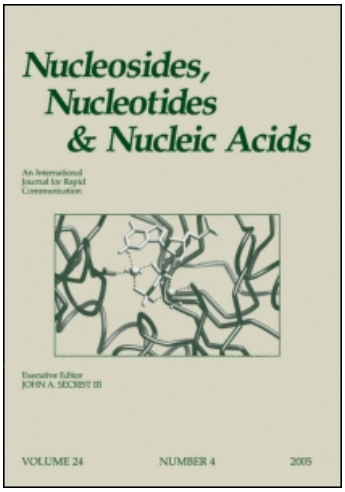

Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

PROCEDURE FOR SELECTING STARTING CONFORMATIONS FOR ENERGY MINIMIZATION OF NUCLEOSIDES AND NUCLEOTIDES

Douglas G. Harris[a]; Jianyin Shao[a]; Jonathan M. Anderson[a]; Douglas P. Marx[a]; S. Scott Zimmerman[a]
[a] Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT, U.S.A.

Online publication date: 12 February 2002

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016

NUCLEOSIDES, NUCLEOTIDES & NUCLEIC ACIDS
Vol. 21, Nos. 11 & 12, pp. 803–812, 2002

PROCEDURE FOR SELECTING STARTING CONFORMATIONS FOR ENERGY MINIMIZATION OF NUCLEOSIDES AND NUCLEOTIDES

Douglas G. Harris, Jianyin Shao, Jonathan M. Anderson, Douglas P. Marx, and S. Scott Zimmerman*

Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT 84602–5700

ABSTRACT

The purpose of this study was to carry out a thorough search of the conformational space of various adenine-containing nucleotides, applying a previously published searching procedure, known as the representative method. This method, which reduces the number of starting conformations required to explore all the important regions of conformational space, appears to be successful in finding all (or nearly all) the putative low-energy conformations of each molecule.

INTRODUCTION

Molecular mechanics (MM) is an important method for analyzing the conformations of nucleosides, nucleotides, and nucleic acids.[1–3] Such an analysis requires a force field parameterized for nucleic acids and their constituents. These force fields are employed in carrying out three major

*Corresponding author. Fax: (801) 422-0153; E-mail: scott_zimmerman@byu.edu

DOI: 10.1081/NCN-120016482 1525-7770 (Print); 1532-2335 (Online)

types of MM simulation of nucleic acids: minimizations (including normal mode analyses), Monte Carlo searches, and molecular dynamics.[2]

In conformational analysis using minimization, one must select a set of starting conformations, minimize the total conformational energy of each of these starting conformations, and thereby produce a set of conformational energy minima. Selecting enough starting conformations to find all low-energy minima is impossible for all but the simplest molecules, due to the large amount of conformational space. This is specifically true for the nucleotides in which we are interested, namely, adenosine 5′-monophosphate (AMP), adenosine 5′-diphosphate (ADP), and adenosine 5′-triphosphate (ATP), as shown in Table 1. All these molecules contain the purine base adenine, a planar ring system with no rotatable torsional bonds. The only rotatable bond involving the adenine in nucleosides and nucleotides is χ, which joins the adenine base to ribose. Furthermore, the conformational space of cAMP is relatively small because (1) the phosphate-containing ring adopts a typical chair conformation, (2) the cyclic phosphate group locks the adjacent ribose ring into the $C_{3'}$-*endo*-$C_{4'}$-*exo* conformation, and (3) only one hydroxyl group (OH2′) is free to rotate. On the other hand, when the phosphate group is not cyclized, as in AMP, ADP, and ATP, the phosphate group becomes more flexible, and concomitantly, the ribose ring is no longer locked into a single conformation. In solution, therefore, the ring can take on two well-characterized conformational states, $C_{2'}$-*endo* and $C_{3'}$-*endo*. Furthermore, these more complex nucleotides have two freely rotating hydroxyl groups. Thus, as the size of the nucleotide increases, the size of the conformational space increases, requiring more and more starting conformations in an energy-minimization analysis.

To circumvent this problem of investigating such a large area of conformational space, researchers have developed various specialized searching methods, which Leach[4] classified as follows: systematic methods, model-building methods, random search methods, distance geometry methods, and molecular dynamics. These conformational search methods are efficient at locating important low-energy minima of small- and medium-sized molecules.

Table 1. Problem of Increasing Complexity

Molecule	Ribose Conformations	Dihedral Angles	# of Starting Conformations
cAMP	1	2	144
Adenosine	2	4	4×10^4
AMP	2	6	6×10^6
ADP	2	8	8.6×10^8
ATP	2	10	1.24×10^{11}

To cite one example, Hingerty et al.[5] used a model-building procedure known as the build-up method to predict DNA structure from its nucleotide sequence. This method utilizes global searches of the conformational space of each nucleic acid monomer unit, and then "builds up" the particular nucleotide sequence by combining the minimum-energy conformations of the monomer nucleic acid building blocks. This approach was successful in locating B- and Z- forms of DNA for the $d(CG)_6 \cdot d(CG)_6$ segment investigated.[5]

In this study, we tested another model-building procedure known as the representative method, originally developed for the conformational analysis of peptides.[6] The details of this method as applied in this study appear in the Methods section below. Using the representative method to choose conformation for subsequent minimization, we minimized the total conformational energy of these starting conformations to obtain an ensemble of minima for each molecule investigated. From this ensemble, we then calculated the probability of occurrence of each low-energy conformation. From an analysis of the ensemble of low-energy conformations, we developed a "conformational profile" to describe the entire collection of calculated dihedral probabilities of the molecule under investigation.

Conformational profiles could have direct application to drug activity analysis and drug design. When the conformational properties of a drug are known (or suspected) to play an important role in the therapeutic value of the drug, any newly designed analogue should have similar conformational properties to enhance its probability of binding to the receptor target. For example, Mickle and Nair[7] recently demonstrated a correlation between conformational preferences and antiviral activity of various dideoxynucleosides. The success, however, of using conformational profiles depends on the accuracy of the computer simulation. Therefore, in this study, we select familiar nucleosides and nucleotides to ascertain the accuracy of the representative method in duplicating the experimentally determined conformational preferences.

We can then compare our calculated conformational profiles with reported experimental conformations.[1,8] In a similar way, Kollman[3] recently compared the known experimentally determined conformational preferences of nucleosides and nucleotides with theoretical results obtained using the AMBER force field.

Our goals, therefore, in this study are to (1) generate all (or most of) the low-energy minima of each nucleotide by selecting appropriate starting conformations and by minimizing their energy, (2) use the ensemble of low-energy minima to develop conformational profiles of nucleosides and nucleotides, (3) ascertain the accuracy of the conformational profiles in comparison with experimental results, and (4) determine the applicability of the representative method in the conformational analysis of other nucleosides and nucleotides.

This work will serve as a foundation for our future studies on the binding of nucleoside and nucleotide analogues to target enzymes and the applicability of our methodologies in drug design.

METHODS

Throughout this work, we use the nomenclature and conventions adopted by the IUPAC-IUB Commission.[9]

We carried out all calculations using the AMBER force field of the Discover/Insight II program suite from Accelrys, Inc., on Silicon Graphics O2, Indigo2, and Power Challenge workstations.

We did not explicitly include solvation in the calculations but attempted to mimic the effects of solvent by utilizing a distance-dependent dielectric constant ($\varepsilon = 4R_{ij}$ as suggested in the Weiner version of the AMBER force field.[10] Weiner et al.[10] found that specific conformations of the adenosine OH2′ orientation and highly damped intramolecular electrostatic interactions were necessary to reproduce experimental results and adjusted the dielectric constant accordingly. Using this dielectric constant, they computed an energy difference between the northern ($C_{3'}$-*endo*) and southern ($C_{2'}$-*endo*) orientations of the nucleoside ring of adenosine to be 0.29 kcal/mol, which compares favorably with the experimentally observed value of 0.19–0.42 kcal/mol.

In our study, we subjected each starting conformation to a Newton-Raphson minimization until the average gradient was less than 0.01 kcal mol^{-1} $Å^{-1}$. When the minimization was complete (due to convergence), we assumed without further testing that the resulting conformation was a minimum in conformational space.

Minimizing the total conformational energy from each starting conformation yielded a minimum-energy conformation with a total conformational energy E. The value of E for the assumed global minimum of each molecule was designated E_0. The relative energy of each minimum, therefore, was calculated to be $\Delta E = E - E_0$.

We applied a systematic grid search with subsequent minimization to the nucleoside adenosine. This type of search was possible due to the small number (four) of rotatable dihedral angles in the molecule. A plot of the locations of adenosine minima in the χ-γ and OH2′-ε planes of conformational space revealed that the minima were clustered into a few areas, as shown in Figs. 1 and 2. These clusters of minima in conformational space suggested the potential success of the representative method in the nucleosides and nucleotides. In this method, we selected one point to represent all the points in each densely populated region of the adenosine χ-γ and OH2′-ε maps. (Although these different minima are close together in one plane of conformational space, they are not necessarily close together in other planes;

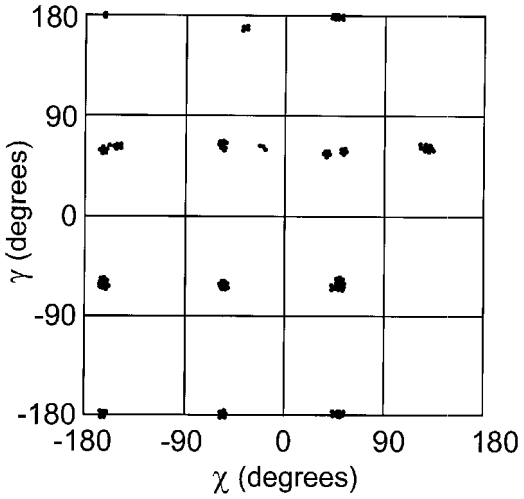

Figure 1. Plot of the locations of adenosine minima in the χ-γ plane of conformational space.

each point is a distinct minimum in multidimensional conformational space.) Thus, we chose 10 representative χ-γ values and 9 representative OH2′-ε values for adenosine, one for each of the 10 clusters in Fig. 1 and 9 clusters in Fig. 2, respectively.

We also used the two ribose-ring conformations ($C_{2'}$-*endo* and $C_{3'}$-*endo*) in generating the list of representative points from the adenosine minima for subsequent minimization of AMP, ADP, and ATP. In our initial studies, we also used the higher energy $O_{4'}$-endo conformation of adenosine as a starting

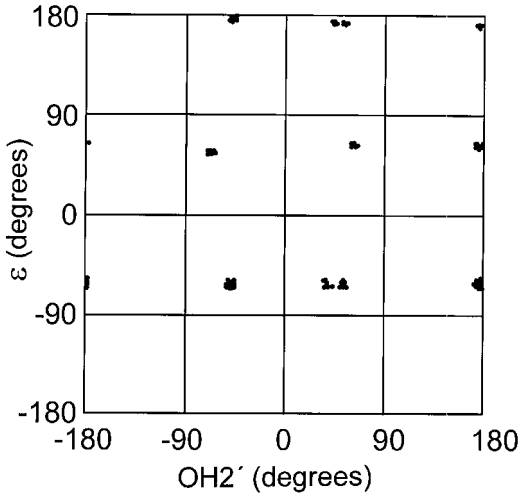

Figure 2. Plot of the locations of adenosine minima in the Epsilon-2′OH plane of conformational space.

point for the furanose ring, but in all cases, energy minimization yielded the more stable $C_{2'}$-endo or $C_{3'}$-endo conformations. Therefore, in our complete analysis, we considered only the two ribose low-energy ring conformations described above, $C_{2'}$-endo and $C_{3'}$-endo.

To model the more complex nucleotides of AMP, ADP, and ATP, we used minima from the fragments of methyl mono, di, and triphospate to obtain the representative points for starting conformations in the subsequent minimization, similar to the build-up method. Even the simple molecule of methyl tri-phosphate contains six dihedral angles that can vary in the conformational searches and subsequent minimizations. If we rotated each of these dihedral angles for the entire 360° in 30° increments, we would generate nearly three million starting conformations. With the representative method, however, we carried out the minimization with only 96 starting conformations.

The representative method differs from the build-up method in the following way. The representative method uses all combinations of representative χ, γ, OH2′, and ε values of adenosine as starting conformations for the nucleotides AMP, ADP, and ATP, independent of which combinations of those dihedral angles actually lead to adenosine minima. Many of those combinations are, in fact, adenosine minima, but many of them are not. The build-up method, on the other hand, uses only combinations of adenosine minima and therefore would not generate as many starting conformations.

Once we obtained low-energy minima, we used Boltzmann statistics to determine an average energy for the ensemble of minima and probabilities of dihedral angle distribution for each molecule, following the methods described earlier.[11] The specific conformational regions of each dihedral are defined as follows: gauche plus (g+) around 60°, gauche minus (g−) around −60°, and trans (t) around 180°. The compilation of the relative preferences for these and other conformations determines the previously defined conformational profile of each molecule.

RESULTS AND DISCUSSION

Table 2 lists the numbers of low-energy minima found for adenosine, cAMP, AMP, ADP, and ATP within the specified energy ranges relative to the global minimum energy E_0. Note that we locate many other minima for $\Delta E > 6.0$ kcal/mol, but did not include them in the table. As shown in the table, for adenosine, we found 488 unique minima within 6.0 kcal/mol of the global minimum. For this molecule, we employed a systematic grid search, with subsequent minimization, to locate these minima, whereas we employed the representative method to analyze the molecules AMP, ADP, and ATP.

Table 2. The Number of Minimum-Energy Conformations Generated for the Adenine Containing Nucleoside and Nucleotides within the Specified Energy Cutoff (ΔE)

ΔE (kcal/mol)	Number of Minima				
	Adenosine	cAMP	AMP	ADP	ATP
1.0	17	2	4	3	8
2.0	72	1	31	13	21
3.0	229	1	62	47	63
4.0	390	–	109	102	122
5.0	474	–	182	180	221
6.0	488	–	259	280	328

One would expect that a systematic grid search would produce a greater number of unique minima than other search methods, including the representative method. This proved to be the case, as seen by that fact that the simple adenosine, which was analyzed using a grid search, yielded more minima (488) with $\Delta E \leq 6$ kcal/mol than the more complex AMP (with 259 minima), ADP (with 280 minima), and ATP (with 328 minima) (see Table 2), which were analyzed using the representative method.

On the other hand, the representative method was more efficient than the grid search method because, with the former, a higher percentage of the starting conformations lead to unique minima. Specifically, only 3% of the starting conformations obtained using the grid search method for adenosine yielded unique minima, whereas 14% of the starting conformations obtained using the representative method for AMP yield unique minima, and this percentage increases for the more complex nucleotides—20% for ADP and 22% for ATP.

We cannot be certain, though, if our application of the representative method locates all the important low-energy minima for these building blocks of nucleic acids, because there is no way of knowing what percent of all local minima were found or even if the global minimum was located.[6] An exhaustive grid search of the conformational space would certainly cast more light upon these questions, but given the number of dimensions of conformational space and the number of intervals required for each angle, the problem of finding all local minima seems intractable except for small nucleosides.

The advent of increased computational power facilitates the application of ab initio methods in determining more accurate energies for global minima. The representative method could prove useful in determining structures to be used as starting conformations for subsequent ab initio calculations.

Although ab initio calculations might provide more accurate energies of nucleosides and nucleotides, molecular mechanics calculations are still

valuable because they allow (1) generation of a large ensemble of low-energy minima (which would require excessive computer time to generate using ab initio methods) when details of all the low-energy minima are of interest, and (2) a basis for comparing conformational energies when small- and medium-sized ligand molecules are used as part of a larger calculation involving, for example, ligand-enzymes complexes (which are too large for ab initio calculations but which are a major emphasis of the research in our laboratory).

Table 3 shows the calculated conformational profiles generated by the representative method for each of the studied molecules. Rotations of the α phosphodiester linkage in ADP and ATP favor the g+ conformation, with 50% probability of occurrence, with the t conformation also significant (about 40% probability). Rotations of the β phosphodiester linkage in AMP, ADP, and ATP favor the t conformation almost exclusively (60–90% probability), where P···C4′ repulsions are minimized. The γ bond of all the studied molecules greatly prefers the g+ conformation. Saenger[1] points out that the high probability of finding the nucleotides in the g+ conformation is due to extra stabilization provided by a base C-H···O5′ hydrogen bond, which also influences the syn/anti orientation of the base. The dihedral bond χ can exist in both anti and syn positions for purines. The anti position is highly favored for the purine nucleotides due to the previously mentioned C-H···O5′ interaction. Calculated χ probabilities indicate that the anti conformation is also favored in the studied nucleotides.

Saenger[1] describes the conformational preferences of the nucleotide cAMP. Experimental observations indicate that the dihedral angle χ for this molecule can exist in either the anti or syn conformation. Surprisingly, the crystal structure, which contains two molecules of cAMP in the asymmetric unit cell, has both the anti and syn conformations of χ represented.[12] Our calculations are compatible with this observation, in that the two lowest energy minima of cAMP consist of one conformation with χ in the anti and one with χ in the syn. The other two minima (of the four total minima)

Table 3. Conformational Profiles of Adenosine, AMP, ADP, and ATP. Values Are Calculated Probabilities Conformations Corresponding to each Type of Conformation

	α			β			γ			χ	
	g+	g−	t	g+	g−	t	g+	g−	t	Anti	Syn
Adenosine	–	–	–	–	–	–	0.59	0.04	0.37	0.38	0.62
AMP	–	–	–	0.07	0.31	0.61	0.62	0.37	0.01	0.78	0.22
ADP	0.50	0.09	0.41	0.07	0.11	0.83	0.92	0.08	0.00	0.88	0.12
ATP	0.54	0.08	0.37	0.06	0.07	0.87	0.91	0.05	0.05	0.67	0.33

found for cAMP also have χ in the anti and syn but with different orientations of OH2′.

CONCLUSION

We have successfully applied the conformational searching procedure known as the representative method to analyze the conformational properties of selected adenine containing nucleotides and nucleosides. The representative method permits a reasonably thorough search of the conformational space of nucleotides and produces ensembles of minima that possess experimentally observed characteristics. Using the results of the representative method, therefore, allows us to determine reasonable conformational profiles of nucleotides. We are currently applying this method to the conformational analysis of nucleoside and nucleotide analogues and using the results to help compute binding energies of these analogues with enzymes of therapeutic interest.

ACKNOWLEDGMENTS

This work was partially supported by a fellowship provided by the Brigham Young University Department of Chemistry and Biochemistry. We would also like to thank Ira Fulton for donating the SGI O2 workstation used in this research.

REFERENCES

1. Saenger, W. Structures and Conformational Properties of Bases, Furanose Sugars, and Phosphate Groups. In *Principles of Nucleic Acid Structure*; Springer Advanced Texts in Chemistry; Cantor, C.R., Eds.; Springer-Verlag New York Inc.: New York, 1984; 51–104.
2. Miller, J.L.; Cheatham III, T.E.; Kollman, P.A. Simulation of Nucleic Acid Structure. In *Oxford Handbook of Nucleic Acid Structure*; Neidle, S., Ed.; Oxford University Press Inc.: New York, 1999; 95–115.
3. Kollman, P.A. Theoretical Methods. In *Nucleic Acids: Structures, Properties, and Functions*; Bloomfield, V.A., Crothers, D.M., Tinoco, Jr., I., Eds.; University Science Books: California, 2000; 223–258.
4. Leach, A.R. A Survey of Methods for Searching the Conformational Space of Small and Medium-Sized Molecules. In *Reviews in Computational Chemistry*; Lipkowitz, K.B., Boyd, D.B., Eds.; VCH Publishers: New York, 1991; 2, 1–55.
5. Hingerty, B.E.; Figueroa, S.; Hayden, T.L.; Broyde, S. Prediction of DNA Structure from Sequence: A Build-up Technique. Biopolymers **1989**, *28*, 1195–1222.

6. Davidson, R.B.; Zimmerman, S.S. Evaluation of Two Procedures for Selecting Starting Conformations for Energy Minimization of Peptides. J. Chem. Inf. Comput. Sci. **1994**, *34*, 1009–1013.
7. Mickle, T.; Nair, V. Correlation of Anti-HIV Activity with Structure: Use of Electrostatic Potential and Conformational Analysis. Bioorg. Med. Chem. Lett. **1999**, *9*, 1963–1968.
8. Saenger, W. Modified Nucleosides and Nucleotides; Nucleoside Di- and Triphosphates; Coenzymes and Antibiotics. In *Principles of Nucleic Acid Structure*; Springer Advanced Texts in Chemistry; Cantor, C.R., Eds., Springer: New York, 1984; 159–200.
9. IUPAC-IUB Abbreviations and Symbols for the Description of Conformation of Polynucleotide Chains. Eur. J. Biochem. **1983**, *131*, 9–15.
10. Weiner, S.J.; Kollman, P.A.; Nguyen, D.T.; Case, D.A. An All Atom Force Field for Simulations of Proteins and Nucleic Acids. J. Comp. Chem. **1986**, *2*, 230–252.
11. Zimmerman, S.S.; Shipman, L.L.; Scheraga, H.A. Bends in Globular Proteins. A Statistical Mechanical Analysis of the Conformational Space of Dipeptides and Proteins. J. Phys. Chem. **1977**, *81* (7), 614–622.
12. Watenpaugh, K.; Dow, J.; Jensen, L.H.; Furberg, S. Crystal and Molecular Structure of Adenosine 3′,5′-Cyclic Phosphate. Science **1968**, *159*, 206–207.

Received March 11, 2002
Accepted August 22, 2002