or between multiple MD trajectories, all displaying similar RMS values. The results of these studies support the principle of structural substates at the subnanosecond time scale.[4,11] Thus, two dimensional RMSD plots highlight hidden structural topologies present in the one dimensional plots. The presence of multiple structural substates at the nanosecond time scale brings back to fore the question of the definition of an equilibrated state of the system and of the time scale on which structural transitions occur.

A recent paper describing a 1 ns simulation of BPTI in its crystalline environment, using an Ewald summation approach for the treatment of long-range interactions, achieving an RMS

[⊗] Abstract published in *Advance ACS Abstracts*, June 1, 1995.

(1) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids;* Cambridge University Press: New York, 1987.

(2) Brooks, C. L.; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics;* J. Wiley & Sons: New York, 1988; Vol. LXXI.

(3) van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 992−1023.

(4) Beveridge, D. L.; Ravishanker, G. *Curr. Opin. Struct. Biol.* **1994**, *4*, 246−255.

(5) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, 1987.

(6) Pearlman, D. A.; Case, D. A.; Caldwell, J. C.; Seibel, G. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. AMBER 4.0 Manual, San Francisco, CA, 1991.

(7) Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, *15*, 667−683.

(8) Avbelj, F.; Moult, J.; Kitson, D. H.; James, M. N. G.; Hagler, A. T. *Biochemistry* **1990**, *29*, 8658−8676.

(9) Kitson, D. H.; Avbelj, F.; Moult, J.; Nguyen, D. T.; Mertz, J. E.; Hadzi, D.; Hagler, A. T. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8920−8924.

(10) Kitchen, D. B.; Hirata, F.; Westbrook, J. D.; Levy, R.; Kofke, D.; Yarmush, M. *J. Comput. Chem.* **1990**, *11*, 1169−1180.

(11) McConnel, K. J.; Nirmala, R.; Young, M.; Ravishanker, G.; Beveridge, D. L. *J. Am. Chem. Soc.* **1994**, *116*, 4461−4462.

0002-7863/95/1517-6720$09.00/0

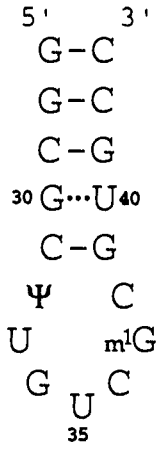

Figure 1. Two dimensional representation of the anticodon loop of $tRNA^{Asp}$. Ψ and m^1G notations are referring to the nonstandard pseudouridine and 1-methylguanine residues.

backbone deviation from the X-ray structure of 0.33 Å,[12] could serve to set a higher standard in evaluation of less approximate MD methodologies. The most pertinent result of this study seems to be that as the description of the system becomes more explicit and accurate, the structural and dynamical properties of the system are seen to remain close to those of the crystal structure or starting configuration. This implies that conservation of the structural interactions present in the initial system configuration, at least on short time scales, signals a true system equilibrium. It follows that undue rupturing or reorganization of these interactions, while tempting one to postulate unique alternate conformations or possible mechanistic implications, could possibly be the sign of an inferior theoretical model and care should be taken in their interpretation.

Present computational means allow for simulations of larger and more complex, fully solvated systems with explicit inclusion of mobile ions. Thus far, the trend has been toward pushing the limits of both MD methodology and computational means in single extended simulations in which all environmental effects are included explicitly and the trajectory computed for several hundreds of picoseconds up to 1 ns. While this is required in order to calculate properties of the system with long relaxation times, the accumulation of statistical errors or effects of methodological perturbations over longer trajectories is not trivial and should be considered. Therefore, it is essential to include systematic testing for each new protocol used in order to determine its range of applicability.

In the following, we describe a nanosecond of MD simulations on the solvated anticodon loop of $tRNA^{Asp}$ with the aim of addressing the ramifications of the differences observed in ten 100 ps molecular dynamics trajectories generated by slight modifications to the initial protocol of the system. The constancy of the set of trajectories, the utility of using multiple runs as a diagnostic of protocol stability and as an alternative means of sampling the theoretical conformational space, and finally the conscientious critical examination of a set of divergent MD trajectories are explored and discussed.

Computational Procedure

Ten 100 ps MD trajectories on the solvated anticodon loop of $tRNA^{Asp}$ were run. The system configuration consisted of the 17 bases of the anticodon hairpin of $tRNA^{Asp}$ (Figure 1) neutralized by 16 NH_4^+ counterions and solvated by 2856 SPC/E water molecules,[13] filling a 55.7 × 43.2 × 42.7 $Å^3$ box. The starting atomic positions of the hairpin were extracted from the crystal structure of $tRNA^{Asp}$.[14] The counterions were placed 5 Å from the phosphorus atom along the OPO bisector. The AMBER4.1 package was used[15] to run several simulations at a constant temperature of 298 K, a constant pressure of 1 atm, and a time step of 2 fs, with the same starting coordinates for the RNA fragment, counterions, and water molecules for each simulation unless specified otherwise. In order to minimize effects due to truncation of long-range electrostatic interactions, particularly important in highly charged systems like nucleic acids, no cutoffs were applied to *solute−solute* interactions, the solute being a neutral group of molecules comprising the anticodon hairpin and the NH_4^+ counterions. Therefore all electrostatic interactions between solute residues, including the charged counterions and phosphate groups, were explicitly evaluated. No periodic boundary conditions were applied on *solute−solute* interactions. An 8 Å residue based cutoff was used for the shorter ranged *solute−solvent* and *solvent−solvent* interactions. For all simulations, unless otherwise specified, the Pearlman and Kim set of charges, derived from low temperature X-ray data of isolated nucleotides, was used.[16] Differences between these two sets for the ribose and cytosine groups are shown Figure 2. The charges of the Ψ32 and m^1G37 residues

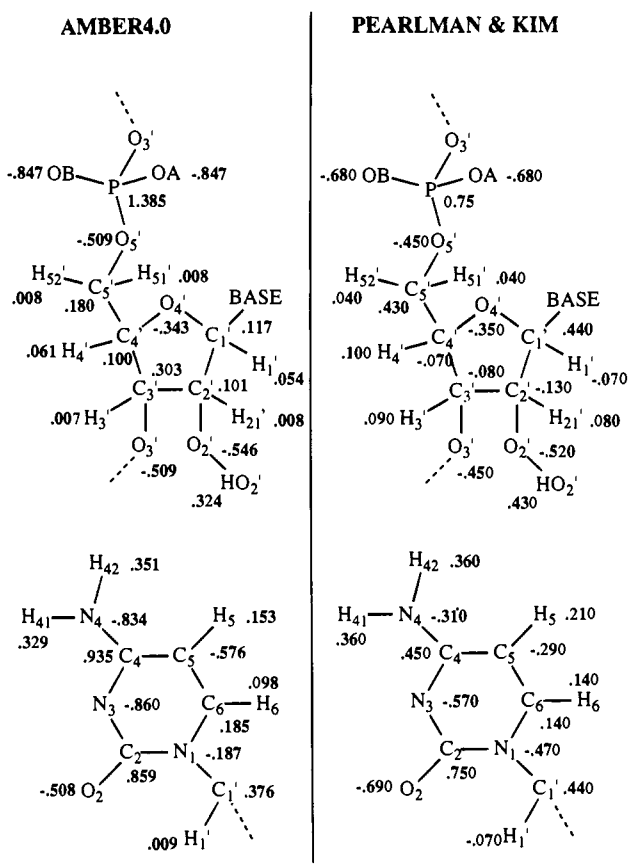

Figure 2. Ribose and cytosine charges from the AMBER4.0 force field[20] and from the Pearlman and Kim[16] set of charges.

(12) York, D. M.; Wlodawer, A.; Pedersen, L. G.; Darden, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8715−8718.

(13) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *97*, 6269−6271.

(14) Westhof, E.; Dumas, P.; Moras, D. *J. Mol. Biol.* **1985**, *184*, 119−145.

(15) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham III, T. E.; Ferguson, D. M.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. AMBER 4.1, San Francisco, CA, 1994.

(16) Pearlman, D. A.; Kim, S. H. *J. Mol. Biol.* **1990**, *211*, 171−187.

Table 1. Differences in Protocol between Simulations **S1**−**S10** and RMS Deviations Averaged over the Last 5 ps of the Simulation with Respect to the Crystal Structure[14] for All the Atoms of the Anticodon Loop

simulation	differences in protocol for simulations **S1**−**S10**	RMS (Å)
S1	reference	1.8
S2	different random seed from S1 at each change of temperature	2.3
S3	different random seed from S1 only at t = 12 ps (50 K restart)	2.2
S4	no restarts were applied	3.4
S5	different number of water molecules	2.5
S6	same as S1, but equilibration times were doubled	2.4
S7	same as S1, but use of 25 K steps to reach 300 K	3.8
S8	same as S7, but all the equilibration times were doubled	2.7
S9	same as S1, but using Amber4.0 with Kollman set of charges	3.1
S10	same as S1, but using Amber4.0 with Pearlman and Kim set of charges	2.1

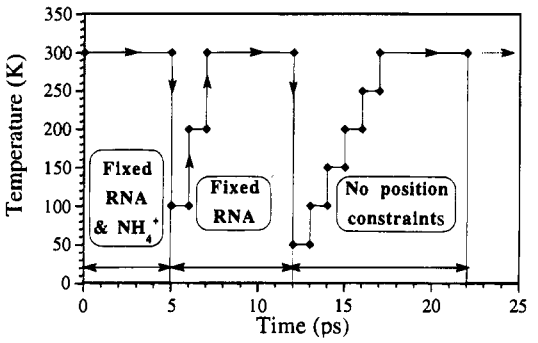

Figure 3. Schematic representation of the different phases in the equilibration protocol.

were adapted from those of the standard bases, and the NH_4^+ counterion charges were extracted from the work of Singh et al.[17]

The equilibration protocol (Figure 3) consisted of 100 steps of steepest descent minimization applied to the solvent molecules in order to relax possible steric clashes present at the interface of the RNA fragment and the solvent. This was followed by 5 ps of molecular dynamics at 300 K with fixed RNA and counterions positions where only the water molecules were allowed to reorient around the solute atoms. Next followed several picoseconds of molecular dynamics with fixed RNA atomic positions, mobile NH_4^+ counterions, and water molecules at temperatures of 100 K (1 ps), 200 K (1 ps), and 300 K (5 ps). In the following steps, no position constraints were applied to the system, and the temperature was progressively increased to 298 K in steps of 50 K with 1 ps of MD at each step. Finally, at 298 K, 5 ps of dynamics were run in order to allow the system to equilibrate at room temperature. The thermalization and equilibration phases of the MD protocol thus lasted 22 ps (Figure 3) for each simulation unless otherwise specified. At each temperature increase, the velocities were reassigned randomly according to a Boltzmann distribution. During thermalization, equilibration, and for the 100 ps of production that followed, the SHAKE algorithm[18] was used to constrain all the X−H bonds of the system and 10 kcal/mol harmonic distance constraints were applied to the three hydrogen bonds of the last GC base pair of the helix in order to prevent fraying of the stem (Figure 1). The MDdraw program[19] was used to visualize the generated trajectories on a Silicon Graphics workstation.

Eight simulations, denoted **S1**−**S8** (Table 1) were investigated in which systematic modifications were made in the equilibration protocol. Two additional simulations, **S9** and **S10** utilizing different electrostatic charge parameters, were also carried out in order to evaluate the effects associated with these force field changes. Simulation **S1** is chosen as the reference because it has the lowest calculated final RMS deviations from the crystal structure of the set (Table 1). **S2** differs from **S1** by a change in the random seed at every restart of the equilibration protocol which generates a different velocity distribution at each change of temperature. **S3** differs from **S1** only in the use of a different random seed at the velocity assignment at 50 K (t = 12 ps, Figure 3). In simulation **S4**, during the equilibration steps between 50 and 300 K, the velocities from each preceeding step were carried forward to the next step instead of being reassigned randomly from a Boltzmann distribution as in **S1**−**S3**. In simulation **S5**, the number of water molecules solvating the solute was increased by about 2% resulting in addition of 50 water molecules to the system. The effects of longer and more gentle thermalization protocols were investigated in simulations **S6**, **S7**, and **S8**. In **S6**, the length of each equilibration step at each temperature increment was doubled resulting in 44 ps equilibration instead of 22 ps (Figure 4). In **S7**, the temperature was increased by steps of 25 K, instead of 50 K, thus extending the length of the equilibration protocol from 22 to 28 ps. In simulation **S8**, the time of each equilibration step was doubled, and the temperature was increased in steps of 25 K resulting in a 56 ps equilibration phase. Simulations **S9** and **S10** were generated with the same starting coordinates and protocols as **S1** but with the earlier 4.0 version of AMBER. Additionally, **S9** utilized the standard AMBER 4.0 set of charges,[20] whereas simulations **S1**−**S8** and **S10** utilized the Pearlman and Kim set of charges.[16]

The temperature curves associated with the different equilibration protocols are shown in Figure 4. The trajectories were analyzed by calculating the time course of the RMS deviations from the crystal structure starting at the 50 K restart (Figure 5) and the RMS deviation for each base residue, i.e., base plus sugar in AMBER, calculated over the last 5 ps of the simulations (Figure 6).

Simulations S1−**S4—Effects of Velocity Redistributions.** The protocol modifications made in the thermalization steps of simulations **S1**−**S4** concern the effect of the velocity distribution at a change of temperature. In **S1**−**S3** the velocities were reassigned from a Boltzmann distribution at various temperature changes, through variations in the random seed (Table 1). In **S4** the velocities were not reassigned at each temperature increase. These protocol variations led to four distinct trajectories in the theoretically accessible phase space as evident in the time course plots of the RMS deviations (RMSD) from the crystal structure (Figure 5). Simulation **S1** stays the closest to the crystal structure in value and remains reasonably stable. The RMSD of **S2** increases throughout the simulation and for **S3** attains a high RMS value yet seems to begin to plateau toward

(17) Singh, U. C.; Brown, F. K.; Bash, P. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 1607−1614.

(18) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−336.

(19) Engler, E.; Wipff, G. MDdraw: a program to visualize MD trajectories, Université de Strasbourg, France, 1994.

(20) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230−252.

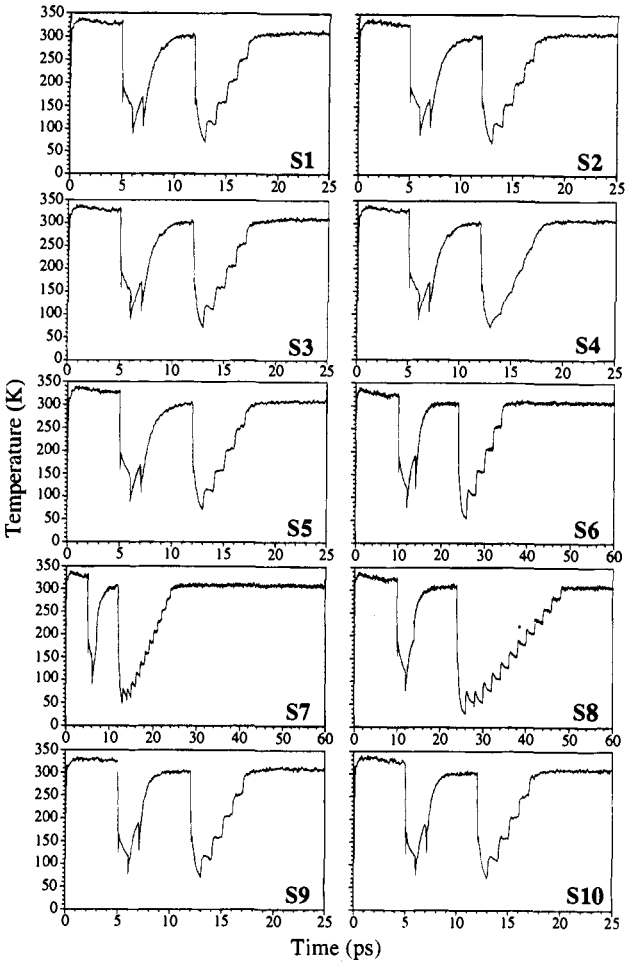

Figure 4. Curves representing the evolution of the temperature with time during the equilibration procedure (the different time scales for simulation **S6**, **S7**, and **S8**, result from different equilibration lengths).

the end of the trajectory. **S4**, in which there are no velocity redistributions, increases dramatically in RMSD value throughout the latter half of the trajectory. The total RMSD from the starting structure of these four simulations averaged over the last 5 ps of the trajectories are 1.8, 2.3, 2.2, and 3.4 Å for **S1**–**S4**, respectively (Table 1). The RMSD per residue averaged over the last 5 ps of the simulations shown in Figure 5 give indications of a different dynamical behavior of the anticodon loop in these four trajectories. There is an overall more jagged profile observed for **S4**, and we also find that the three anticodon bases G34, U35, and C36 show a higher deviation in **S3** and **S4** than in **S1** or **S2**.

From the comparison of these four runs, we were inclined to believe that the divergent trajectories were a result of a not well equilibrated system. This was suggested to us in particular by comparison of simulation **S4** with **S1**–**S3**. In **S4**, the dramatic increase in RMSD throughout the time course of the trajectory could be due to impulse forces present at the boundaries of the solute–solvent components of the system not repeatedly corrected for by uniform velocity reassignment of the entire system from a smooth statistical distribution. Thus, attempts were made to lengthen and smooth the thermalization protocol to allow further the dissipation of impulse forces, if present, and to coax the system more gently toward its preferred equilibrium state.

Simulations S6–S8—Extended Equilibration Protocols. The length of equilibration at each temperature change was increased in trajectory **S6**, the temperature increment at each step was reduced in trajectory **S7** and both these strategies were combined in trajectory **S8** (Table 1). The RMS deviations of

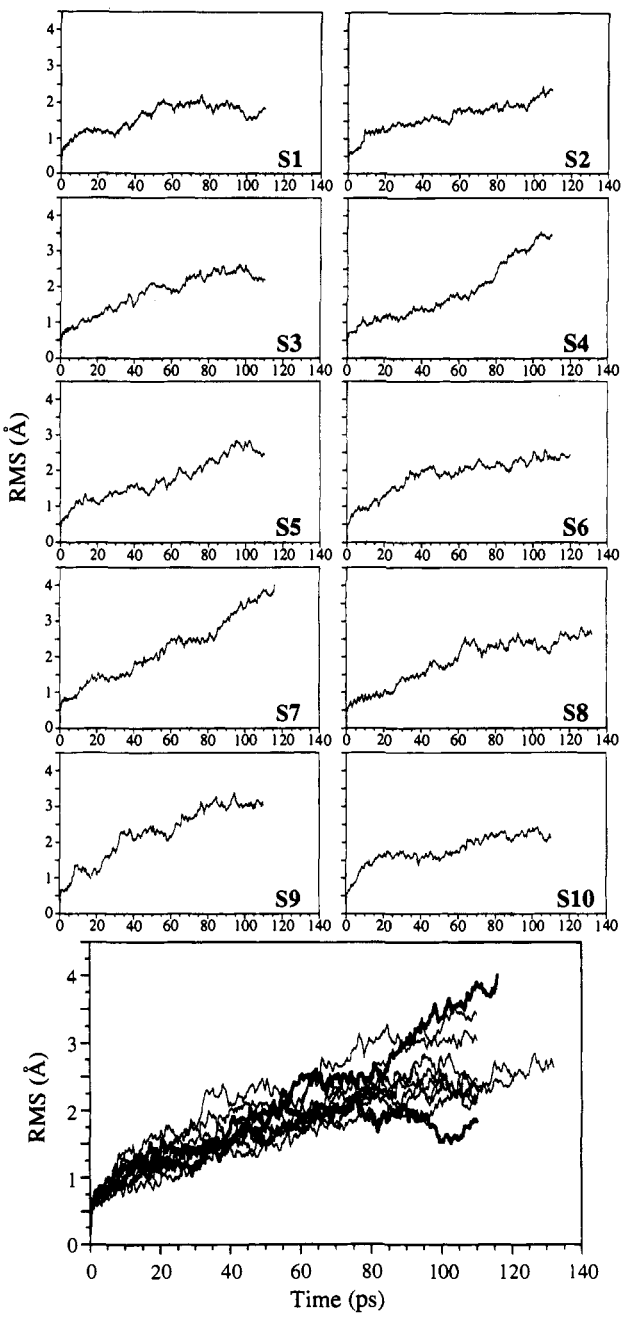

Figure 5. At the top, curves representing the total RMS deviations with respect to the crystal structure starting at the 50 K restart (the different time scales for simulation **S6**, **S7**, and **S8**, result from different equilibration lengths). At the bottom, superposition of the RMS curves for simulations **S1**–**S10**. **S1** and **S7**, with the lowest and highest final RMS deviations, are plotted in bold.

trajectories **S6** and **S8** are comparable to that calculated for **S2** and **S3**, still climbing to higher RMS values at the end of the calculated runs. The RMSD averaged over the last 5 ps were 2.4 and 2.7 Å, respectively. Simulation **S7**, however, shows dramatic increases in RMSD values, comparable in behavior to simulation **S4**, with a final RMSD value averaged over the last 5 ps of 3.8 Å. The per residue RMSD profiles for trajectories **S6**–**S8**, likewise match the respective characters of trajectories **S2**–**S4** discussed previously.

Simulations S9, S10, and S5—Change in Parameters or System. At the beginning of this series of simulations, our aim was to compare the results obtained with the set of charges of the AMBER4.0 package[20] to those obtained with the Pearlman and Kim set of charges.[16] Some of the main differences between these two sets can be appreciated from Figure 2. For example,

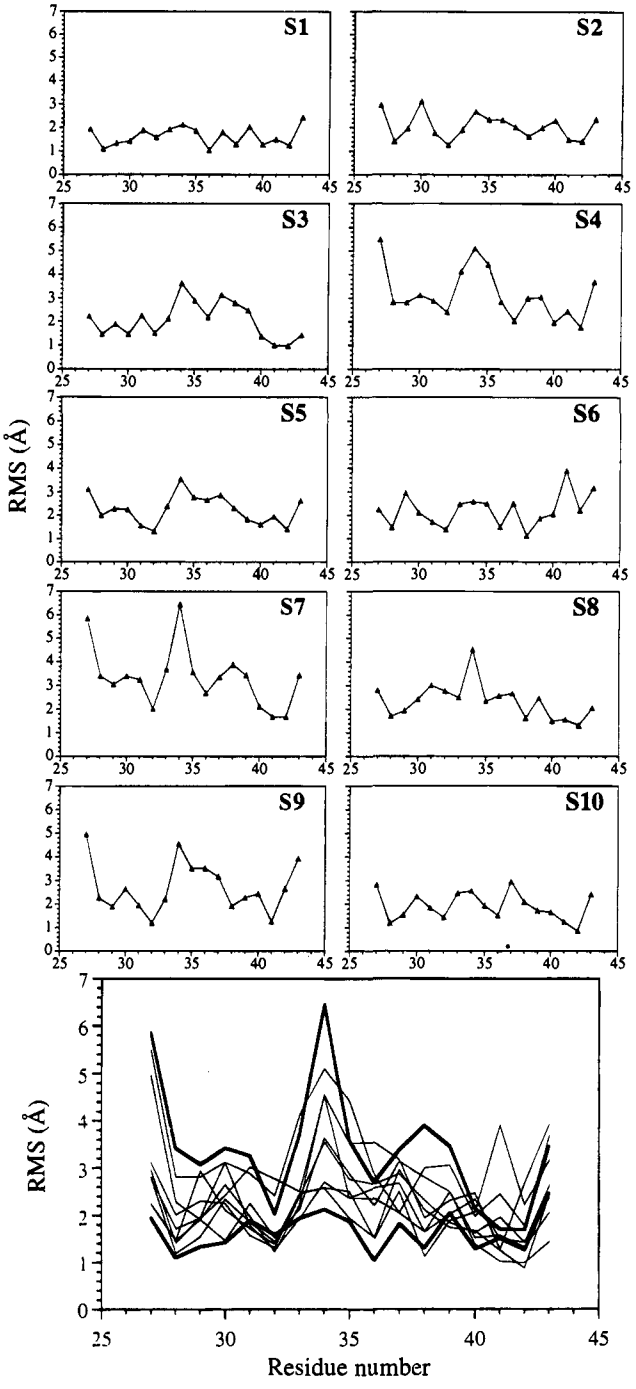

Figure 6. At the top, RMS deviations per residue for simulations **S1**–**S10**, averaged over the last 5 ps of the trajectories. At the bottom, superposition of all the RMS per residue curves. **S1** and **S7**, with the lowest and highest final RMS deviations, are plotted in bold.

the charge on the phosphate atoms are much larger for the AMBER4.0 than for the Pearlman and Kim set of parameters, and the $-NH_2$ groups carry an overall negative charge of -0.154 for AMBER4.0 while it carries a positive charge of $+0.410$ for the Pearlman and Kim set of parameters. An initial simulation, **S9**, conducted with the AMBER charges did not lead to satisfactory results. The structure was distorted, and the RMS deviations continued to increase markedly at the end of the 100 ps simulation. Thus, the experimentally derived Pearlman and Kim set of charges was investigated in simulation **S10**. This attempt led to much lower RMS deviations, 2.1 Å compared to 3.1 Å for **S10** and **S9**, respectively (Table 1). But

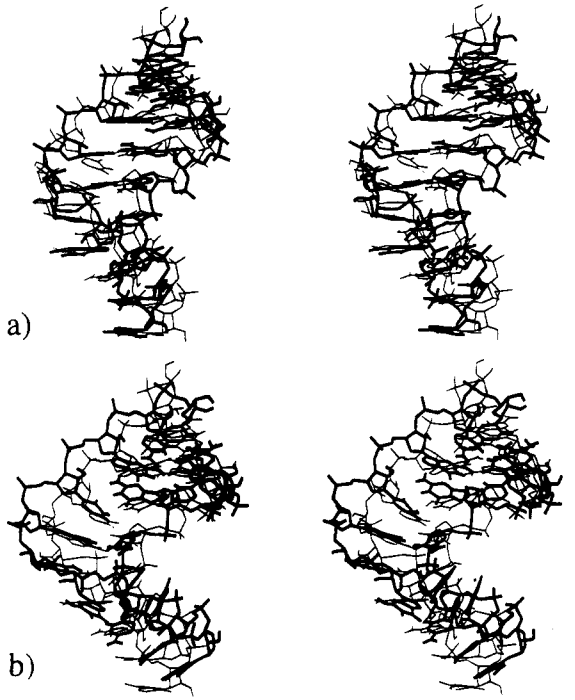

Figure 7. Stereoview of (a) the final structure of simulation **S1** and (b) the final structure of simulation **S7** superposed with the crystal structure. **S1** and **S7** plotted in bold.

additional endeavors to improve the MD protocol using the Pearlman and Kim set of charges and the updated 4.1 version of AMBER in simulations **S1**–**S4** and **S6**–**S8** led to divergent trajectories. The divergences within the set were found comparable to the differences between simulations using different force field parameters. As an additional test of the fidelity of our chosen protocol, the system itself was modified slightly in **S5** by increasing the number of water molecules solvating the solute by 2% (roughly 50 molecules) but without changing any other parameters from **S1**. The time course of the RMSD shows increasing values through to the end of the calculated trajectory similar to run **S8** and a per residue RMSD profile similar to run **S3**.

Figure 7 displays the last structure of the calculated trajectory superposed on the crystal starting structure for the best case (**S1**) and worst case (**S7**) trajectories of this set of simulations. The differences in structural characteristics of these two trajectories, generated from essentially the same simulation protocol, are obvious. Such a result shifted our focus from systematically trying to locate a stable protocol for our chosen system to acknowledging and addressing the ramifications of highly divergent trajectories resulting from slight modifications to the standard protocol that we had been following.

Discussion

In our initial aim to characterize the structural and dynamical characteristics of the solvated anticodon loop of $tRNA^{Asp}$ via solvated molecular dynamics trajectory using standard protocol methods, two related concerns guided our systematic search for a reliable simulation protocol: how to avoid structurally divergent trajectories and how to achieve an equilibrated system most expediently. After changing to the Pearlman and Kim set of charges, minor protocol variations were introduced into the system as perturbations, the results of which we hoped would help us to characterize the range of structural and dynamical properties of the system. Subsequently, thermalization steps were lengthened and smoothed in an attempt to remain closer

to the vicinity of the starting structure as the system approached equilibrium. Although there resulted at least one and possibly several trajectories that could be considered equilibrated or nearing equilibration, the fidelity of the protocol itself was called into question by the divergence of certain trajectories (**S4** and **S7**), especially when changes to the system (**S5**) or force field parameters (**S9**) produced trajectories statistically indistinguishable considering the set as a whole.

Using the nanosecond of MD trajectories generated for the purpose of addressing pertinent issues of critical evaluation of MD protocols and methods for detection of protocol instabilities or methodological artifacts seemed more warranted than a detailed structural report on the most optimistic of our trajectories. We have been able to identify several aspects of current MD methodology for which greater appreciation of the possible difficulties or limits encounterable is of general benefit. We have also demonstrated the utility of multiple MD trajectories in realizing a more critical and rigorous evaluation of present MD methodologies.

The RMSD Criteria. A plot of the time dependent root mean square deviation from the starting configuration of an MD trajectory yields a quick measure of the relative distance from the onset to which the system has evolved (its value) as well as its rate of evolution (its slope). When the rate of evolution of the system becomes negligible over a reasonable period of time, the system is considered to have reached equilibrium. When, in addition to equilibrium, the RMS deviation from the initial system configuration is low, it can be assumed that the trajectory reflects sampling in the vicinity of the starting structure with the advantages being that it is usually an experimentally derived reference. Thus the properties of the MD trajectory become useful in interpretation of the experimental results. However, when either of these conditions are not satisfied within the calculated trajectory, the relevance of properties deriving from the trajectory is not straightforward. Increasing RMSD values signal that the system has not achieved equilibrium, while high, yet stable RMSD values imply that one is not sampling the system in the vicinity of the experimental configuration. The RMSD criteria for evaluations of single MD trajectories of a system has limited value in that there is only one reference point, the static starting configuration. Multiple MD runs allow the advantage of a better appreciation of the range of dynamical behaviors and structural properties before the temptation to interpret certain structural transitions, which could be due to artifacts, manifests itself.

For simulations **S2**, **S4**, and **S5**−**S9** it is clear that the system is still evolving. In some cases, **S4**, **S7**, and **S9**, the rate of evolution from the initial configuration is quite rapid and the system would not be expected to return to the vicinity of the initial configuration within a reasonable MD time scale. Simulations **S1**, **S3**, and **S10** constitute good candidates for equilibrated or close to equilibrated trajectories within the current protocol with final RMSD values of 1.8, 2.2, and 2.2 Å and rates of evolution away from the initial system considerably reduced. However, the per residue RMSD profiles in Figure 5 for these three simulations show distinct characteristics which would lead to different interpretations of the structural properties and dynamical behaviors of the anticodon structure. For simulations **S1** and **S10**, the uniform deviation of all residues of the structure would lead one to emphasize that the stability of the loop region is commensurate with that of the stem, while interpretation of **S3** would prompt one to emphasize the increased mobility of the wobble base G34. Although these interpretations are not unreasonable, neither can be definitively confirmed nor ruled out by the experimental data.

Efforts Toward Achieving Equilibrium. In a truly ergodic space, no matter which distinct path the trajectory takes, after sufficient time the time averaged properties of different trajectories should be the same if the trajectory is sampling the system at equilibrium. In the construction of the molecular dynamics system, much care is given to trying to start as close to an equilibrium state as possible for each of the separate components; the molecule of interest, the mobile counterions, and solvent as well as for the ensemble. The molecule or system of interest is often taken from a crystal structure, where a heterogenous mixture of diverse molecules, salts, and solvents are present, however, many more components than can be included in the theoretical system. The positions of the mobile ions are not established by experimental means in other than descriptive sense and present a challenge to the object of constructing a near equilibrated system.[21] The solute, the molecule plus counterions, is usually immersed in a box of equilibrated bulk solvent to complete the system. The interdependence between stabilizing and destabilizing forces of the solute, mobile counterions and solvent is magnified in the initial part of the trajectory when the system energy has not yet settled, and impulse forces at the boundaries of the components may be present. If the initial configuration of the system is at an inherently unstable or high energy feature of the theoretical potential surface, the divergences of the initial trajectories may result in mutually exclusive regions of the potential surface being explored with extension of the trajectory because of the nonreversibility of structural barriers passed at an early high energy state of the system once the system locates regions of lower potential energy.

Multiple divergent trajectories, drifting far away from the initial structure, may signify a badly equilibrated system and possibly artifactual behavior. Our attempts to thermalize the system gently so as not to bump it from the region of the starting configuration did not lead to definitive improvement of the structural properties of the system or to an RMSD profile that indicated approaching equilibrium. Trajectory **S6** was prepared by doubling the dynamics equilibration time at each temperature increase throughout the thermalization period; however, this had no noticeable effect on the RMSD profile of the trajectory in comparison with **S1**, **S2**, and **S3**. Trajectory **S7**, prepared by halving the temperature increment, led to the most rapidly divergent structural trajectory. Employment of both of these methods in **S8** led to an intermediate result, neither demonstrably better nor worse than trajectories **S6** and **S7**.

Minor protocol changes, such as use of a different random seed in trajectories **S2** and **S3**, resulted in commensurate RMS deviation values as those resulting from larger changes, such as the use of a different version of the AMBER program (**S10**) or a different equilibration protocol (**S6**). Likewise, major structural divergences were encountered by modifications as large as the use of a different electrostatic charge set (**S9**) or as small as not using restarts in the thermalization steps (**S4**). Thus, pursuit of an extended trajectory for the purposes of sampling equilibrium properties was not thought prudent using the present protocol after considerations of its widely divergent trajectories resulting from systematic attempts to improve upon it.

In fact, a recent set of six trajectories generated in our laboratory on the same system with the same parameters used for **S1**, with the sole change of a *solute*−*solvent* and a *solvent*−*solvent* truncation distance of 16 Å instead of 8 Å,[22] revealed a better constancy of the structural results and an average lower

(21) Marlow, G. E.; Perkyns, J. S.; Pettitt, B. M. *Chem. Rev.* **1993**, *93*, 2503−2521.

(22) Auffinger, P.; Louise-May, S.; Westhof, E. Manuscript in preparation.

RMS deviation from the crystal structure from that calculated in the present work. The success of using more accurate treatments of the electrostatic long-range forces has been shown in several recent studies as also leading to trajectories remaining very close to the starting crystal structure.[9,12,23]

Can Different Parameter Sets Be Compared Using Single Trajectories? In the sampling of a multidimensional phase space which often shows extreme sensitivity to the initial state of the system, conclusions drawn from comparisons of two single trajectories may be misleading. In our ensemble of trajectories, comparison of trajectory **S1** with trajectory **S9** leads to the seemingly straightforward conclusion that the Pearlman and Kim set of charges is superior to the Amber set of charges, yet when trajectory **S4** or **S7** is added into consideration, such a conclusion cannot be definitively drawn. This result suggests the use of multiple MD runs for the evaluation of the respective qualities of different force fields, not on the basis of two single simulations, but in terms of comparison of average properties estimated from two separate sets of simulations. This will lead to the development of more reliable force fields and to an improved knowledge of their respective characteristics.

Detection of Artifacts. The trend toward the construction of increasingly more explicit and complex biological systems underlines the need for rapid testing of protocol variations as well as early detection of possible artifacts. As was shown in previous studies,[7,23-26] approximations of the long-range forces beyond a certain distance can lead to various artifacts manifesting themselves on various time scales. There is also reason to believe that certain artifactual behaviors related to long-range electrostatic forces may manifest their influence slowly, becoming apparent only after considerable computational investment has been made. Thus development of rigorous statistical analyses to test for the possible presence of artifacts early on in molecular dynamics simulations needs to be pursued. In the absence of rigorous tests, shorter multiple trajectories can be used to statistically sample up to a nanosecond of MD trajectory of the theoretical phase space in the vicinity of the starting configuration while allowing characterization of the effects of yet undetected perturbations. Our results suggest further that inconstancy of structural and dynamical properties across a set of trajectories could possibly be indicative of the presence of artifactual behaviors due to unstable protocols and therefore have to be analyzed and interpreted with great care.

(23) York, D. M.; Darden, T.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *99*, 8345-8348.
(24) Smith, P. E.; Pettitt, B. M. *J. Chem. Phys.* **1991**, *95*, 8430-8441.
(25) Schreiber, H.; Steinhauser, O. *J. Mol. Biol.* **1992**, *228*, 909-923.
(26) Auffinger, P.; Beveridge, D. L. *Chem. Phys. Lett.* **1995**, *234*, 413-415.

Conclusions

We report on a nanosecond of MD simulations on the solvated anticodon loop of $tRNA^{Asp}$ with the aim of addressing the ramifications of the differences observed in ten 100 ps molecular dynamics trajectories generated by modifications to the protocol of the system. We used the speed and economy of short multiple MD trajectories to systematically test various protocol variations. However, the constancy of the simulation results was not found sufficient to warrant extension or a detailed structural analysis of any of the present trajectories. We suggest the use of multiple MD runs, which minimize the possible accumulation of yet undiscovered methodological artifacts, as an alternative means of statistically sampling the available conformational space over longer single trajectories. Such a strategy helps in an estimation of the reliability of the results and affords a better estimate of the characteristics of the theoretical phase space defined by the parameters and simulation protocol used. Finally, it allows for a more complete evaluation of the dynamical and structural properties of the system of interest.

Acknowledgment. P.A. is supported by a fellowship from ORGANIBIO (28, rue Saint Dominique, Paris, France) in the program CM_2AO. S.L.-M. is grateful to the French government for providing a Chateaubriand fellowship. The authors aknowledge also the IDRIS computing center which provided computer time, and E.W. is thankful to the CEE for providing funds through the Biotech contract BIO 2-CT93-0345.

JA9500825