

# CHARMM Force Field Parameters for Simulation of Reactive Intermediates in Native and Thio-Substituted Ribozymes\*

EVELYN MAYAAN,<sup>1</sup> ADAM MOSER,<sup>1</sup> ALEXANDER D. MACKERELL, JR.,<sup>2</sup> DARRIN M. YORK<sup>1</sup>

<sup>1</sup>Department of Chemistry, University of Minnesota, 207 Pleasant St. SE,  
Minneapolis, Minnesota 55455-0431

<sup>2</sup>Department of Pharmaceutical Chemistry, School of Pharmacy, University  
of Maryland, Baltimore, Maryland 21201

Received 26 September 2005; Revised 28 January 2006; Accepted 7 February 2006

DOI 10.1002/jcc.20474

Published online 21 December 2006 in Wiley InterScience (www.interscience.wiley.com).

**Abstract:** Force field parameters specifically optimized for residues important in the study of RNA catalysis are derived from density-functional calculations, in a fashion consistent with the CHARMM27 all-atom empirical force field. Parameters are presented for residues that model reactive RNA intermediates and transition state analogs, thio-substituted phosphates and phosphoranes, and bound  $Mg^{2+}$  and di-metal bridge complexes. Target data was generated via density-functional calculations at the B3LYP/6-311++G(3df,2p)//B3LYP/6-31++G(d,p) level. Partial atomic charges were initially derived from CHelpG electrostatic potential fitting and subsequently adjusted to be consistent with the CHARMM27 charges. Lennard-Jones parameters were determined to reproduce interaction energies with water molecules. Bond, angle, and torsion parameters were derived from the density-functional calculations and renormalized to maintain compatibility with the existing CHARMM27 parameters for standard residues. The extension of the CHARMM27 force field parameters for the nonstandard biological residues presented here will have considerable use in simulations of ribozymes, including the study of freeze-trapped catalytic intermediates, metal ion binding and occupation, and thio effects.

© 2006 Wiley Periodicals, Inc. J Comput Chem 28: 495–507, 2007

**Key words:** molecular dynamics; RNA; metal ions; thio effect; phosphorane

## Introduction

Over the last several decades, a wealth of data has accumulated that demonstrates the central role RNA catalysis plays in many biological processes. Starting in the late 1970s, it was shown that RNA could catalyze fairly complex biological reactions in ribonuclease P<sup>1,2</sup> and Tetrahymena,<sup>3,4</sup> with an efficiency that rivaled many protein enzymes. These discoveries sparked a wave of interest in the scientific community focused on unraveling the details of how RNA enzymes (ribozymes) function. An understanding of the catalytic mechanisms of ribozymes, and their relation to sequence and tertiary structure, is opening up a variety of new frontiers. In biomedical technology, gene expression inhibitors that target viral and genetic diseases<sup>5</sup> such as HIV<sup>6</sup> and cancer<sup>7</sup> are being developed, and new biotechnologies such as RNA chips<sup>8</sup> and allosteric molecular switches in nanodevices<sup>9</sup> are being explored.

A common biological reaction catalyzed by several prototype ribozyme systems, such as the hammerhead,<sup>10,11</sup> hairpin,<sup>12,13</sup> and hepatitis delta virus<sup>14,15</sup> ribozymes, involves cleavage of a

phosphate group through a transesterification reaction<sup>16,17</sup> (Fig. 1). In this reaction, the 2'OH of the ribose sugar ring becomes activated via deprotonation and makes an in-line attack to the adjacent 3'-phosphate along the phosphodiester backbone. The attack produces a trigonal bipyramidal phosphorane intermediate/transition state that is accompanied by an inversion of configuration about the

\*This article contains supplementary material available via the Internet at <http://www.interscience.wiley.com/jpages/0192-8651/suppmat>.

**Correspondence to:** D. M. York; e-mail: york@chem.umn.edu

Contract/grant sponsor: National Institutes of Health; contract/grant numbers: GM62248, GM51501

Contract/grant sponsor: Army High Performance Computing Research Center (AHPARC)

Contract/grant sponsor: Department of the Army, Army Research Laboratory (ARL); contract/grant number: DAAD19-01-2-0014

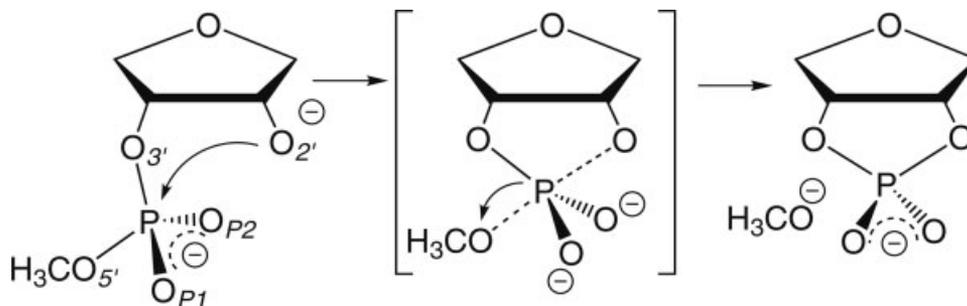


Figure 1. Model RNA transesterification reaction.

phosphorus center as the exocyclic P-O5' bond is cleaved. The product of the transesterification is a 5'-OH terminus and a 2', 3'-cyclic phosphate. Additional information about this mechanism has been obtained via kinetic isotope studies<sup>18–20</sup> and chemical modifications such as thio substitution<sup>21–23</sup> at the scissile phosphate, although the mechanistic interpretation of these experimental studies remains a topic of discussion and some debate. Time-resolved X-ray crystallography<sup>10,24,25</sup> has become a powerful tool to elucidate structural information at different stages along the catalytic reaction coordinate that provides valuable insight into ribozyme activity. However, this area is considerably challenging due to the difficulty of trapping a reactive intermediate and obtaining quality crystals, as well as uncertainties due to the nature of effects that arise from the crystallization conditions. For these reasons, there is considerable interest in the development of theoretical methods that can aid in the refinement and interpretation of existing experimental data, and provide structural insight into systems where such data is not yet available.

Molecular simulation, along with experimental structural data, provides an avenue for the characterization of ribozyme dynamics in solution and refinement of key mechanistic details. Molecular simulation force fields for nucleic acids continue to improve,<sup>26–31</sup> and a variety of simulations involving ribozymes have been carried out in recent years.<sup>32–40</sup> In order to study the structure and dynamics of different catalytic states along the reaction path of a ribozyme, however, reliable empirical force fields parameters must be developed for the transition states and reactive intermediates of these reactions. Furthermore, in order to use molecular simulation to aid in the interpretation of experimentally measured thio effects, parameters for thio-substituted phosphate and phosphorane models and their interactions with metal ions are required. In the present work, new force field parameters for residues important to the study of RNA catalysis are derived from density-functional calculations, to be consistent with the CHARMM27<sup>41</sup> all-atom empirical force field. These parameters will allow molecular dynamics (MD) simulations of ribozymes in reactive states to be performed, leading to a better understanding of the structure and dynamics of catalysis.

## Methods

The potential energy function used for the CHARMM27 empirical force field for nucleic acids,<sup>42,43</sup> and for the new modified RNA

residues of the present work, has the general form<sup>44</sup>:

$$\begin{aligned}
 U(\mathbf{r}_1, \mathbf{r}_2 \cdots \mathbf{r}_N) = & \sum_{\text{bonds}} K_b (b - b_0)^2 + \sum_{\text{angles}} K_\theta (\theta - \theta_0)^2 \\
 & + \sum_{\text{Urey-Bradley}} K_{UB} (S - S_0)^2 + \sum_{\text{impropers}} K_\varphi (\varphi - \varphi_0)^2 \\
 & + \sum_{\text{dihedrals}} K_\chi [1 + \cos(n\chi - \chi_0)] \\
 & + \sum_{i,j < i} \varepsilon_{ij} \left[ \left( \frac{R_{0,ij}}{r_{ij}} \right)^{12} - 2 \left( \frac{R_{0,ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{\varepsilon r_{ij}} \quad (1)
 \end{aligned}$$

The first four summations are quadratic terms that give rise to energy penalties for geometrical deviations about equilibrium coordinate values. The variables  $b$ ,  $\theta$ ,  $S$ , and  $\varphi$  are the bond length, bond angle, Urey–Bradley 1,3-distance, and improper torsion angle coordinates, respectively, while  $b_0$ ,  $\theta_0$ ,  $S_0$ ,  $\varphi_0$  and  $K_b$ ,  $K_\theta$ ,  $K_S$ ,  $K_\varphi$  are the corresponding force field parameters for the equilibrium geometries and force constants, respectively. The fifth summation is a trigonometric term that adjusts the periodic dihedral angle rotational barriers. The coordinate  $\chi$  is the dihedral coordinate,  $n$  determines the periodicity,  $\chi_0$  is a phase factor, and  $K_\chi$  is the amplitude force constant. The terms involving a sum over atom pairs  $i, j < i$  (neglecting nonbonded exclusions) are the nonbonded van der Waals/Lennard-Jones (L-J) and electrostatic terms. The parameters  $\varepsilon_{ij}$  and  $R_{0,ij}$  are the van der Waals well depth and minimum distance between the  $ij$  atom pair, respectively, and are by default calculated via the Lorentz–Berthelot combining rules<sup>45</sup> from the corresponding one-body parameters as  $\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$  and  $R_{0,ij} = R_{0,i} + R_{0,j}$ , respectively, although this default can be explicitly overridden using the nonbond fix (NBFIX) option, if fine tuning of specific pairwise interactions is desired. The last summation is the electrostatic energy determined between atomic partial charges  $q_i$ , which normally are calculated with a unit dielectric constant,  $\varepsilon = 1$ , as in the present work.

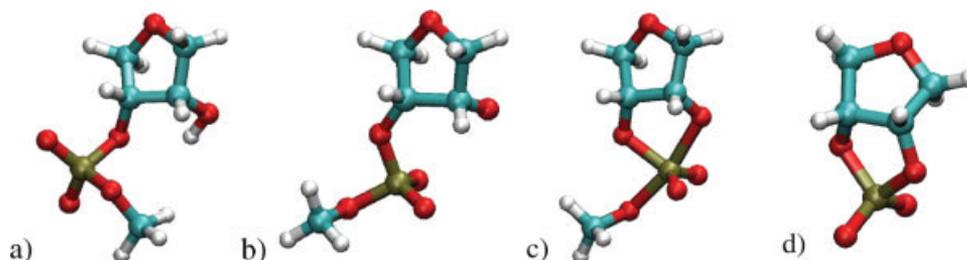
Parameterization of the CHARMM27 force field was based on *ab initio* and experimental data for small molecules,<sup>42</sup> as well as macromolecular simulation data of DNA and RNA<sup>46</sup> in order to more accurately capture experimental condensed phase properties. This was an important improvement to the CHARMM22 force field<sup>47</sup> that was parameterized more heavily on the basis of small

molecule target data alone. While this previous approach succeeded in capturing the target data of the selected small compounds, later simulations of duplex DNA in solution showed some disagreement with experiment in regard to the relative conformational stability of A-DNA versus B-DNA.<sup>48,49</sup>

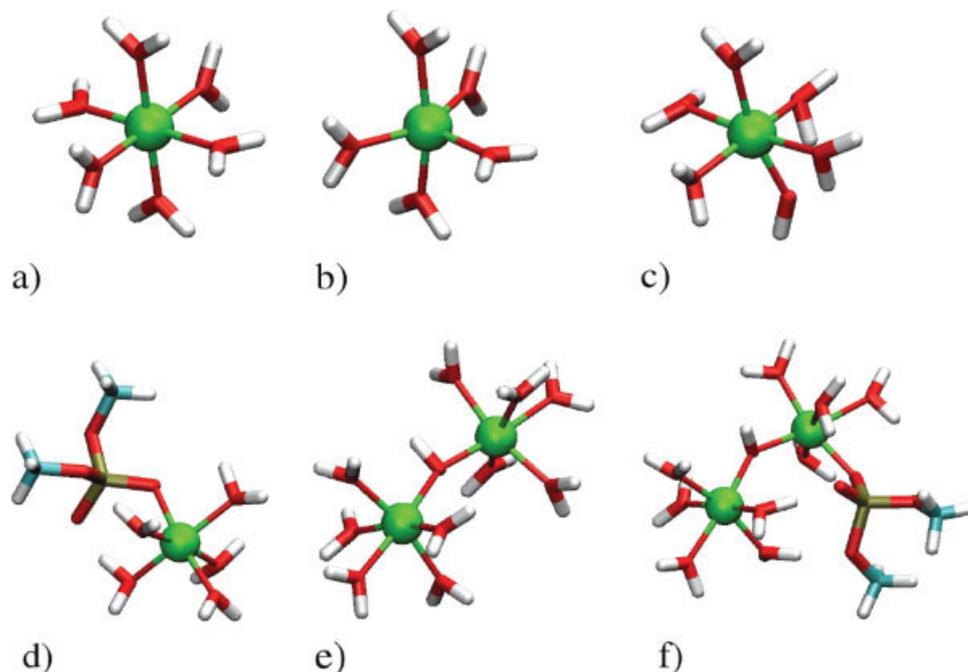
The CHARMM27 force field was developed using a self-consistent, step-wise strategy.<sup>31,42</sup> Partial atomic charges for the CHARMM27 force field were initially obtained from a Mulliken population analysis of the HF/6-31G(d) wave function. These were then adjusted to better reproduce scaled HF/6-31G(d) TIP3P<sup>50</sup> water interaction energies, experimental dipole moments, and heats of sublimation, where available. Equilibrium geometry parameters for the selected small molecules were initially optimized to reproduce the experimental geometries of microwave, electron diffraction, and X-ray crystal survey data, where possible. Equilibrium geometries and force constants were optimized iteratively until satisfactory fitting to the target data was achieved. Once this optimization was complete, iterative adjustment of the charges and L-J parameters was coupled with the internal parameters, until overall convergence was reached. A survey of RNA and DNA crystal structures from the nucleic acids database (NDB)<sup>51</sup> was taken, so that fitting to target macromolecular experimental properties, such as sugar puckering phase and dihedral angle distributions, could be made through crystal MD simulation. During this stage, in the original CHARMM27 all-atom empirical force field parameterization for nucleic acids,<sup>42</sup> dihedral angle parameters were adjusted to lower or raise energy barriers, such that target-condensed phase properties were better reproduced. While this sometimes sacrificed the quality of fitting to small molecule *ab initio* data, it accomplished the goal of more accurately reproducing experimental condensed phase properties. This macromolecular fitting was coupled with the internal parameter optimization, until satisfactory convergence was achieved.

Owing to the lack of high-resolution experimental data for RNA reactive intermediates and chemically modified nucleic acids, the parameterization of the present study is based solely on *ab initio* data, the optimized structures found in Figs. 2–5. Parameter optimization for modified CHARMM27 RNA,  $\text{Mg}^{2+}$ , and  $\text{OH}^-$  residues was based on density functional theory (DFT) calculations, for training sets of small molecules that represented the desired target

systems. Density-functional calculations were performed using the Becke three-parameter hybrid functional combined with the Lee–Yang–Parr exchange–correlation functional (B3LYP)<sup>52,53</sup> with the 6–31++G(d,p) basis set for geometry and frequency calculations followed by single-point electronic structure refinement with the 6–311++G(3df,2p) basis set, in a manner analogous to recent studies of biological phosphates.<sup>54–58</sup> The B3LYP model employed here neglects proper treatment of long-range dispersion interactions; however, these weak dispersion interactions are small in comparison with the highly polar (e.g., hydrogen bonded) and ionic interactions investigated here, for which B3LYP has been demonstrated previously to be generally reliable. All DFT calculations were performed using the GAUSSIAN03<sup>59</sup> package. Partial atomic charges were based on the CHarges from ELectrostatic Potentials using a Grid method (CHelpG)<sup>60,61</sup> which fits atom-centered point charges to the molecular electrostatic potential. The CHelpG charges were also constrained to reproduce molecular dipole moments. Similar approaches to force field charge determination have been applied and validated previously<sup>60,62</sup>. These calculations were performed with the 6–311++G(3df,2p) basis set, and the CHelpG charges were subsequently modified to be consistent with the original force field by taking into account only the charge differences from standard CHARMM residues, and making further adjustments for polar and nonpolar hydrogens and integer-charged groups, in accord with the original parameterization procedure<sup>31,42</sup> (see Results and Discussion section for complete details). As in the original CHARMM27 force field, transferring charges from the small model compounds to the nucleic acid fragments was accomplished by adding the charge of the removed hydrogen atom to the heavy atom from which it was deleted. Once charges were obtained, L-J parameters were determined by fitting CHARMM residue-TIP3P water interaction energies and distances to the density-functional target values. This was accomplished through scaling of the interaction energies and shifting of the binding distances, such that the DFT values matched those of the original CHARMM27 force field for unsubstituted dimethyl phosphate ( $\text{DMP}^-$ ). Although the original CHARMM27 force field parameterization fixed waters to the TIP3P geometry in *ab initio* calculations, it was found in this work that the relative differences in geometries predicted by the DFT level of



**Figure 2.** CHARMM27 standard and modified nucleotide residues; (a) A standard CHARMM27 ribonucleotide analog (no base) with the 2' position protonated; (b) 2' deprotonated ribose phosphate (i.e., activated at the 2' position); (c) ribose phosphorane (i.e., intermediate/transition state model); (d) 2', 3'-cyclic ribose phosphate (i.e., transesterification product). Color coding of atoms is as follows: carbon = turquoise, oxygen = red, phosphorus = brown, hydrogen = white. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]



**Figure 3.** Complexes used for  $\text{Mg}^{2+}$  and  $\text{OH}^-$  parameterization. (a)  $\text{Mg}(\text{H}_2\text{O})_6^{2+}$ ; (b)  $\text{Mg}(\text{H}_2\text{O})_5^{2+}$ ; (c)  $\text{Mg}(\text{H}_2\text{O})_5(\text{OH}^-)^{1+}$ ; (d)  $\text{Mg}(\text{H}_2\text{O})_5:\text{DMP}^{1+}$ ; (e)  $\text{Mg}(\text{H}_2\text{O})_5 \cdot (\text{OH}^-) \cdot \text{Mg}(\text{H}_2\text{O})_3^{3+}$ ; (f)  $\text{Mg}(\text{H}_2\text{O})_5 \cdot (\text{OH}^-) \cdot \text{Mg}(\text{H}_2\text{O})_4:\text{DMP}^{2+}$ . Color coding of atoms is as follows: magnesium = green, oxygen = red, carbon = turquoise, hydrogen = white. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

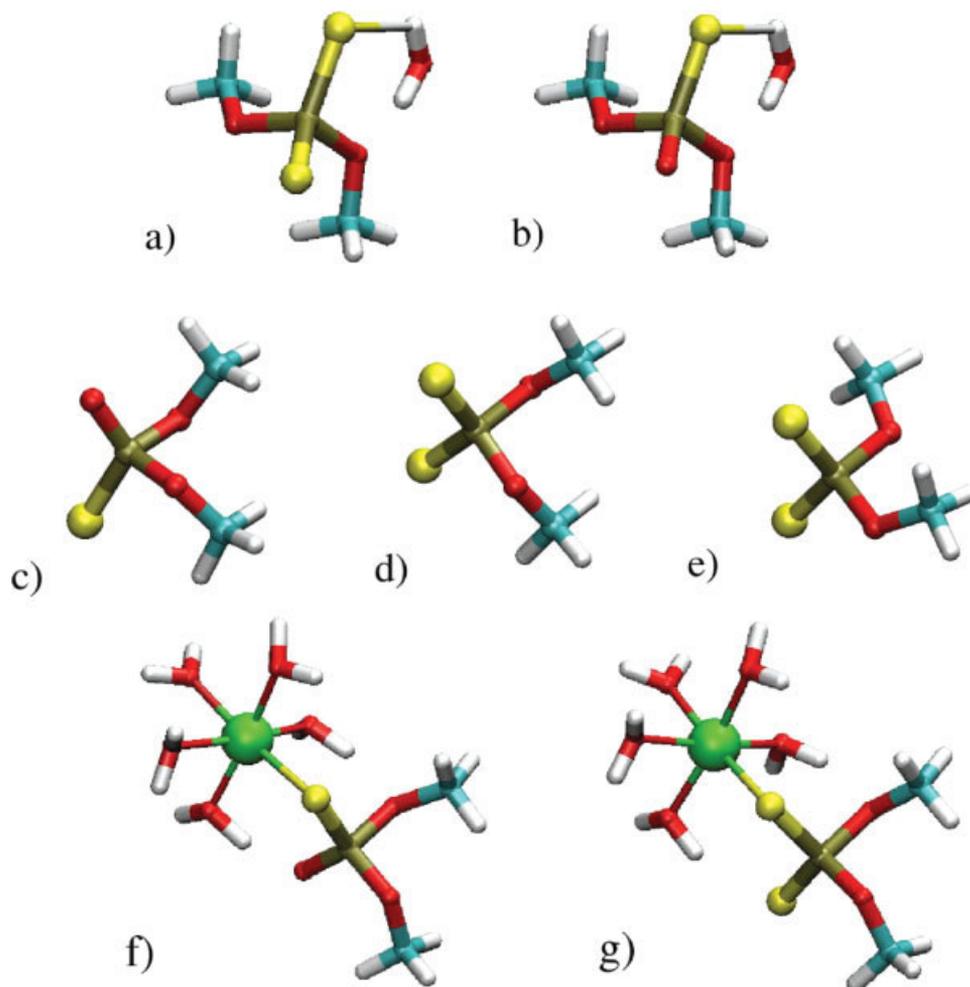
theory used compared better with the HF/6-31G(d) used in the original force field when the waters remained unconstrained (see Table S2 in Supporting Information). Internal parameters were fit to the density-functional geometries and energies for the model compounds.

The CHelpG charges used in the present work to obtain the charge differences with respect to similar standard CHARMM residues are different than the Mulliken charges that were used as a starting point for optimization in the original CHARMM27 force field. In general, the CHelpG charges are larger in magnitude, and due to the constraints, preserve the molecular dipole moments. The Mulliken charges on the other hand are considerably basis-set dependent, particularly when diffuse functions are used, and generally smaller in magnitude. It should be emphasized that the Mulliken charges used as a starting point for parameter optimization with CHARMM27 were often significantly altered upon optimization to obtain intermolecular interactions with TIP3P water molecules, the final charges in many cases being closer to the CHelpG charges (e.g., for the nonbridging phosphate oxygens).

RNA generally has more accessible “folded” conformations than does DNA<sup>63</sup> and often conformational deformation in the phosphate backbone is requisite for catalysis.<sup>64</sup> The CHARMM27 force field contains dihedral parameters for unsubstituted RNA based on *ab initio* gas-phase calculations of a variety of test compounds and adjusted with respect to A and B form DNA and RNA within the NDB.<sup>51</sup> These adjustments included the lowering of torsional energy barriers in certain regions after the observation that direct parameterization to the *ab initio* torsion profiles resulted in

simulated dihedral distributions that were inconsistent (e.g., too rigid) compared with those from a large NDB survey of RNA and DNA<sup>42</sup>. In the case of nonbridging thio substitutions, no CHARMM parameters exist. Therefore, attention was given to the C—O—P—O (i.e.,  $\alpha$  and  $\zeta$  in Fig. 6) dihedral parameters, because of their role along the phosphate backbone and the possibility of significant changes upon thio substitution. As in CHARMM27, DMP (dimethyl phosphate) was used as the model compound for native RNA as well as singly and doubly thio-substituted DMP, to determine the effects of the nonbridging sulfur(s) on this torsion. DMP and thio-substituted torsional potential energy surfaces (PESs) were generated using B3LYP/6-311++G(3df,2p)//B3LYP/6-31++G(d,p) with all degrees of freedom relaxed, except for the dihedral of interest and the symmetric  $\zeta$  dihedral, which is fixed in the trans conformation to induce symmetry and facilitate parameterization, see Figure 7.

As in the original force field parameterization,<sup>42</sup> a self-consistent step-wise optimization approach was taken that involved the iterative adjustments of L-J and internal parameters (not including torsion parameters), until convergence of the fitting function was obtained. Initial force constant values were taken from harmonic fitting to PES scans of bonds and angles, with distortions of 0.2 Å and 2.0°, respectively (see example in Fig. 8). Starting geometries for the new CHARMM\* model compounds were taken directly from the density-functional results and read into CHARMM to perform an adopted basis Newton–Raphson (ABNR) minimization,<sup>65</sup> until a gradient of  $<10^{-6}$  kcal/mol Å was reached. The results of the CHARMM minimization with the current parameters were used to



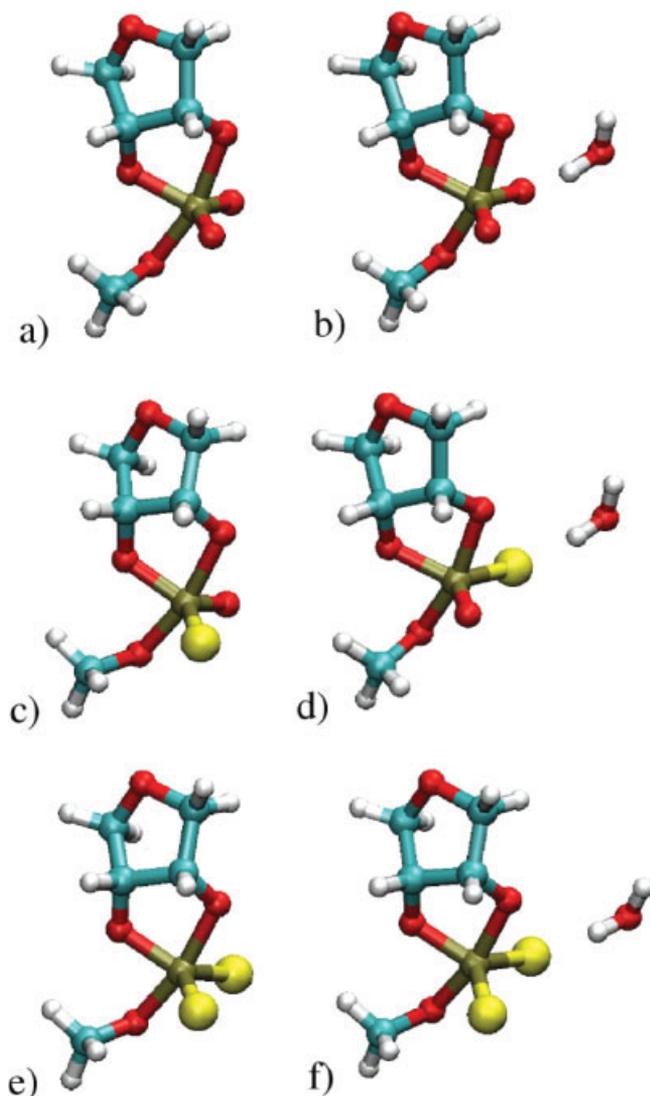
**Figure 4.** Complexes used for S and Mg<sup>2+</sup> parameterization. (a) DMP-ss<sup>-</sup>:HOH; (b) DMP-so<sup>-</sup>:HOH; (c) DMP-so<sup>-</sup>; (d) DMP-ss<sup>-</sup><sub>g-g</sub>; (e) DMP-ss<sup>-</sup><sub>g-r</sub>; (f) DMP-so<sup>-</sup>: [Mg(HOH)<sub>5</sub>]<sup>2+</sup>; (g) DMP-ss<sup>-</sup>: [Mg(HOH)<sub>5</sub>]<sup>2+</sup>. Color coding of atoms is as follows: magnesium = green, oxygen = red, carbon = turquoise, hydrogen = white, phosphorous = brown, sulfur = yellow.

construct a  $\chi^2$  fitting function. In the present work, this function took the form:

$$\chi^2(\lambda) = \sum_{k=1}^{\text{bonds}} \left( \frac{b_k^* - b_k^{\text{DFT}}}{\sigma_{b,k}^2} \right)^2 + \sum_{k=1}^{\text{angles}} \left( \frac{\theta_k^* - \theta_k^{\text{DFT}}}{\sigma_{\theta,k}^2} \right)^2 + \sum_{k=1}^{\text{torsions}} \left( \frac{\chi_k^* - \chi_k^{\text{DFT}}}{\sigma_{\chi,k}^2} \right)^2 + \sum_{k=1}^{\text{bind}} \left( \frac{\Delta E_k^* - \Delta E_k^{\text{DFT}}}{\sigma_{E,k}^2} \right)^2 \quad (2)$$

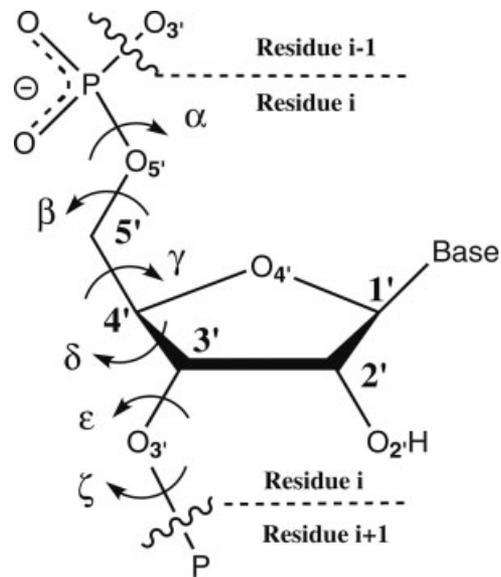
where the  $b_k^*$ ,  $\theta_k^*$ ,  $\chi_k^*$ , and  $\Delta E_k^*$  are the CHARMM energy-minimized bond lengths (Å), bond angles (degrees), torsion angles (degrees), and interaction energies (kcal/mol), respectively, and  $b_k^{\text{DFT}}$ ,  $\theta_k^{\text{DFT}}$ ,  $\chi_k^{\text{DFT}}$ , and  $\Delta E_k^{\text{DFT}}$  are the corresponding DFT reference values. The  $\sigma$  parameters, have the same units as their associated geometrical or energetic quantities (indicated by superscript), the inverse squares for which serve as weights in the unitless  $\chi^2(\lambda)$  merit function.

A nonlinear optimization procedure was used to minimize the  $\chi^2(\lambda)$  merit function with respect to the CHARMM\* parameters (this work) denoted by the vector  $\lambda$ . The  $\sigma$  parameters (and hence the weights) were adjusted empirically over the course of a stepwise minimization procedure, in order to drive the parameter search in the most realistic and robust region of parameter space (see Supporting Information for more details). In most cases, the last stepwise minimization procedure used to arrive at the final set of parameters was a quadratically convergent direction set minimization method<sup>66</sup> that does not require explicit gradient information. Sets of conjugate directions were generated with the algorithm due to Powell in a series of successive line minimizations achieved by parabolic interpolation. The convergence criterion was  $10^{-4}$  on the relative change in the function value, with respect to its minimum value after a series of line minimizations. While this method works very well for finding the local minima of a multivariable function, it is limited in its ability to find the absolute minimum of such a function. For this reason, the minimization procedure was carried out at a large



**Figure 5.** Complexes used for thio-substituted phosphorane parameterization. (a) ribose  $\text{MePA}^{2-}$ , (b) ribose  $\text{MePA}^{2-}:\text{HOH}$ , (c) ribose  $\text{MePA}\text{-so}^{2-}$ , (d) ribose  $\text{MePA}\text{-so}^{2-}:\text{HOH}$ , (e) ribose  $\text{MePA}\text{-ss}^{2-}$ , (f) ribose  $\text{MePA}\text{-ss}^{-}:\text{HOH}$ . Color coding of atoms is as follows: oxygen = red, carbon = turquoise, hydrogen = white, phosphorous = brown, sulfur = yellow. [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

number of starting points for each variable until the best overall minima could be found. Once a minimum was reached the torsion angle parameters were adjusted empirically to obtain a reasonable balance between the trends in the relative *ab initio* curves (native versus sulfur-substituted) while maintaining consistency with the standard CHARMM27 torsion profile, which does not correspond directly to an *ab initio* curve. Course-grained adjustment of the parameters (0.05 kcal) was seen to be sufficient to provide an acceptable balance. The torsion parameters were adjusted iteratively with the parameters of the nonlinear optimization. The complete topology and parameter files of this work are available as Supporting Information.



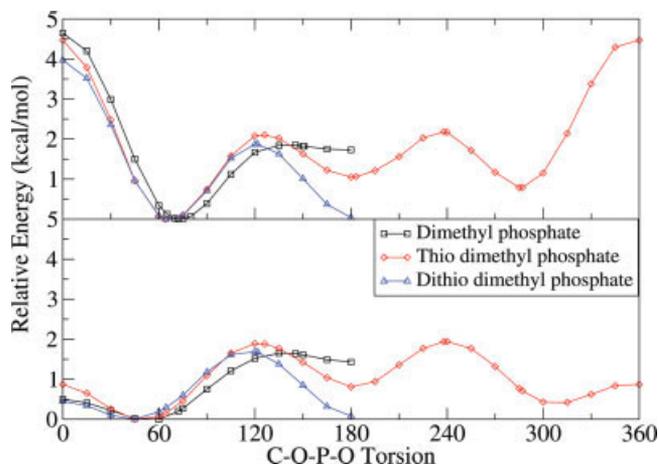
**Figure 6.** Phosphate backbone of RNA with torsions labeled.

## Results and Discussion

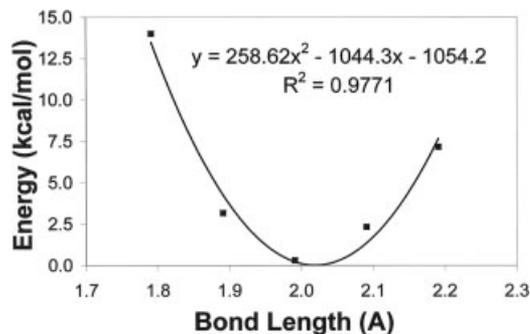
### Magnesium Complexes

RNA enzymes show diversity in the number as well as the role of metal ions they require. Metal ions can have both structural and chemical roles in catalysis. Possible (chemical) catalytic roles for metal ions include acting as:<sup>11,17</sup>

- General acid–base catalyst by accepting/donating protons from coordinated water molecules to the nucleophilic (2') and leaving group (5') positions



**Figure 7.** *Ab initio* (top frame) and CHARMM (bottom frame) torsional potential energy surface for the C–O–P–O dihedral of dimethyl phosphate (OO), nonbridging thio-substituted dimethyl phosphate (SO), and nonbridging dithio-substituted dimethyl phosphate (SS). [Color figure can be viewed in the online issue, which is available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]



**Figure 8.** Harmonic fitting to *ab initio* potential energy surface scan of thio-substituted dimethyl phosphate P–S bond.

- Lewis acid catalyst by direct coordination to either the nucleophilic (2') or leaving group (5') positions
- Electrophilic catalyst by coordinating to the nonbridging phosphoryl oxygens and assisting in formation of a phosphorane-like transition state or intermediate

One way for a metal ion to act as a general base is to lose a proton from a coordinated water to form a metal hydroxide. This  $\text{OH}^-$  has been postulated to play the role of abstracting a proton from a ribose 2'-OH,<sup>67</sup> although the  $\text{p}K_a$  difference between a phosphate-coordinated  $\text{Mg}^{2+}$  and a ribose 2'OH remains a topic of current interest.<sup>68</sup> In one postulated two metal ion mechanism for the hammerhead ribozyme (HR), it has been proposed that a metal hydroxide exists as a  $\mu$ -bridging  $\text{OH}^-$  between two magnesium ions.<sup>39</sup> Other studies suggest a single metal ion mechanism with an outer-sphere coordination of the  $\text{Mg}^{2+}$  ion to the 2'OH,<sup>69</sup> and still others suggest a nonbridging two metal ion mechanism with inner-sphere binding of  $\text{Mg}^{2+}$ .<sup>70</sup> Clearly, MD and QM/MM simulation could help resolve these mechanisms and the catalytic role played by  $\text{Mg}^{2+}$  ions.

The force field parameters for  $\text{Mg}^{2+}$  and  $\text{OH}^-$  are important to obtain accurate geometries and energies for  $\text{Mg}^{2+}$  ligand binding

and the formation of dimetal  $\text{Mg}^{2+}$  complexes containing a bridging hydroxide. Current nonbonded parameters exist for  $\text{Mg}^{2+}$  in the CHARMM27 force field<sup>46</sup> and  $\text{OH}^-$  parameters are available from Lee et al.<sup>71</sup> However, these parameters are not optimal for the specific problems that involve dimetal bridges or RNA-phosphate binding. Hence, the parameters for  $\text{OH}^-$  and  $\text{Mg}^{2+}$  have been re-optimized specifically for these types of systems.

DFT calculations have been run on several model systems<sup>55</sup> that formed the training set for the  $\text{Mg}^{2+}$  and  $\text{OH}^-$  parameterization. Each structure of the training set (Fig. 3) contained one or two  $\text{Mg}^{2+}$  ions, with a first coordination sphere combination of  $\text{H}_2\text{O}$ ,  $\text{OH}^-$ , or  $\text{DMP}^-$  ligands. In the structures with two  $\text{Mg}^{2+}$  ions present, a  $\text{OH}^-$  ion was placed in a  $\mu$ -bridging position between the two metal ions. CHelpG charges were calculated for  $\text{OH}^-$ , and the L-J parameters for  $\text{Mg}^{2+}$  and  $\text{OH}^-$  were iteratively adjusted until the best fit with geometry and water binding energy was obtained (Table 1).  $\text{Mg}^{2+}$  coordination distances were significantly improved with the new force field parameters for ligand binding with  $\text{H}_2\text{O}$ ,  $\text{OH}^-$ , and  $\text{DMP}^-$ . This improvement was most pronounced for the dimetal bridging hydroxide structures which lengthened the  $\text{Mg}^{2+} \cdots \text{OH}^-$  interaction by almost 0.30 Å and contract the  $\text{Mg}^{2+} \cdots \text{OH}^- \cdots \text{Mg}^{2+}$  angle by almost 10°. In the current study, only complexes that involve inner-sphere coordination of  $\text{Mg}^{2+}$  were used in the parameterization, and transferability to the important case of outer-sphere coordination remains to be tested with simulation.

### 2'O Deprotonated Sugar Ring

The first step of transesterification for a RNA phosphate is deprotonation of the sugar 2'OH. This creates a nucleophile of sufficient strength to attack the phosphorus of an adjacent or nearby phosphate group.<sup>16</sup> A "patch"<sup>42</sup> has been prepared for the standard CHARMM27 RNA nucleotide, which removes the 2'OH hydrogen and modifies the 2'O<sup>-</sup> charge, L-J interactions, and geometry of nearby atoms.

To derive new charges for a deprotonated nucleotide residue, a ribose methyl phosphate in its protonated and deprotonated states

**Table 1.**  $\text{Mg}^{2+}$  and  $\text{OH}^-$  Parameter Fitting Results.

|  | CHARMM <sup>1</sup> | DFT   | CHARMM*      |
|--|---------------------|-------|--------------|
| Bond (Å)   |                     |       |              |
| HT–OW  | 0.97 (0.00)         | 0.97  | 0.97 (0.00)  |
| MG $\cdots$ OH2  | 1.99 (–0.13)        | 2.12  | 2.13 (0.01)  |
| MG $\cdots$ OW   | 1.81 (–0.19)        | 2.00  | 2.00 (0.00)  |
| MG $\cdots$ ON3  | 1.86 (–0.17)        | 2.03  | 2.01 (–0.02) |
| MG $\cdots$ MG   | 3.56 (–0.33)        | 3.89  | 3.91 (0.02)  |
| Angle (°)  |                     |       |              |
| MG $\cdots$ OW $\cdots$ MG   | 155.8 (7.6)         | 148.2 | 152.7 (4.5)  |
| Energy (kcal/mol)  |                     |       |              |
| ( $\text{Mg}(\text{H}_2\text{O})_5 + \text{H}_2\text{O}$ ) $\rightarrow$ $\text{Mg}(\text{H}_2\text{O})_6$ | –27.4 (0.0)         | –27.4 | –28.3 (–0.9) |

<sup>1</sup>Training set (Figure 3) average errors for geometry and reaction energy fitting. With the exception of the hydroxide OW and HT atom types from Lee et al.,<sup>71</sup> the remaining parameters were originally from the standard CHARMM27 force field.<sup>42</sup> OH2 is the CHARMM27 water oxygen and ON3 is the non-bridging oxygen of  $\text{DMP}^-$ . See the Supporting Information for more data on the fitting function weight values used.

**Table 2.** CHelpG Charge Fitting for Deprotonated Ribose Phosphate.

| Atom | $q_0$ | $q_P$ | $q'_P$ | $q_D$ | $q'_D$ | $\delta q$ | $q_D^*$ |
|------|-------|-------|--------|-------|--------|------------|---------|
| C1'  | 0.07  | 0.19  | 0.00   | 0.31  | -0.19  | -0.19      | -0.12   |
| H1'  | 0.09  | 0.03  | 0.09   | -0.12 | 0.09   | 0.00       | 0.09    |
| H1'' | 0.09  | -0.04 | 0.09   | -0.20 | 0.09   | 0.00       | 0.09    |
| C2'  | 0.14  | 0.35  | 0.21   | 0.70  | 0.30   | 0.09       | 0.23    |
| H2'' | 0.09  | -0.05 | 0.09   | -0.31 | 0.09   | 0.00       | 0.09    |
| O2'  | -0.66 | -0.71 | -0.72  | -0.97 | -0.97  | -0.25      | -0.91   |
| H2'  | 0.43  | 0.42  | 0.43   | -     | -      | -0.43      | -       |
| C3'  | 0.01  | 0.30  | 0.20   | 0.65  | 0.37   | 0.17       | 0.18    |
| H3'  | 0.09  | -0.01 | 0.09   | -0.19 | 0.09   | 0.00       | 0.09    |
| O3'  | -0.57 | -0.57 | -0.57  | -0.63 | -0.63  | -0.06      | -0.63   |
| P    | 1.50  | 1.23  | 1.23   | 1.37  | 1.37   | 0.14       | 1.64    |
| O1P  | -0.78 | -0.77 | -0.77  | -0.84 | -0.83  | -0.06      | -0.84   |
| O2P  | -0.78 | -0.76 | -0.76  | -0.81 | -0.83  | -0.06      | -0.84   |
| O3P  | -0.57 | -0.50 | -0.50  | -0.58 | -0.58  | -0.08      | -0.65   |
| C3T  | -0.17 | 0.31  | -0.10  | 0.36  | -0.16  | -0.06      | -0.23   |
| H3T1 | 0.09  | -0.06 | 0.09   | -0.08 | 0.09   | 0.00       | 0.09    |
| H3T2 | 0.09  | -0.05 | 0.09   | -0.08 | 0.09   | 0.00       | 0.09    |
| H3T3 | 0.09  | -0.03 | 0.09   | -0.09 | 0.09   | 0.00       | 0.09    |
| C4'  | 0.07  | 0.27  | 0.06   | 0.14  | -0.12  | -0.18      | -0.11   |
| H4'  | 0.09  | -0.03 | 0.09   | -0.07 | 0.09   | 0.00       | 0.09    |
| H4'' | 0.09  | 0.00  | 0.09   | -0.01 | 0.09   | 0.00       | 0.09    |
| O4'  | -0.50 | -0.52 | -0.52  | -0.55 | -0.55  | -0.03      | -0.53   |

Adjusted charges ( $q'_P/q'_D$ ) were determined by changing all CHelpG hydrogen charges to be consistent with CHARMM27 by the procedure outlined in eqs. (3) and (4) of the text. This difference between the CHARMM27 ( $q_0$ ) and CHelpG charges ( $q_P/q_D$ ) was then added into the nearest heavy atom charge. Once the adjusted atomic charge differences between the protonated and deprotonated structures were determined ( $\delta q$ ), these differences were then added to the standard CHARMM27 charges to shift for deprotonation ( $q_D^*$ ).

was optimized using DFT. It was found that an “in-line” geometry minimum did not exist in the gas phase without the ribozyme environment. In order to parameterize a deprotonated ribose methyl phosphate (ribose MeP<sup>-</sup>) residue to a structure that models the in-line phosphate attack conformation, a minimization was performed with the corresponding  $\alpha$ ,  $\epsilon$ , and  $\zeta$  dihedrals of the phosphate-sugar backbone fixed to the minimum energy geometry found for the phosphorane structure, as described in Results and Discussion section (Fig. 2). In this way, the backbone of the ribose MeP<sup>-</sup> was constrained to be “in-line”. CHelpG charges, constrained to reproduce molecular dipole moment were then calculated for the atoms of the optimized structures (Table 2). Since the original CHARMM27

charges ( $q_0$ ) were obtained via a slightly different procedure than those of the present DFT/CHelpG method ( $q$ ), adjustments to the CHelpG charges ( $q'$ ) were made to create charges more consistent with the constrained CHARMM27 force field parameterization. For instance, in the CHARMM27 force field, all nonpolar hydrogens are constrained to have a charge of 0.09  $e$ , while polar hydroxyl hydrogens are assigned a charge of 0.43  $e$ . The calculated CHelpG charges were adjusted for this by changing the value of the hydrogens to be consistent with the CHARMM27 force field while adding the difference created from this adjustment to the adjoining heavy atoms (Table 2). The differences between the adjusted CHelpG charges for the protonated and deprotonated

**Table 3.** RNA L-J Parameters.

| Parameter           | Protonated methyl ribose | Deprotonated methyl ribose | Phosphorane methyl ribose | 2',3'-cyclic methyl ribose |
|---------------------|--------------------------|----------------------------|---------------------------|----------------------------|
| $R_{\min}/2O_{2'}$  | 1.77                     | 1.75                       | 1.76                      | 1.77                       |
| $\epsilon_{O_{2'}}$ | -0.1521                  | -0.3236                    | -0.2378                   | -0.1521                    |

L-J parameters for the deprotonated ribose phosphate were fit to the water binding coordination distance (1.713 Å) and the water binding coordination energy (18.4 kcal/mol) with TIP3P water (see Table 4 as well). L-J parameters for axial phosphorane oxygens (bond order  $\approx$  0.5) were determined by averaging between the protonated (bond order  $\approx$  0.0) and 2',3'-cyclic (bond order  $\approx$  1.0) structures.

**Table 4.** Geometry Fitting Results of Deprotonated Ribose Phosphate.

|  | Protonated<br>DFT | Deprotonated<br>DFT | Relative<br>Difference | Protonated<br>CHARMM | Adjusted<br>CHARMM* |             |
|--|-------------------|---------------------|------------------------|----------------------|---------------------|-------------|
| Bond (Å)   |                   |                     |                        |                      |                     |             |
| O <sub>2</sub> '-C <sub>2</sub> '  | 1.43              | 1.33                | -0.10                  | 1.43                 | 1.33 (0.00)         |             |
| O <sub>2</sub> '-C <sub>2</sub> '-H <sub>2</sub> '   | 109.6             | 115.8               | 6.2                    | 109.9                | 116.1 (0.0)         |             |
| O <sub>2</sub> '-C <sub>2</sub> '-C <sub>1</sub> '   | 110.2             | 112.7               | 2.5                    | 114.0                | 116.5 (0.0)         |             |
| O <sub>2</sub> '-C <sub>2</sub> '-C <sub>3</sub> '   | 109.8             | 116.4               | 6.6                    | 112.2                | 118.8 (0.0)         |             |
| Coordination Distance  |                   |                     |                        |                      |                     |             |
| Coordination Energy  |                   |                     |                        |                      |                     |             |
| Reaction   | DFT               |                     | CHARMM*                | DFT                  |                     | CHARMM*     |
| ribose MeP <sup>-</sup> + H <sub>2</sub> O →<br>ribose MeP <sup>-</sup> :H <sub>2</sub> O (kcal/mol) | 1.71              |                     | 1.70 (0.00)            | -18.4                |                     | -18.4 (0.0) |

Ribose MeP<sup>-</sup> stands for the deprotonated ribose methyl phosphate. Final fitting for the adjusted CHARMM\* residue was performed within the RNA sequence active site taken from the “early intermediate” hammerhead ribozyme X-ray crystal structure.<sup>10</sup>

ribose methyl phosphate were then calculated and used to modify the original CHARMM27 charges of a neutral nucleotide as follows:

$$\delta q = q'_P - q'_D \quad (3)$$

$$q_D^* = q_o + \delta q \quad (4)$$

where  $q'_P$  and  $q'_D$  are the adjusted DFT/CHelpG charges for the protonated and deprotonated structures, respectively,  $q_o$  are the CHARMM27 charges, and  $q_D^*$  are the new charges for the deprotonated structure. Final partial atomic charge parameters are shown in Table 2. It should be noted that constraints on the charge fitting to create unit charge groups within residues was not used in this method.

**Table 5.** CHelpG Charge Fitting for Ribose Phosphorane.

| Atom | $q_D^*$ | $q_D$ | $q'_D$ | $q_{TS}$ | $q'_{TS}$ | $\delta q$ | $q_{TS}^*$ |
|------|---------|-------|--------|----------|-----------|------------|------------|
| C1'  | -0.12   | 0.31  | -0.19  | 0.21     | -0.09     | 0.10       | -0.02      |
| H1'  | 0.09    | -0.12 | 0.09   | -0.10    | 0.09      | 0.00       | 0.09       |
| H1'' | 0.09    | -0.20 | 0.09   | -0.02    | 0.09      | 0.00       | 0.09       |
| C2'  | 0.23    | 0.70  | 0.30   | 0.42     | 0.17      | -0.13      | 0.10       |
| H2'' | 0.09    | -0.31 | 0.09   | -0.16    | 0.09      | 0.00       | 0.09       |
| O2'  | -0.91   | -0.97 | -0.97  | -0.74    | -0.74     | 0.23       | -0.68      |
| C3'  | 0.18    | 0.65  | 0.37   | 0.54     | 0.29      | -0.08      | 0.10       |
| H3'  | 0.09    | -0.19 | 0.09   | -0.16    | 0.09      | 0.00       | 0.09       |
| O3'  | -0.63   | -0.63 | -0.63  | -0.62    | -0.62     | 0.01       | -0.62      |
| P    | 1.64    | 1.37  | 1.37   | 1.54     | 1.54      | 0.17       | 1.81       |
| O1P  | -0.84   | -0.84 | -0.83  | -0.88    | -0.90     | -0.08      | -0.92      |
| O2P  | -0.84   | -0.81 | -0.83  | -0.92    | -0.90     | -0.08      | -0.92      |
| O3P  | -0.65   | -0.58 | -0.58  | -0.68    | -0.68     | -0.10      | -0.75      |
| C3T  | -0.23   | 0.36  | -0.16  | 0.54     | -0.21     | -0.05      | -0.28      |
| H3T1 | 0.09    | -0.08 | 0.09   | -0.14    | 0.09      | 0.00       | 0.09       |
| H3T2 | 0.09    | -0.08 | 0.09   | -0.16    | 0.09      | 0.00       | 0.09       |
| H3T3 | 0.09    | -0.09 | 0.09   | -0.18    | 0.09      | 0.00       | 0.09       |
| C4'  | -0.11   | 0.14  | -0.12  | 0.02     | -0.14     | -0.02      | -0.13      |
| H4'  | 0.09    | -0.07 | 0.09   | -0.04    | 0.09      | 0.00       | 0.09       |
| H4'' | 0.09    | -0.01 | 0.09   | 0.02     | 0.09      | 0.00       | 0.09       |
| O4'  | -0.53   | -0.55 | -0.55  | -0.53    | -0.53     | 0.02       | -0.51      |

CHelpG adjusted charges for determining ribose phosphorane partial atomic charges. New charges were created by taking the adjusted CHelpG charge difference  $\delta q$  between the ribose MeP<sup>-</sup> charges calculated in Table 2 and the ribose phosphorane.  $q_{TS}$  is the ribose phosphorane CHelpG charges,  $q'_{TS}$  the adjusted CHelpG charges and,  $q_{TS}^*$  are the final CHARMM\* charges. See eqs. (3) and (4) of text for further details.

**Table 6.** Geometry Fitting Results for Ribose Phosphorane Parameterization.

|  | DFT   | CHARMM*      |
|--|-------|--------------|
| Bond (Å)                                       |       |              |
| C <sub>2</sub> –O <sub>2</sub>                 | 1.361 | 1.361 (0.00) |
| P–O <sub>2</sub>                               | 1.99  | 1.99 (0.00)  |
| P–O <sub>3</sub>                               | 1.79  | 1.79 (0.00)  |
| P–O <sub>5</sub>                               | 1.84  | 1.84 (0.00)  |
| P–O <sub>R/S</sub>                             | 1.53  | 1.53 (0.00)  |
| O <sub>2</sub> –O <sub>5</sub>                 | 3.78  | 3.79 (0.01)  |
| Angle (°)                                      |       |              |
| C <sub>1</sub> –C <sub>2</sub> –O <sub>2</sub> | 113.5 | 113.7 (0.2)  |
| C <sub>3</sub> –C <sub>2</sub> –O <sub>2</sub> | 105.9 | 105.9 (0.0)  |
| C <sub>2</sub> –O <sub>2</sub> –P              | 109.3 | 109.3 (0.0)  |
| C <sub>3</sub> –O <sub>3</sub> –P              | 117.5 | 117.5 (0.0)  |
| O <sub>2</sub> –P–O <sub>5</sub>               | 163.5 | 164.0 (0.5)  |
| O <sub>R</sub> –P–O <sub>S</sub>               | 129.1 | 129.1 (0.0)  |
| O <sub>3</sub> –P–O <sub>R/S</sub>             | 115.3 | 115.3 (0.0)  |
| O <sub>2</sub> –P–O <sub>R/S</sub>             | 91.4  | 91.4 (0.0)   |
| O <sub>5</sub> –P–O <sub>R/S</sub>             | 95.4  | 95.3 (0.1)   |
| O <sub>2</sub> –P–O <sub>3</sub>               | 81.0  | 81.0 (0.0)   |
| O <sub>3</sub> –P–O <sub>5</sub>               | 83.3  | 83.3 (0.0)   |

Final fitting for the adjusted CHARMM\* residue was performed within the RNA sequence active site taken from the “late intermediate” hammerhead ribozyme X-ray crystal structure.<sup>74</sup>

Once charges were calculated for the deprotonated ribose 2'O<sup>-</sup> oxygen, L-J parameters were determined by fitting to the 2'O<sup>-</sup> water coordination distance (1.713 Å) and interaction energy (-18.4 kcal/mol) obtained from DFT optimization of a ribose MeP<sup>-</sup> (Table 3). Bond and angle parameters were fit to the differences between the gas-phase DFT geometry of protonated and deprotonated ribose MeP<sup>-</sup> (Table 4). Force constants were predicted by fitting to relaxed PES scans. All dihedrals, except for H3'-C3'-C2'-O2', were set to zero in the original protonated CHARMM27 structure and remained so in the new parameterization. The H3'-C3'-C2'-O2', which has a  $K_{\chi}$  of 0.195 kcal/mol/radian<sup>2</sup> and 0.0° phase for the threefold term, was retained. Final fitting was performed for the residue patch within the RNA sequence UCA taken from the “early intermediate” active site geometry in the X-ray crystal structure of a HR.<sup>10</sup> ABNR minimization was performed for this sequence with the U and A residues fixed and, where necessary to improve fitting, small adjustments to the internal parameters were made manually. The final results (Table 4) were able to reproduce the DFT structure and water binding energy very closely.

#### Phosphate Transition State Analog

The deprotonated ribose 2'O<sup>-</sup> attack at the phosphate produces a phosphorane that is a transition state in the gas phase (Fig. 2). Partial charges for the phosphorane atoms were determined similar to those for the 2'O<sup>-</sup> ribose MeP<sup>-</sup> described earlier. However, owing to the instability of dianionic phosphorane in the gas phase, DFT optimization was carried out with the axial P-O2' bond length fixed to the P-O2' bond length of a 2'O<sup>-</sup> ribose methyl phosphorane (ribose MePA<sup>2-</sup>) optimized in the aqueous phase (1.986 Å) using the polarizable continuum model solvation model.<sup>72,73</sup> Comparison

of similar neutral and anionic phosphorane structures optimized in both the gas and aqueous phases suggests that the error for this constraint should be less than 0.02 Å for all but the axial P–O bonds may have errors as high as 0.07 Å. Calculated CHelpG charges for hydrogen atoms were adjusted to 0.09 *e*, with heavy atoms taking up the difference as before (Table 5). Final charges ( $q_{\text{TSG}}^*$ ) for the phosphorane structure were determined by adding the adjusted CHelpG charge difference ( $\delta q$ ) between the ribose methyl phosphorane and the 2'O<sup>-</sup> ribose MeP<sup>-</sup> to the previously determined CHARMM\* charges ( $q_{\text{D}}^*$ ) for the the 2'O<sup>-</sup> ribose MeP<sup>-</sup> (Table 2).

Bond length and angle equilibrium parameters for the ribose MePA<sup>2-</sup> were determined from fitting to the gas phase DFT optimization of the partially frozen dianionic ribose MePA<sup>2-</sup> structure described previously. Force constants for bonds and angles were calculated by fitting to relaxed PES scans for each P–O bond and O–P–O angle of a monoanionic phosphorane (which is stable in the gas phase). The final ribose MePA<sup>2-</sup> residue was fit within the RNA sequence UCA taken from the near in-line geometry in the X-ray crystal structure of a HR “late-intermediate.”<sup>74</sup> ABNR minimization was performed for this sequence with the U and A residues fixed and, where necessary, small adjustments to the internal parameters were made to improve fitting. Because the penta-coordinated geometry of the phosphorane is quite rigid, the dihedrals had negligible effect, and hence were set to zero. The final fitting results matched with the DFT results almost exactly (Table 6).

#### 2',3'-Cyclic Phosphate

Transesterification terminates in the creation of a 2',3'-cyclic phosphate after exocyclic P-O5' bond cleavage. The 2',3'-cyclic phosphate is similar to a standard RNA ribose methyl phosphate residue, except for the cyclization of the O2'-P-O3'-C3'-C2' atoms (Fig. 2 and the parameter file in the Supporting Information). In the cyclic residue, the O2' is bonded to the phosphorus in the same manner as the O3'. Owing to the near symmetric equivalence between the O2' and O3' cyclic ribose phosphate sequences, the bond length, angle, and dihedral parameters are assumed to be identical. Therefore, O2' was assigned the same atom type as O3' (type

**Table 7.** Geometry Fitting Results for 2',3'-Cyclic Phosphate Parameterization.

|              | DFT   | CHARMM*     |
|--------------|-------|-------------|
| Bond (Å)     |       |             |
| O2'–P        | 1.71  | 1.71 (0.00) |
| O3'–P        | 1.71  | 1.71 (0.00) |
| P–O1P        | 1.50  | 1.50 (0.00) |
| P–O2P        | 1.50  | 1.50 (0.00) |
| Angle (°)    |       |             |
| C2'–O2'–PC   | 109.4 | 109.4 (0.0) |
| C3'–O3'–PC   | 109.4 | 109.4 (0.0) |
| O2'–PC–O3'   | 91.0  | 91.0 (0.0)  |
| O2'–PC–O1/2C | 109.0 | 109.0 (0.0) |
| O3'–PC–O1/2C | 109.0 | 109.0 (0.0) |
| O1C–PC–O2C   | 124.5 | 124.5 (0.0) |

Training set average geometries for geometry fitting of 2',3'-cyclic phosphate.

**Table 8.** Geometry and Binding Energy Results for Phosphate Thio-Substitution Parameterization.

|  | Relative DFT | CHARMM*      |
|--|--------------|--------------|
| Bond (Å)                                   |              |              |
| SO...OT                                    | 2.65         | 2.41 (−0.24) |
| SS...OT                                    | 2.65         | 2.34 (−0.31) |
| SO...MG                                    | 2.56         | 2.06 (−0.50) |
| SS...MG                                    | 2.54         | 2.05 (−0.49) |
| SO—P                                       | 2.01         | 2.00 (−0.01) |
| SS—P                                       | 2.00         | 2.00 (0.00)  |
| Angle (°)                                  |              |              |
|  | DFT          | CHARMM*      |
| ON2—P—SO                                   | 107.9        | 107.9 (0.0)  |
| ON2—P—SS                                   | 108.4        | 108.0 (−0.4) |
| ONS—P—SO                                   | 121.0        | 120.6 (−0.4) |
| SS—P—SS                                    | 121.4        | 121.1 (−0.3) |
| Binding Energy (kcal/mol)                  |              |              |
|  | Relative DFT | CHARMM*      |
| DMP(oo)/(so)...H <sub>2</sub> O            | −2.7         | −2.8 (−0.1)  |
| DMP(oo)/(ss)...H <sub>2</sub> O            | −3.4         | −3.8 (−0.4)  |
| DMP(oo)/(ss)-(g-t)/(g-g)                   | −1.8         | −1.5 (0.2)   |
| Mg(HOH)5(DMP)-(oo)/(so)...H <sub>2</sub> O | −4.5         | −5.2 (−0.7)  |
| Mg(HOH)5(DMP)-(oo)/(ss)...H <sub>2</sub> O | −6.8         | −7.5 (−0.7)  |

Training set average geometries and reaction energy for L-J fitting. DMP-oo/ss...H<sub>2</sub>O above is the relative energy difference between water binding to DMP<sup>−</sup> with and without thio substitution at the non-bridging oxygens. The binding distance for SO...MG and SS...MG were fit using a nonbond fix (NBFIX). Geometries were fit to the shift predicted by DFT relative to the unsubstituted DMP<sup>−</sup>. (g-t) vs. (g-g) indicates the gauche–trans vs. gauche–gauche conformations. See the Supporting Information for more data on the fitting function weight values used.

ON2). Adjusted CHelpG charges for the 2', 3'-cyclic phosphate ring were calculated, and force constants were determined through fitting to relaxed PES scans. Parameters were fit to the DFT data using ABNR minimization of the CHARMM\* residue. New dihedrals were set to zero, analogous to those of the CHARMM27 ribose ring. Final parameters fit with very small errors with respect to the DFT results (Table 7).

### Thio Substitutions

To determine the role of divalent metal ions in a reaction, it is often useful to perform thio-effect experiments.<sup>75,76</sup> Changing the oxygen atom to a sulfur effects how strongly a divalent metal can bind to

this position due to the differences between their sizes, polarizabilities, and bond lengths with phosphorus. While a “hard” ion like Mg<sup>2+</sup> will bind very tightly to a correspondingly “hard” oxygen, it will bind much more weakly to a “softer,” more diffuse sulfur<sup>77</sup>. If the divalent metal ion binding at a certain position is required for catalysis, a large decrease in the reaction rate should be seen upon thio substitution at this position. However, the interpretation of thio effect results are sometimes inconclusive,<sup>21,23,78,79</sup> especially when possible conformational changes induced by sulfur are considered. Molecular simulation of thio-substituted phosphates and phosphoranes and their role on structure would aid in the interpretation of experimental thio effects.

A training set of phosphate compounds, bound and unbound to hexa-coordinated Mg<sup>2+</sup>, with both single and double sulfur substitutions, was constructed (Fig. 4). Adjusted CHelpG partial atomic charge calculations were calculated for both mono- and di-thio DMP<sup>−</sup> substitutions. Iterative optimization of the L-J parameters for sulfur was made, until the best possible fit to the DFT-calculated sulfur water-binding distance and relative water-binding energy between DMP<sup>−</sup> and di-thio substituted DMP<sup>−</sup> was obtained. Force constants and equilibrium geometry parameters were then optimized iteratively along with the L-J parameters, until minimization of the  $\chi^2$  function was achieved.

In order to improve the differential binding of Mg<sup>2+</sup> to the nonbridging phosphoryl O and S atoms, a NBFIX term<sup>42</sup> was used to parameterize specifically for Mg<sup>2+</sup>–S interactions; i.e., a specific value for the two-body nonbonded van der Waals parameters was included explicitly for Mg<sup>2+</sup>–S, rather than using the Lorentz–Berthelot combining rules to derive the two-body parameters from one-body Mg<sup>2+</sup> and S parameters. The resulting fit was further improved by the introduction of three new atom types. Two for the nonbridging oxygen (ONS) and sulfur (SO) of a monosubstituted phosphate and one for the nonbridging sulfur (SS) of a dithiosubstituted phosphate. The DFT training set average binding distance for a Mg<sup>2+</sup>–S was 2.55 Å as compared to the CHARMM\* average value of 2.06 Å (Table 8). The shortened distances arose out of the need to strengthen the binding energy results that, in the absence of explicit polarization on the soft sulfur atoms, is considerably underestimated with nonpolarizable force fields. In the case of water binding to the substituted sulfurs, the distances are in better agreement and the relative binding energies are all within 1.0 kcal/mol of the DFT values.

The  $\zeta$  *ab initio* torsional PESs for the native and nonbridging thio-substituted DMP<sup>−</sup> are shown in Figure 7 (top frame). The

**Table 9.** Dihedral Parameters of Nonbridging Oxygen and Sulfur for Thio-Substituted Phosphate.

|                     | Atom types |   |     |     | Force constant | Fold | Phase  |
|---------------------|------------|---|-----|-----|----------------|------|--------|
|                     |            |   |     |     |                |      |        |
| Thio substitution   | ONS        | P | ON2 | CN9 | 0.20           | 2    | 0.00   |
|                     | ONS        | P | ON2 | CN9 | 0.20           | 3    | 0.00   |
|                     | SO         | P | ON2 | CN9 | 0.40           | 1    | 180.00 |
|                     | SO         | P | ON2 | CN9 | 0.10           | 3    | 0.00   |
| Dithio substitution | SS         | P | ON2 | CN9 | 0.45           | 1    | 180.00 |
|                     | SS         | P | ON2 | CN9 | 0.20           | 2    | 0.00   |
|                     | SS         | P | ON2 | CN9 | 0.10           | 3    | 0.00   |

**Table 10.** Geometry Fitting Results for Thio-substituted Ribose Phosphorane Parameterization.

|  | DFT   | CHARMM*     |
|--|-------|-------------|
| Bond (Å)   |       |             |
| P–S <sub>R/S</sub>   | 2.10  | 2.10 (0.00) |
| Angle (°)  |       |             |
| O <sub>R</sub> –P–S <sub>S</sub>   | 125.4 | 125.3 (0.1) |
| S <sub>R</sub> –P–S <sub>S</sub>   | 125.3 | 125.3 (0.0) |
| O <sub>3</sub> –P–S <sub>R/S</sub>   | 113.4 | 113.5 (0.1) |
| O <sub>2</sub> –P–S <sub>R/S</sub>   | 91.4  | 91.4 (0.0)  |
| O <sub>5</sub> –P–S <sub>R/S</sub>   | 97.3  | 97.3 (0.0)  |
| ribose MePA-(oo/os) <sup>2-</sup> + H <sub>2</sub> O →<br>ribose MePA-(oo/os) <sup>2-</sup> :H <sub>2</sub> O (kcal/mol) | 6.6   | 7.9 (1.3)   |
| ribose MePA-(oo/ss) <sup>2-</sup> + H <sub>2</sub> O →<br>ribose MePA-(oo/ss) <sup>2-</sup> :H <sub>2</sub> O (kcal/mol) | 6.4   | 7.7 (1.3)   |

Training set average geometries and energies for geometry fitting and L-J fitting of dianionic ribose methyl phosphorane.

minima identified agree well with fully relaxed structures in previous work by Florián et al.<sup>80</sup> A significant difference between the thio-substituted and unsubstituted surfaces is the minima at the staggered position, 180°, which is negligible for the native DMP<sup>-</sup>, but of almost equal energy to the gauche minima in the case of dithio-substituted DMP<sup>-</sup>. Also noteworthy is the shift of the gauche minima (70°) and gauche–staggered barrier with thio substitution. New torsional parameters, summarized in Table 9, are defined along the S–P–O–C and O–P–O–C dihedrals to qualitatively reproduce the changes indicated by the *ab initio* calculations, while maintaining consistency with the CHARMM27 force field.

A scan of the DMP<sup>-</sup> and thio-substituted torsions with the same constraints as the *ab initio* calculations using the CHARMM27 parameters and those developed in this work are shown in Figure 7 (bottom frame). As indicated in Foloppe and MacKerel,<sup>43</sup> the barrier between gauche–gauche states in the unsubstituted DMP<sup>-</sup> has been significantly lowered to reproduce results from experiment, while the gauche–staggered barrier is left relatively unchanged. The thio-substituted structures retain the lowered gauche–gauche barrier, and the gauche minima is shifted as indicated by the *ab initio* calculations. The staggered conformation has been lowered in energy for both thio-substituted molecules. For the singly substituted case, the minima at 300° was also lowered in addition to the minima at 60°, but the relative energy of the two gauche and eclipsed conformations were maintained. The lowering staggered conformation may be of particular importance to catalytic differences in native and thio-substituted structures, since conformational deformation is a prerequisite to reaction.<sup>32,63,81,82</sup>

L-J parameters for the nonbridging atoms of the penta-coordinated dianionic phosphorane structures were parameterized using a training set that contained methyl ribose phosphorane (ribose MePA<sup>2+</sup>), mono-thio substituted ribose MePA-so<sup>2+</sup> and di-thio substituted ribose MePA-ss<sup>2+</sup>, with and without water bound (Fig. 5). In the case of the unsubstituted phosphorane, it was required to freeze the P-O2' bond, as discussed in Results and Discussion section. Stable gas-phase sulfur-substituted phosphoranones were used for the other structures. Both the nonbridging S and O atoms were

reparameterized in this case because of the large differences in geometry and charge distribution of phosphate and phosphorane. As before, equilibrium geometry values were fit to DFT structures, with force constants for all modified bonds and angles determined through fitting to relaxed PES scans. Geometry fitting results are shown in Table 10.

## Conclusions

Molecular simulation force field parameters are presented for a series of nonstandard residues important in RNA catalysis. Parameters are based on density-functional calculations and developed to be consistent with the CHARMM27 all-atom empirical force field for nucleic acids. Parameters have been developed for an activated (2'O deprotonated) ribose phosphate representing an early reactive state, a penta-coordinate phosphorane intermediate/transition state model, and a 2',3'-cyclic phosphate transesterification product. In addition, parameters for thio-substituted analogs important in the study of experimental thio effects, and specific, single and dimetal Mg<sup>2+</sup> complexes, and μ-bridging OH<sup>-</sup> ion parameters have been developed. These parameters, which are optimized specifically for the present systems, will allow the simulation of reactive intermediates and experimentally modified residues important in the study of ribozyme mechanisms. Further validation and testing of these parameters in simulations is important, although this will likely be a lengthy endeavor because of the lack of available experimental structural data for the reactive intermediates and thio-substituted nucleic acids. Ultimately, it is the hope that simulations of these systems, together with experiment, will help paint a more detailed picture of the catalytic mechanisms of RNA catalysis.

## Acknowledgment

Computational resources were provided by the Minnesota Supercomputing Institute.

## References

1. Stark, B. C.; Kole, R.; Bowman, E. J.; Altman, S. Proc Natl Acad Sci USA 1978, 75, 3717.
2. Guerrier-Takada, C.; Gardiner, K.; Maresh, T. Cell 1983, 35, 849.
3. Cech, T.; Zaug, A.; Grabowski, P. Proc Natl Acad Sci USA 1979, 76, 5051.
4. Zaug, A. J.; Cech, T. R. Science 1986, 231, 849.
5. Novina, C. D.; Murray, M. F.; Dykxhoorn, D. M.; Beresford, P. J.; Riess, J.; Lee, S.-K.; Collman, R. G.; Lieberman, J.; Shankar, P.; Sharp, P. A. Nat Med 2002, 8, 681.
6. Buryanovskii, L. N.; Shved, A. D. Biopolim Kletka 1996, 12, 20.
7. Holmlund, J. T. Curr Opin Mol Ther 1999, 1, 372.
8. Yeakley, J. M.; Fan, J.-B.; Doucet, D.; Luo, L.; Wickham, E.; Ye, Z.; Chee, M. S.; Fu, X.-D. Nat Biotechnol 2002, 20, 353.
9. Soukup, G. A.; Breaker, R. R. Trends Biotechnol 1994, 17, 469.
10. Scott, W. G.; Murray, J. B.; Arnold, J. R. P.; Stoddard, B. L.; Klug, A. Science 1996, 274, 2065.
11. Scott, W. G. Q Rev Biophys 1999, 32, 241.
12. Burke, J. M. Nat Struct Biol 2001, 8, 382.

13. Rupert, P. B.; Massey, A. P.; Sigurdsson, S. T.; Ferré-Dámaré, A. R. *Science* 2002, 298, 1421.
14. Ferré-Dámaré, A. R.; Zhou, K.; Doudna, J. A. *Nature* 1998, 395, 567.
15. hung Shih, I.; Been, M. D. *Annu Rev Biochem* 2002, 71, 887.
16. Doherty, E. A.; Doudna, J. A. *Annu Rev Biophys Biomol Struct* 2001, 30, 457.
17. Takagi, Y.; Ikeda, Y.; Taira, K. *Top Curr Chem* 2004, 232, 213.
18. Takagi, Y.; Taira, K. *J Am Chem Soc* 2002, 124, 3850.
19. He, Q.-C.; Zhou, J.-M.; Zhou, D.-M.; Nakamatsu, Y.; Baba, T.; Taira, K. *Biomacromolecules* 2002, 3, 69.
20. Sawata, S.; Komiyama, M.; Taira, K. *J Am Chem Soc* 1995, 117, 2357.
21. Suzumura, K.; Takagi, Y.; Orita, M.; Taira, K. *J Am Chem Soc* 2004, 126, 15504.
22. Laura, M.; Hunsicker, V. J. D. *J Inorg Biochem* 2000, 80, 271.
23. Scott, E. C.; Uhlenbeck, O. C. *Nucleic Acids Res* 1999, 27, 479.
24. Pley, H. W.; Lindes, D. S.; DeLuca-Flaherty, C.; McKay, D. B. *J Biol Chem* 1993, 268, 19656.
25. Murray, J. B.; Szöke, H.; Szöke, A.; Scott, W. G. *Mol Cell* 2000, 5, 279.
26. Beveridge, D. L.; McConnell, K. J. *Curr Opin Struct Biol* 2000, 10, 182.
27. Auffinger, P.; Westhof, E. *Curr Opin Struct Biol* 1998, 8, 227.
28. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J Am Chem Soc* 1995, 117, 5179.
29. Giudice, E.; Lavery, R. *Acc Chem Res* 2002, 35, 350.
30. Cheatham, T. E., III; Young, A. M. *Biopolymers* 2001, 56, 232.
31. Mackerell, A. D., Jr. *J Comput Chem* 2004, 25, 1584.
32. Torres, R. A.; Bruice, T. C. *J Am Chem Soc* 2000, 122, 781.
33. Sarzynska, J.; Nilsson, L.; Kulinski, T. *Biophys J* 2003, 85, 1522.
34. Dvorsky, R.; Sevcik, J.; Caves, L. S. D.; Hubbard, R. E.; Verma, C. S. *J Phys Chem B* 2000, 104, 10387.
35. Réblová, K.; Špačková, N.; Štefl, R.; Csaszar, K.; Koča, J.; Leontis, N. B.; Šponer, J. *Biophys J* 2003, 84, 3564.
36. Boero, M.; Terakura, K.; Tateno, M. *J Am Chem Soc* 2002, 124, 8949.
37. Le, S.-Y.; Chen, J.-H.; Pattabiraman, N.; Maizel, J. V., Jr. *J Biomol Struct Dyn* 1998, 6, 1.
38. Auffinger, P.; Westhof, E. *J Mol Biol* 1997, 269, 326.
39. Hermann, T.; Auffinger, P.; Scott, W. G.; Westhof, E. *Nucleic Acids Res* 1997, 25, 3421.
40. Hermann, T.; Auffinger, P.; Westhof, E. *Eur Biophys J* 1998, 27, 153.
41. MacKerell, A. D., Jr.; Brooks, B.; Brooks, C. L., III; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. In *Encyclopedia of Computational Chemistry*, Vol. 1, Schleyer, V. R. P.; Allinger, N. L.; Clark, T.; Gasteiger, J.; Kollman, P. A.; Schaefer, H. F., III; Schreiner, P. R., Eds.; Wiley: Chichester, 1998; pp 271–277.
42. Foloppe, N.; MacKerell, A. D., Jr. *J Comput Chem* 2000, 21, 86.
43. MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J Phys Chem B* 1998, 102, 3586.
44. MacKerell, A. D., Jr.; Banavali, N.; Foloppe, N. *Biopolymers* 2001, 56, 257.
45. Stone, A. *International Series of Monographs in Chemistry*, Vol. 32: *The Theory of Intermolecular Forces*; Clarendon Press: Oxford, 1996.
46. MacKerell, A. D., Jr.; Banavali, N. K. *J Comput Chem* 2000, 21, 105.
47. MacKerell, A. D., Jr.; Wiórkiewicz-Kuczera, J.; Karplus, M. *J Am Chem Soc* 1995, 117, 11946.
48. Yang, L.; Pettitt, B. M. *J Phys Chem A* 1996, 100, 100.
49. Feig, M.; Pettitt, B. M. *J Phys Chem B* 1997, 101, 7361.
50. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J Chem Phys* 1983, 79, 926.
51. *Nucleic Acids Database*. <http://ndbserver.rutgers.edu>.
52. Becke, A. D. *J Chem Phys* 1993, 98, 5648.
53. Lee, C.; Yang, W.; Parr, R. G. *Phys Rev B* 1988, 37, 785.
54. Range, K.; McGrath, M. J.; Lopez, X.; York, D. M. *J Am Chem Soc* 2004, 126, 1654.
55. Mayaan, E.; Range, K.; York, D. M. *J Biol Inorg Chem* 2004, 9, 807.
56. López, C. S.; Faza, O. N.; Gregersen, B. A.; Lopez, X.; de Lera, A. R.; York, D. M. *Chem Phys Chem* 2004, 5, 1045.
57. López, C. S.; Faza, O. N.; de Lera, A. R.; York, D. M. *Chem Eur J* 2005, 11, 2081.
58. Liu, Y.; Lopez, X.; York, D. M. *Chem Commun* 2005, 31, 3909.
59. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision B.05*. Gaussian, Inc.: Wallingford, CT, 2003.
60. Chirlian, L. E.; Francl, M. M. *J Comput Chem* 1987, 8, 894.
61. Breneman, C. M.; Wiberg, K. B. *J Comput Chem* 1990, 11, 361.
62. Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J Phys Chem* 1993, 97, 10269.
63. Atereshko, V.; Wallace, S. T.; Usman, N.; Wincott, F. E.; Egli, M. *RNA* 2001, 7, 405.
64. Soukup, G. A.; Breaker, R. R. *RNA* 1999, 5, 1308.
65. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J Comput Chem* 1983, 4, 187.
66. Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, W. P. *Numerical Recipes in FORTRAN*, 2nd ed.; Cambridge University Press: Cambridge, 1992.
67. Hertel, K. J.; Uhlenbeck, O. C. *Biochemistry* 1995, 34, 1744.
68. Lyne, P. D.; Karplus, M. *J Am Chem Soc* 2000, 122, 166.
69. Koizumi, M.; Ohtsuka, E. *Biochemistry* 1991, 30, 5145.
70. Cunningham, L. A.; Li, J.; Lu, Y. *J Am Chem Soc* 1998, 120, 4518.
71. Lee, H.; Darden, T. A.; Pedersen, L. G. *J Chem Phys* 1995, 102, 3830.
72. Mineva, T.; Russo, N.; Sicilia, E. *J Comput Chem* 1998, 19, 290.
73. Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J Chem Phys* 2002, 117, 43.
74. Murray, J. B.; Terwey, D. P.; Maloney, L.; Karpeisky, A.; Usman, N.; Beigelman, L.; Scott, W. G. *Cell* 1998, 92, 665.
75. Herschlag, D.; Piccirilli, J. A.; Cech, T. R. *Biochemistry* 1991, 30, 4844.
76. Oivanen, M.; Kuusela, S.; Lönnberg, H. *Chem Rev* 1998, 98, 961.
77. Pearson, R. G. *J Chem Educ* 1987, 64, 562.
78. Zhou, D.-M.; He, Q.-C.; Zhou, J.-M.; Taira, K. *FEBS Lett* 1998, 431, 154.
79. Yoshinari, K.; Taira, K. *Nucleic Acids Res* 2000, 28, 1730.
80. Florián, J.; Štrajbl, M.; Warshel, A. *J Am Chem Soc* 1998, 120, 7959.
81. Murray, J. B.; Dunham, C. M.; Scott, W. G. *J Mol Biol* 2002, 315, 121.
82. Scott, W. G. *Curr Opin Struct Biol* 1998, 8, 720.