SUPPORTING INFORMATION

The Steric Hypothesis for DNA Replication and Fluorine Hydrogen Bonding Revisited in Light of Structural Data

MARTIN EGLI*

*Department of Biochemistry, School of Medicine, Vanderbilt University, Nashville, Tennessee 37232, USA * E-mail: martin.egli@vanderbilt.edu*

Error Estimation for Atomic Coordinates

Various crystal structures of DNA and RNA duplexes containing the dF and rF nucleotide analogs, respectively, complexes of DNA polymerases (Dpo4 and RB69) bound to DNA duplexes with dF opposite A or G (Dpo4), or dATP/dTTP/dGTP/dCTP (RB69) at the active site, as well as of nucleic acid duplexes with incorporated fluorine-modified residues (i.e. 2′-deoxy-2′-fluoro-ribo- or 2′-deoxy-2′-fluoro-arabinonucleotides) were reported by our own laboratory as well as the Konigsberg group (RB69:DNA:dNTP structures). The crystallographic data were phased by the molecular replacement or anomalous dispersion techniques and refined by standard approaches, using either full-matrix least squares optimization with anomalous B-factors and minimal geometry restraints [i.e. the RNA dodecamer duplex with G:rF pairs (PDB 2Q1O,[1] **Table S1**)],[6] or related strategies, such as the maximum likelihood approach in combination with bond lengths and angles, planarity etc. restraints and isotropic B-factors.[7] Detailed descriptions of the methodologies used in each individual case can be found in the original literature along with examples of the quality of the final electron density and tabulated crystal data, diffraction data collection and refinement statistics. Selected refinement parameters for a subset of the structures discussed in the review are listed in **Table S1** along with errors for atomic coordinates. The latter were estimated with the so-called diffraction-component precision index (DPI) introduced by Cruickshank,[8] using the program SFCHECK in the CCP4 suite.[9]

Table S1. Selected refinement parameters and DPI estimates for distances based on R-free for crystal structures containing dF (DNA) or rF (RNA).

Structure (ref.)	PDB ID	Resol. [Å]	R-work	R-free	r.m.s. bonds [Å]	r.m.s. angles [°]	DPI [Å]
RNA[1]	2Q1O	1.10	0.112	0.142	0.014	0.029*	0.006
RNA[2]	2G92	1.61	0.196	0.231	0.011	1.8	0.110
DNA:RNaseH[3]	3I8D	1.61	0.197	0.238	0.030	2.5	0.108
DNA:RB69[4]	3QEP	1.80	0.167	0.203	0.008	2.7	0.193
DNA:Dpo4[5]	2VA2	2.80	0.211	0.249	0.007	1.1	0.609

*1⋯3 distance

REFERENCES

1. Li, F.; Pallan, P. S.; Maier, M. A.; Rajeev, K. G.; Mathieu, S. L.; Kreutz, C.; Fan, Y.; Sanghvi, J.; Micura, R.; Rozners, E.; Manoharan, M.; Egli, M. Crystal Structure, Stability and in vitro RNAi Activity of Oligoribonucleotides Containing the Ribo-difluorotoluyl Nucleotide: Insights into Substrate Requirements by the Human RISC Ago2 Enzyme. *Nucleic Acids Res.* **2007**, *35*, 6424-6438.
2. Xia, J.; Noronha, A.; Toudjarska, I.; Li, F.; Akinc, A.; Braich, R.; Rajeev, K. G.; Egli, M.; Manoharan, M. siRNAs with a Ribo-difluorotoluyl Nucleotide: Structure, RISC-mediated Recognition, and Silencing. *ACS Chem. Biol.* **2006**, *1*, 176-183.
3. Pallan, P. S.; Egli, M. The Pairing Geometry of the Hydrophobic Thymine Analog 2,4-Difluorotoluene in Duplex DNA as Analyzed by X-ray Crystallography. *J. Am. Chem. Soc.* **2009**, *131*, 12548-12549.
4. Xia, S.; Eom, S. H.; Konigsberg, W. H.; Wang, J. Structural Basis for Differential Insertion Kinetics of dNMPs Opposite a Difluorotoluene Nucleotide Residue. *Bochemistry* **2012**, *51*, 1476-1485.
5. Irimia, A.; Eoff, R. L.; Pallan, P. S.; Guengerich, F. P.; Egli, M. Structure and Activity of Y-class DNA Polymerase Dpo4 from *Sulfolobus solfataricus* with Templates Containing the Hydrophobic Thymine Analog 2,4-Difluorotoluene. *J. Biol. Chem.* **2007**, *282*, 36421-36433.
6. Sheldrick, G. M.; Schneider, T. R. SHELXL: High-Resolution Refinement. *Meth. Enzymol.* **1997**, *277*, 319-343.
7. Murshudov, G .N.; Vagin, A. A.; Dodson, E. J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst. D* **1999***, 53*, 240-255.
8. Cruickshank, D. W. J. Remarks About Protein Structure Precision. *Acta Cryst. D* **1999**, *55*, 583-601.
9. Collaborative Computational Project, Number 4. The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst. D* **1994**, *50*, 760-763.