RNA backbone rotamers – finding your way in seven dimensions

L.J.W. Murray, J.S. Richardson[1], W.B. Arendall III and D.C. Richardson

Department of Biochemistry, Duke University, Durham, NC 27710-3711, U.S.A.

Abstract

Despite the importance of local structural detail for a mechanistic understanding of RNA catalysis and binding functions, RNA backbone conformation has been recalcitrant to analysis. There are too many variable dihedral angles per residue, and their raw empirical distributions are poorly clustered. This study applies quality-filtering techniques (using resolution, crystallographic B factor and all-atom-steric clashes) to the backbone dihedral angle distributions from a selected 8636 residue RNA database. With noise levels significantly decreased, clear signal appears for the underlying angle preferences. We analyse the multidimensional backbone dihedral distributions within sugar-to-sugar 'suites' rather than chemical residues due to the greater base interaction and steric interdependence within the suite. The final result is a small library of RNA backbone rotamers, each represented by a data cluster in seven-dimensional dihedral space, which should provide valid conformations for nearly all RNA backbones encountered in experimental structures. We are in the process of improving that library, and developing tools and applications for it in structure determination and analysis.

Introduction

Detailed structural analysis has always been much more difficult for nucleic acid backbone than for nucleic acid bases: the backbone has a high number of degrees of freedom, frequently underdetermined in both X-ray crystallography and NMR, and the resulting challenges to accurate structure determination in turn hinder systematic study that would assist in the development of better tools for structure determination. Previous approaches to RNA-backbone analysis either gave up on categorization (e.g. RNABase at http://www.rnabase.org) or achieved motif identification by major simplification such as reduction to two pseudo-torsion angles [1]. We are interested, however, in a categorized but complete backbone description that can support both crystallographic location of individual atoms and biological inference about molecular interactions that see the full level of detail. Availability of coordinates for the ribosome structures [2–4] increased the size of the database to a point where we could apply our quality-filtering techniques to develop a rotamer library and related tools for RNA-backbone fitting, analysis and correction.

Methods

Our approach is based on techniques previously developed in the Richardson laboratory for application to proteins, using all-atom-contact analysis [5] and other quality filters in the critical analysis of dihedral distributions [6]. Corrections based on these techniques have resulted in simultaneous improvements to geometrical criteria and fit to experimental data in protein structures [7].

Figure 1 | Cross-window point picking in MAGE: one of the two dinucleotide-platform rotamers

(**a**) Picking an α-β-γ point (from a plot with sugar pucker specified) automatically selects the corresponding δ^{-1}-ε^{-1}-ζ^{-1} point for that suite. (**b**) Examples of this rotamer: rr0033 [14] chain 0 suite 59–60 (picked in 1**a**), pr0133 [15] suite 63–64.
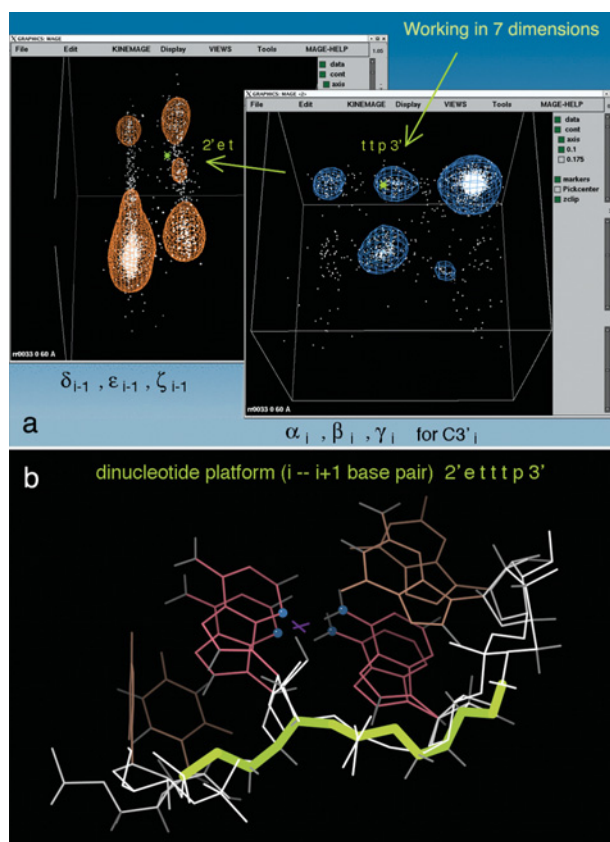
Figure 2 | Correction of rr0033 chain 0 suite 870–871

(**a**) Published conformation with severe steric clashes (red spikes). (**b**) Rotatable model, adjusted from ideal rotamer for A-form to fit bases, phosphorus and chain connections of original. (**c**) Proposed correction with good van der Waals contacts. (**d**) Original and altered conformation with electron density overlaid. Note that sterics have improved without sacrificing fit to data.
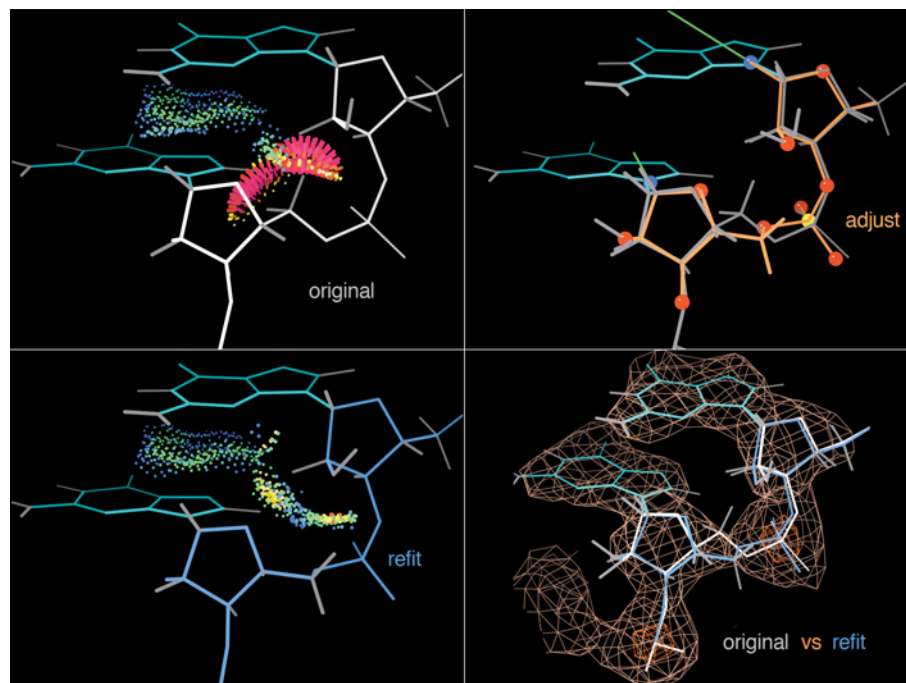
We began by selecting RNA crystal structures from the Nucleic Acid Database [8] at $\leqslant 3.0$ Å (1 Å $= 10^{-10}$ nm) resolution, omitting duplications and backbone modifications. We used resolution and B factor as criteria for quality filters, as well as all-atom-contact analysis: adding and optimizing hydrogens using REDUCE and examining the resulting van der Waals surfaces for steric conflicts using PROBE [9]. Suites with a steric clash of $\geqslant 0.4$ Å between backbone atoms ± 1 residue number were filtered out due to physical impossibility.

The seven-dihedral suite from δ^{-1} through δ is our unit of analysis. To extend our three-dimensional visualizations across the entire suite, we used the plot of filtered δ^{-1}–ε^{-1}–ζ^{-1} values to select seven peak clusters and plotted the corresponding α-β-γ values for each cluster, split by the value of δ, which is strongly bimodal in correspondence with sugar pucker and has a notable effect on the α-β-γ distributions. From each of these 14 plots, we were able to identify clusters of believable data points representing full seven-dihedral rotamer conformations. See Figure 1(a) for an illustration of another method of correlating the plots, using the cross-window picking feature in MAGE [10].

Results and discussion

RNA backbone cannot be said to be entirely relaxed, as steric restrictions lead to ε being nearly eclipsed on a regular basis.

Once noise in the form of random and systematic error is largely filtered out, however, the dihedral distributions show clearly that RNA backbone takes on a relatively manageable number of discrete and distinguishable rotameric conformations.

We were able to identify 42 backbone rotamers in our initial analysis, tabulated in [11]. Because they are based on data filtered for accuracy and reliability, these conformations should represent nearly all legitimate backbone suites in well-fit structures and may be used as starting points in initial fitting or in correcting known problems. We are working with the two other groups who have recently published on the data-driven classification of RNA backbone [12,13] to reach a consensus on terminology and major conformations. In examining several well-known RNA motifs defined primarily on base-pairing criteria, we find that backbone conformation is a consistent aspect and can be made a useful part of their description; see Figure 1(b) for examples of one of the two common dinucleotide-platform rotamers.

A major goal of the present study is to apply the backbone rotamers, geometrical analysis, and all-atom-contact analysis to improve the accuracy of RNA crystal structures, both by correcting problems in existing structures and by creating tools (both computational and rule-based) for easier initial fitting and rebuilding of RNA backbone models into electron density. Figure 2 shows an example of the current stepwise correction process and its results: a badly clashing suite in

the 23 S ribosomal RNA is rebuilt in a rotameric, non-clashing conformation that changes α and γ by 120° but fits the density equally well. Protein residues in the RNA interface can also be corrected, such as arginine side-chains with the guanidinium group flipped over in its triangular density. Multiple changes will be further refined to confirm their validity, as is now successfully done for proteins [7]. This experience will guide the development of more automated tools, which will be made available on our website (http://kinemage. biochem.duke.edu) and its MOLPROBITY services [9].

References

1 Duarte, C.M., Wadley, L.M. and Pyle, A.M. (2003) Nucleic Acids Res. **31**, 4755–4761

2 Ban, N., Nissen, P., Hansen, J., Moore, P.B. and Steitz, T.A. (2000) Science **289**, 905–920

3 Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F. and Yonath, A. (2000) Cell (Cambridge, Mass.) **102**, 615–623

4 Carter, A.P., Clemons, W.M., Brodersen, D.E., Morgan-Warren, R.J., Wimberly, B.T. and Ramakrishnan, V. (2000) Nature (London) **407**, 340–348

5 Word, J.M., Lovell, S.C., LaBean, T.H., Taylor, H.C., Zalis, M.E., Presley, B.K., Richardson, J.S. and Richardson, D.C. (1999a) J. Mol. Biol. **285**, 1711–1733

6 Lovell, S.C., Davis, I.W., Arendall, III, W.B., de Bakker, P.I.W., Word, J.M., Prisant, M.G., Richardson, J.S. and Richardson, D.C. (2003) Proteins: Struct Funct. Gen. **50**, 437–450

7 Richardson, J.S., Arendall, III, W.B. and Richardson, D.C. (2003) in Methods in Enzymology: Macromolecular Crystallography, vol. 374 (Carter, Jr, C.W. and Sweet, R.M., eds.), pp. 385–412, Academic Press, New York

8 Berman, H.M., Olson, W.K., Beveridge, D.L., Westbrook, J., Gelbin, A., Demeny, T., Hsieh, S.-H., Srinivasan, A.R. and Schneider, B. (1992) Biophys. J. **63**, 751–759

9 Davis, I.W., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2004) Nucleic Acids Res. **32** Web Server Issue, W615–W619

10 Richardson, J.S. and Richardson, D.C. (2002) in International Tables for Crystallography, vol. F: Crystallography of Biological Macromolecules (Rossmann, M.G. and Arnold, E., eds.), pp. 727–730, Kluwer Academic Publishers, Dordrecht

11 Murray, L.J.W., Arendall, III, W.B., Richardson, D.C. and Richardson, J.S. (2003) Proc. Natl. Acad. Sci. U.S.A. **100**, 13904–13909

12 Hershkovitz, E., Tannenbaum, E., Howerton, S.B., Sheth, A., Tannenbaum, A. and Williams, L.D. (2003) Nucleic Acids Res. **31**, 6249–6257

13 Schneider, B., Zdenek, M. and Berman, H.M. (2004) Nucleic Acids Res. **32**, 1666–1677

14 Klein, D.J., Schmeing, T.M., Moore, P.B. and Steitz, T.A. (2001) EMBO J. **20**, 4214–4221

15 Adams, P.L., Stahley, M.R., Kosek, A.B., Wang, J. and Strobel, S.A. (2004) Nature (London) **430**, 45–50