

Web 3DNA—a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures

Guohui Zheng¹, Xiang-Jun Lu^{1,2}, and Wilma K. Olson^{1,*}

¹Rutgers, the State University of New Jersey, Department of Chemistry & Chemical Biology, BioMaPS Institute for Quantitative Biology, Wright-Rieman Laboratories, 610 Taylor Road, Piscataway, NJ 08854 and ²Department of Biological Sciences, Center for Computational Biology and Bioinformatics, Columbia University, New York, NY 10027, USA

Received March 2, 2009; Revised April 14, 2009; Accepted April 22, 2009

ABSTRACT

The w3DNA (web 3DNA) server is a user-friendly web-based interface to the 3DNA suite of programs for the analysis, reconstruction, and visualization of three-dimensional (3D) nucleic-acid-containing structures, including their complexes with proteins and other ligands. The server allows the user to determine a wide variety of conformational parameters in a given structure—such as the identities and rigid-body parameters of interacting nucleic-acid bases and base-pair steps, the nucleotides comprising helical fragments, etc. It is also possible to build 3D models of arbitrary nucleotide sequences and helical types, customized single-stranded and double-helical structures with user-defined base-pair parameters and sequences, and models of DNA ‘decorated’ at user-defined sites with proteins and other molecules. The visualization component offers unique, publication-quality representations of nucleic-acid structures, such as ‘block’ images of bases and base pairs and stacking diagrams of interacting nucleotides. The w3DNA web server, located at <http://w3dna.rutgers.edu>, is free and open to all users with no login requirement.

INTRODUCTION

DNA and RNA contain layers of biological information, interspersed between or superimposed on the text written in the three-letter codes that provide instructions for making proteins. For example, the structure of the constituent nucleotides governs access to the sites on DNA and RNA targeted by enzymes and regulatory proteins. Understanding how the nucleic acids fold and how proteins and other ligands recognize and deform the 3D

structure are important for comprehending the dynamics of the cell.

Interest in understanding the relationship between the global folding of nucleic acids and the sequence-dependent arrangements of the constituent bases and base pairs has stimulated the development of new approaches to analyze and depict DNA and RNA structures. The characterization and visualization of such structures requires detailed knowledge of both the spatial disposition of the constituent bases and bases pairs and the conformation of the intervening sugar-phosphate backbone. Models that take advantage of this information are useful in the formulation of nucleic-acid binding ligands, the interpretation of various nucleic-acid configurational properties, etc.

The 3DNA suite of programs (1,2) was designed for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid-containing structures, including their complexes with proteins and other ligands. At its core, the software uses a simple matrix-based scheme (3–7) to calculate the complete set of rigid-body parameters that characterize the orientation and displacement of the base pairs, base-pair steps, and single-stranded nucleotide steps that make up a DNA or RNA structure. The description of structure is geometrically straightforward and the computation of parameters is mathematically rigorous (1,6,7), allowing for the exact construction of molecular models based on the derived parameters. Although the software has gained wide use in the scientific community over the past decade, its command-line-driven style is not especially user-friendly, for either novices, i.e. non-Linux/Unix users, or educational purposes.

Here, we report a new, web-based interface that offers user-friendly access to some of the most popular features of the 3DNA package, including: (i) the conformational analysis of arbitrary nucleic-acid-containing structures; (ii) the construction of nucleic-acid models from derived conformational parameters and classic fiber-diffraction models; and (iii) the visualization of local and global

*To whom correspondence should be addressed. Tel: +1 732 445 3993; Fax: +1 732 445-5958; Email: wilma.olson@rutgers.edu

nucleic-acid structure from novel and precisely controlled spatial perspectives. The server also contains a database of pre-analyzed nucleic-acid-containing structures stored in the Protein Data Bank (PDB) (8) and Nucleic Acid Database (NDB) (9) to facilitate user access. The data include conformational information for the asymmetric and biological units of crystal structures and the complete sets of structures determined in NMR studies. Treatment of individual molecular models or ensembles of simulated structures is also possible. The server functions robustly and includes a well-documented tutorial of the program functionalities. To the best of our knowledge, there are no other web servers with the same integrated structural-analysis, modeling, and visualization capabilities.

MATERIALS AND METHODS

Base coordinate frames

The rigid-body parameters commonly used to characterize the 3D arrangements of the bases and base pairs in a nucleic-acid structure quantify the pairwise orientation and displacement of local, orthogonal reference frames embedded in the constituent nucleotides. The set of parameters and the coordinate frames used in the 3DNA software follow established, community-developed guidelines (10,11). The software performs a least-squares fitting of a standard planar base structure with an embedded coordinate frame on its experimental counterpart, following the approach of Babcock *et al.* (12,13), to place the requisite reference frames on the bases in a structure.

Base-pair identification

The identification of interacting residues is based on the computed spatial disposition of the bases, in particular: (i) the distance d between the origins of the reference frames embedded in pairs of bases; (ii) the magnitude of the vertical offset of the base planes, the so-called Stagger (see text below); (iii) the angle Λ between the normals of the base planes; (iv) the distance d_{N1-N9} between the glycosidic base atoms, i.e. the purine N9 and pyrimidine N1 atoms linked to the sugar-phosphate backbone; and (v) the presence of one or more pairs of nitrogen/oxygen base atoms within a 'hydrogen-bonding' distance d_{HB} . The default values employed in the webserver calculations— $d \leq 15 \text{ \AA}$; Stagger $\leq 1.5 \text{ \AA}$; $\Lambda \leq 30^\circ$ or $\geq 150^\circ$; $d_{N1-N9} \geq 4.5 \text{ \AA}$; and $d_{HB} \leq 5.5 \text{ \AA}$ —identify both canonical and non-canonical nucleotides interactions (1,2,14), e.g. Watson-Crick (15), Hoogsteen (16), and other base pairs.

Rigid-body parameters

The w3DNA server reports three sets of rigid-body parameters: (i) the six base-pair parameters describing the spatial arrangements of associated bases—three angles called Buckle, Propeller, and Opening and three displacements called Shear, Stretch, and Stagger; (ii) the six base-pair-step parameters specifying the configurations of spatially adjacent base pairs—two bending angles called

Tilt and Roll, the dimeric rotation angle Twist, two in-plane dislocations termed Shift and Slide, and the vertical displacement Rise; and (iii) the six parameters that relate the positions of successive base pairs relative to a local helical frame—the angles Inclination and Tip and the distances x -displacement and y -displacement describing the orientation and translation of the base planes with respect to the helical axis, and the rotation about and displacement along the helical axis, referred to as Helical Twist and Helical Rise (1,10). The numerical values describe the deviations of the base pairs in a given structure from the planar Watson-Crick base pairs in an ideal B-DNA helix, where the base-pair parameters, the dimeric bending components, and in-plane dislocations of adjacent base pairs are null (11). A fourth set of rigid-body variables—the dinucleotide Tilt, Roll, Twist, Shift, Slide, and Rise—specifies the arrangements of adjacent bases along individual strands. The computations of rigid-body parameters use the mathematical definitions of El Hassan and Calladine (5). The identification of the helical axis between adjacent base pairs follows the methodology introduced by Babcock *et al.* (13).

The reported output also includes the areas of overlap of adjacent bases and base pairs and the positioning of phosphorus atoms within each base-pair step. The former values quantify the stacking of neighboring base pairs, and the latter discriminate between A and B double-helical steps (17). The base-pairing information is complemented by more conventional structural data, such as the identities and lengths of hydrogen bonds, the distances and angles between atoms in hydrogen-bonded and adjacent nucleotides, the torsion angles along the chain backbone, the amplitude and phase angle of sugar pseudorotation (i.e. puckering geometry), the glycosyl torsions orienting the sugars and bases, and the widths of the major and minor grooves.

WEBSERVER

Analysis component

The analysis component of the w3DNA server determines the aforementioned conformational parameters for the paired bases, stacked base pairs, and sequential bases in a user-uploaded, PDB-formatted coordinate file, i.e. the standard listing of chemical information and atomic positions reported for the atoms in a structure (see the RCSB PDB website for a detailed description). The input of a PDB/NDB ID, i.e. the identifiers used respectively in the Protein Data Bank and the Nucleic Acid Database to denote individual structures, yields the same information. A simple keyword/author search and pop-up links to the PDB and NDB search engines and to the NDB Atlas facilitate the selection of archived structures.

Output page. The output page, illustrated in Figure 1 for the structure, deduced from multidimensional heteronuclear NMR spectroscopic studies, of a 13 base-pair DNA duplex bound to the human TTAGGG-repeat binding factor TRF1 (PDB ID: 1IV6) (18), contains four sections: (i) a brief summary of the structure;

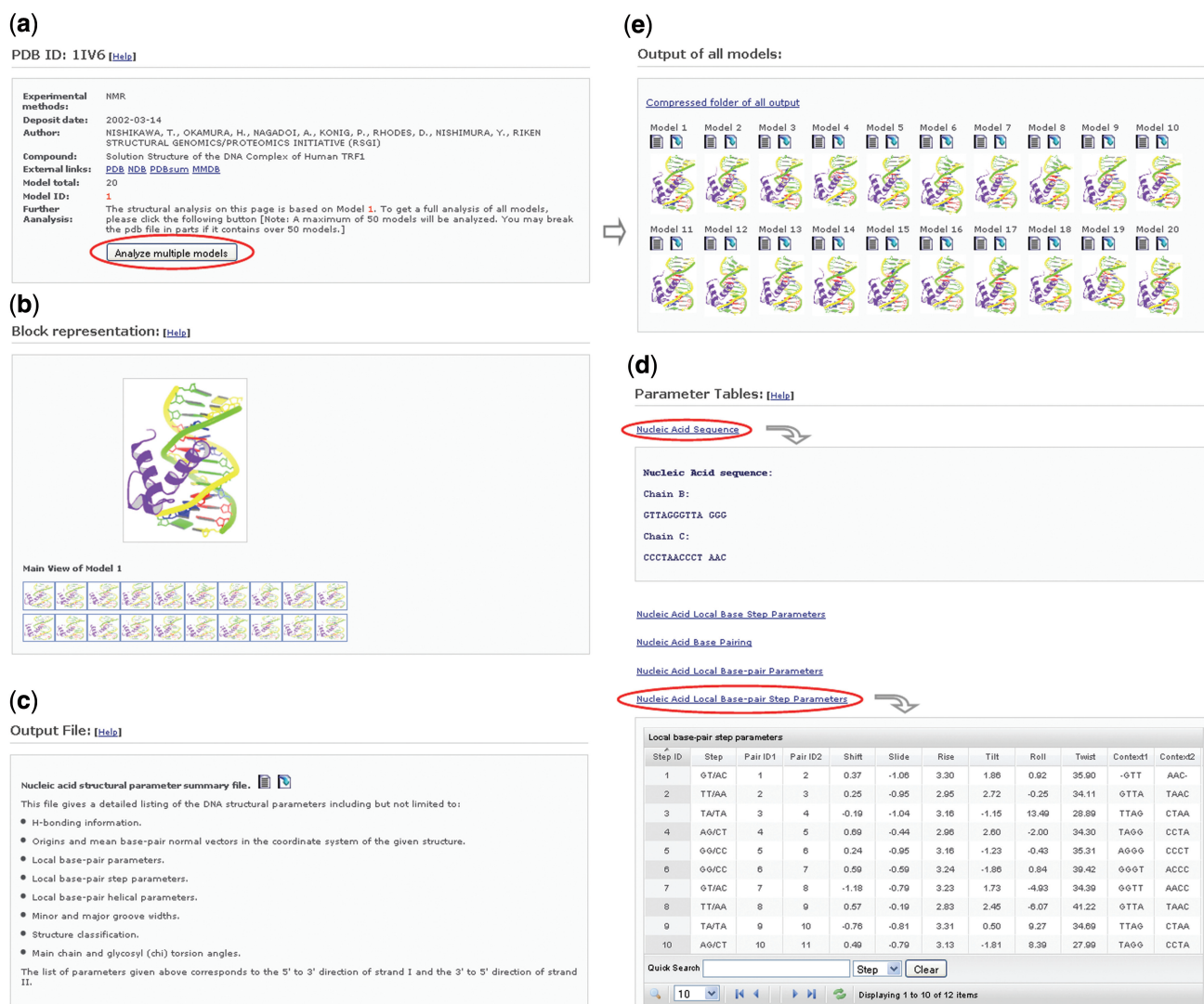


Figure 1. Screenshots illustrating the information provided in the analysis of nucleic-acid-containing structures with valid Protein Data Bank identifiers (PDB ID), here the ensemble of 20 structures of the complex of DNA with the human TTAGGG-repeat binding factor TRF1 determined by multidimensional heteronuclear NMR spectroscopy (18) (PDB ID: 1IV6). The output comprises, but is not limited to: (a) a brief description of the structural file, including the author(s), compound(s), number of models, external links, etc., (b) a gallery of block representations of each model in the file, which by moving the mouse over the icons, reveals the fluctuations in the structural ensemble and by clicking a specific icon, points to the set of parameters describing the chosen model, (c) a summary file with a comprehensive list of structural parameters, which can be viewed or downloaded by clicking the appropriate icons, (d) tables of selected parameters, which can be displayed by clicking on one of the links and sorted by clicking on the headers of the columns in the selected table, (e) the page, redirected from (a) via the 'Analyze multiple models' button, with links to the summary files for all of the models of the protein-DNA complex.

(ii) a schematic representation of the 3D fold; (iii) a link to the complete listing of 3DNA-derived parameters; and (iv) a set of interactive tables for selected parameters.

Structural summary. The structural summary includes the PDB ID and the NDB ID (if any), the methodology used to determine the structure, the resolution (if an X-ray structure), the deposition date (if the structural file is curated in the PDB or NDB), the author(s), the name of the compounds that make up the structure, and the links to several useful websites. If the input coordinate file contains more than one model, the summary also lists the number of models in the file, 20 in the case of

the TRF1-DNA complex presented in Figure 1, and provides a link that gives the user the option to analyze multiple models.

Structural representation(s). The structural representation on the output page is a composite image, with color-coded 'blocks' superimposed on the bases, an atomic depiction of backbone atoms, and color-coded tubes connecting the phosphorus atoms along individual strands. Proteins, if present, are represented by violet ribbons, and small molecules by ball-and-stick images. The same 3DNA-generated representations are found on the PDB and NDB websites. Each illustrated structure is

automatically projected in the plane containing the two longest principal axes of the nucleic-acid fragment, but can be viewed from different viewpoints as described below. These and all other molecular images generated on the webserver can be saved by clicking on the appropriate download link.

Files, like HIV6, with multiple models include a large gallery of small image icons depicting up to 50 structures in the file. Moving the mouse across different icons reveals the structural differences among the models. The location of the mouse determines the model that is enlarged on the output page. Clicking the icon generates the complete output for the selected structure. Icons of the same style allow the user to reorient and view the one model offered for most X-ray crystal structures in different principal-axes planes.

Derived parameters and interactive tables. The listing of derived parameters in the output file can be viewed on the web or downloaded. The 3DNA user's manual, found at <http://3dna.rutgers.edu>, includes a brief description of each type of parameter. The parameter tables contain information about base sequence, interactions, and structure. Users can click each link to show/hide contents. The composition of base pairs and the rigid-body parameters relating sequential and paired bases and neighboring base pairs are presented in interactive, Grid-View tables, with angles expressed in degrees and distances in Ångstrom units. Data can be sorted by pressing an arrow at the top of each column. A simple quick search facilitates the examination of long nucleotide fragments. The example in Figure 1 shows the information included in the table of local base-pair step parameters—the numerical identities and chemical composition of the first 10 of the 12 base-pair steps of TRF1-bound DNA, the rigid-body parameters describing each step, and the tetrameric sequence context in which the step occurs. The user can control the number of steps that are displayed, with up to 100 entries per page.

Reconstruction component

The reconstruction component allows the user to build 3D models of arbitrary sequence and helical type, including: (i) 55 different fiber-diffraction models of regular DNA, RNA, and hybrid DNA/RNA helices; (ii) customized single- and double-stranded structures with bases, base pairs, and base-pair steps arranged according to user-supplied rigid-body parameters; (iii) curved DNA structures constructed from fragments of canonical A-, B-, and C-type helices; and (iv) models of DNA 'decorated' at user-defined locations with proteins and other molecules in the arrangements found in known NMR and crystal complexes. The various structures provide useful starting points for atomic-level calculations.

Fiber-diffraction models. The 55 helical models include single-, double-, and multi-stranded structures based on the fiber-diffraction studies of Arnott and co-workers (19,20) (43 models), Alexeev *et al.* (21) (two models), van Dam and Levitt (22) (two models), and Premilat and Albers (23–28) (eight models). Model choices are

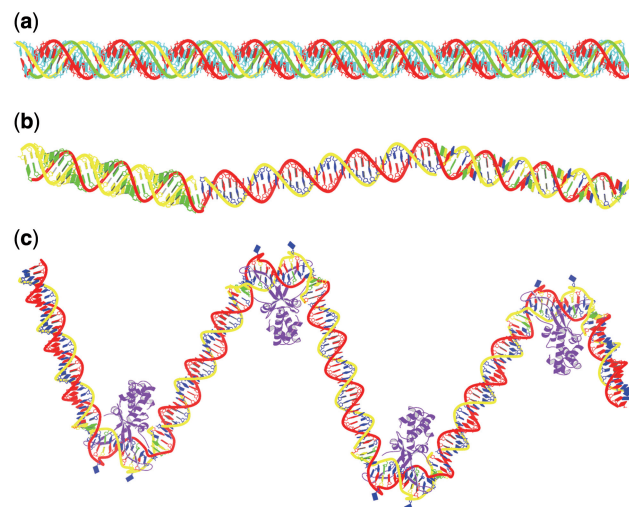


Figure 2. Representations of nucleic-acid-containing structures generated with the reconstruction component of the w3DNA server. (a) Two 100 nucleotide strands of poly(rU) complexed with the same length of poly(rA) in the classic 11-fold RNA poly(rU)-poly(rA)-poly(rU) triple helical structure (19,20) (3DNA fiber model 32). (b) A 100 base-pair curved DNA block copolymer made up respectively of A-, B-, and C-form double-helical fragments (19,20) (3DNA fiber models 1, 4, and 7) of $G_{35}\cdot C_{35}$, $A_{35}\cdot T_{35}$, and $(GA)_{15}\cdot (TC)_{15}$. (c) A 210 base-pair B-form DNA 'decorated' with 4 HU proteins. The protein-bound steps—centered at base pairs 41, 89, 137, and 185—are assigned the sequence and rigid-body parameters of the central 17 base pairs in the 1.90-Å crystal complex with *Anabaena* HU (29) (PDB ID: 1P71). The protein-free DNA steps are fixed in the canonical B form and assigned a homopolymeric repeating sequence $(A_n\cdot T_n)$, where n is respectively 31, 28, 28, 28, and 16 base pairs. Color-coded tubes on RNA and DNA trace the progression of the backbone defined by the phosphorus atoms: strand I (red); strand II (yellow); strand III (green). Color coding of nucleotide sequence conforms to the Nucleic Acid Database standard (9): A (red); C (yellow); G (green); T (blue). Violet ribbons connecting protein C^α carbons generated with MolScript (<http://www.avatar.se/molscript/>).

listed on a pull-down menu and described more fully in a table provided in the user tutorial. The models fall into two categories: generic helices that accommodate arbitrary base sequences of any length and non-generic helices that allow only the repetition of a pre-defined sequence. The example presented in Figure 2a is a non-generic, triple-helical RNA complex made up of two 100-nt fragments of poly rU and a fragment of poly rA of the same length, held in place by Watson-Crick (15) and Hoogsteen (16) A·U pairing. The collection of fiber models includes 39 DNA double- or triple-helical structures, 12 RNA single-, double-, triple-, or quadruple-helical structures, and 4 DNA-RNA hybrid duplexes

Customized models. The input files of sequence information and rigid-body parameters needed to generate customized nucleic-acid models are of two types, depending on the nature of the desired structure. The construction of a folded, single-stranded structure, such as one adopted by an RNA molecule, requires the set of base-step parameters describing the spatial disposition of successive nucleotides. Building a double-stranded structure entails detailed specification of both the base-pair parameters between interacting nucleotides and the base-pair-step parameters between stacked pairs.

Details of the necessary format are found on the tutorial page. The user selects the desired model type—either a full atomic model with an approximate, rigidly attached backbone or a model containing only base and P atoms, both in PDB format—from a pull-down menu.

The curved DNA pathways formed by the concatenation of regular A-, B-, and C-type models depend upon the chosen length, helical composition, and spacing of the structural components. For example, the slight zig-zag of the DNA duplex in Figure 2b reflects the opposing directions of dimeric bending and dislocation (Roll and Slide) in the A- and C-DNA fragments on either side of the central 35 base-pair stretch of B DNA. The all-atom backbones introduced in the model mirror the choice of helical types. The constructs accommodate any base sequence. The sequence within each conformational segment can be specified in two ways, as an arbitrary string of bases or as a string of repeated base-paired units.

Ligand-decorated DNA. The construction of protein- or ligand-decorated DNA models, such as the HU-bound DNA in Figure 2c, entails specification of the DNA chain length and sequence, the number of bound species, the locations at which the molecules are bound, and the requisite protein- or ligand-bound DNA structural templates, such as the crystal complex of DNA with *Anabaena* HU (29) (PDB ID: 1P71) shown here. The bound fragments adopt the conformational parameters of the selected complexes, specified by a PDB or NDB ID or uploaded as a customized PDB-formatted structure. The unbound DNA, including rigidly attached backbone atoms, assumes the user-selected helical form (A, B or C DNA). The binding positions correspond to the locations along the DNA of the central base pair or base-pair step of the chosen ligand-bound DNA structures. The location of the center point depends upon the length of the bound duplex, namely the middle base pair of a bound fragment with an odd number of bases pairs and the central base-pair step of a fragment with an even number of pairs. The software checks the user request for the potential overlap of proteins/ligands on the selected sequence and returns an error message if the proposed binding sites cover the same base pair(s). Only double-stranded structures can be treated and only all-atom models are generated. The DNA is built in two stages, with bases first positioned in accordance with the rigid-body parameters of the protein-bound and free chain segments and the atoms of the sugar-phosphate backbone and associated ligands subsequently superimposed on the base framework. The resulting models reveal the interdependence of the bound species, chosen sites of binding, unbound DNA conformation, and overall macromolecular fold. For example, the undertwisted HU-bound DNA binding sites must be spaced at non-integral helical turns along B-form DNA to generate a planar, zig-zag pathway like the one shown in Figure 2c.

Output features. All molecular constructs share the same three output features: (i) a composite representation of the overall structure in the above described principal-axis frame; (ii) a coordinate file in PDB format, which can be

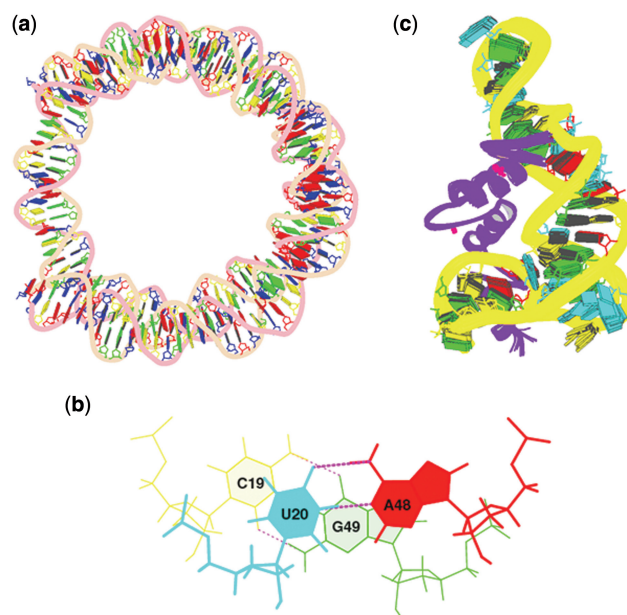


Figure 3. Examples of the unique representations of nucleic-acid structures available through the w3DNA server. (a) Color-coded composite block/backbone/tube representation of the DNA in the currently best-resolved nucleosome core-particle structure (31) (PDB ID: 1KX5). (b) Stacking diagram illustrating the overlap and hydrogen-bonding patterns of adjacent base pairs (C₁₉-G₄₉ and U₂₀-A₄₈) in model 1 from the ensemble of NMR structures of the complex of a 55 nucleotide fragment of *X. laevis* 5S rRNA with three zinc fingers of transcription factor TFIIIA (32) (PDB ID: 2HGH). (c) Schematic 'NMR-ensemble' image of the bases, RNA backbones, and protein ribbons in 15 of the 20 models of the aforementioned RNA-protein complex (32). Color coding identical to that in Figure 2, save for the depiction of U in aqua.

downloaded for further study; and (iii) a link to visualize the final structure via WebMol (30) (best done on a computer running a Java Runtime Environment) or Jmol (<http://www.jmol.org/>). In the interest of computational efficiency, models are limited in size to 1000 base pairs or 2000 nucleotides.

Visualization component

The visualization component creates vector-based drawings and scenes that can be rendered as raster-graphics images, allowing for easy generation of publication-quality figures. The server takes a user-uploaded PDB-formatted file or a PDB/NDB ID, and returns novel representations of the structure or parts of it. The images include: (i) composite block/tube/backbone representations of the type used to illustrate nucleosomal DNA (31) (PDB ID: 1KX5) in Figure 3a; (ii) stacking diagrams of associated base pairs like that shown for neighboring C-G and A-U pairs in Figure 3b; and (iii) composite block/ribbon/backbone representations of structural ensembles, such as the NMR-based models of the 5S RNA-TFIIIA complex (32) (PDB ID: 2HGH) depicted in Figure 3c.

Composite images. The composite images include informative color-coding of the nucleic acid. The user can

choose the parts of the structure to be plotted, such as the nucleic-acid atoms in the nucleosome complex in Figure 3a or the protein ribbons, and can rotate the structure as a whole by arbitrary amounts about one of the principal axes of the nucleic-acid structure. The axes—designated *x*, *y*, and *z*—correspond respectively to the directions of the longest, intermediate, and shortest principal axes of the system.

Stacking diagrams. The stacking diagrams depict the hydrogen bonds between paired bases and reveal the overlap and relative disposition of stacked bases. The associated base pairs are automatically oriented in a top-down view such that the long axis of the step is horizontal and the leading strand lies on the left of the image, i.e. the average (middle) base-pair plane of the step coincides with the plane on which the structure is projected. The software identifies all stacked base pairs in the file and provides a list of the identified steps. The user specifies the step of interest and whether the bases should be labeled, as in Figure 3b, or unlabeled in the diagram.

Ensemble visualization. The ensemble-visualization tool generates a composite block/ribbon/backbone image of a user-selected set of models in a file with multiple NMR-based or computer-generated structures, such as the 15 coordinate files of 5S RNA-TFIIA structures depicted in Figure 3c. The function returns error messages if applied to structures with a single model. The user selects the starting and ending models from a supplied list of model numbers.

Tutorial

The tutorial includes step-by-step instructions and worked-out examples to help the user take advantage of the available functions of the w3DNA server. The information pages address each of the functional categories—i.e. analysis, reconstruction, and visualization. The server also provides links to the 3DNA Forum, a website where users pose and respond to assorted questions dealing with the use and application of the software, and to additional citations for users interested in learning more about (i) the content and capabilities of the software, (ii) the standard coordinate frame used in the determination of rigid-body parameters, (iii) the conformational parameters typical of nucleic-acid structures, and (iv) the differences among programs used in the analysis of nucleic-acid structures.

CONCLUDING REMARKS

The w3DNA server provides straightforward access to some of the most popular features of the 3DNA suite of programs. The server integrates various 3DNA utilities to carry out the pre- and post-processing of data necessary for the analysis and presentation of nucleic-acid structural information.

Other new subroutines working in the background allow the user to search for and manipulate input files, analyze structural data, generate the coordinates of molecular models, display assorted images, and

manipulate tables on the fly. The various components make direct use of commands within w3DNA through graphic input options. The model reconstruction tools include new software for structure superposition and interactive visualization from multiple perspectives.

The server is intended for a broad range of users and educational purposes. Advanced users are encouraged to download the software package from the 3DNA web site and explore more of its functions.

ACKNOWLEDGEMENTS

We wish to thank Feng Cui, Luke Czapla, Thomas Gaillard, Yun Li, Difei Wang, Daniel Weingard, Fei Xu, Yurong Xin, and the anonymous reviewers for testing the server and offering valuable comments.

FUNDING

The U.S. Public Health Service (research grant GM20861, instrumentation grant RR022375) has generously supported this work. G.Z. acknowledges support of a pre-doctoral traineeship from the U.S. Public Health Service (New Interdisciplinary Research Workforce Training Grant DK071502). Partial funding for open access charge: GM20861 and the Mary I. Bunting Fund of Rutgers, the State University of New Jersey.

Conflict of interest statement. None declared.

REFERENCES

- Lu, X.-J. and Olson, W.K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, **31**, 5108–5121.
- Lu, X.-J. and Olson, W.K. (2008) 3DNA: a versatile, integrated software system for the analysis, rebuilding and visualization of three-dimensional nucleic-acid structures. *Nat. Protoc.*, **3**, 1213–1227.
- Zhurkin, V.B., Lysov, Y.P. and Ivanov, V.I. (1979) Anisotropic flexibility of DNA and the nucleosomal structure. *Nucleic Acids Res.*, **6**, 1081–1096.
- Bolshoy, A., McNamara, P., Harrington, R.E. and Trifonov, E.N. (1991) Curved DNA without A–A: experimental estimation of all 16 DNA wedge angles. *Proc. Natl Acad. Sci. USA*, **88**, 2312–2316.
- El Hassan, M.A. and Calladine, C.R. (1995) The assessment of the geometry of dinucleotide steps in double-helical DNA: a new local calculation scheme. *J. Mol. Biol.*, **251**, 648–664.
- Lu, X.-J., El Hassan, M.A. and Hunter, C.A. (1997) Structure and conformation of helical nucleic acids: analysis program (SCHNAaP). *J. Mol. Biol.*, **273**, 668–680.
- Lu, X.-J., El Hassan, M.A. and Hunter, C.A. (1997) Structure and conformation of helical nucleic acids: rebuilding program (SCHNArP). *J. Mol. Biol.*, **273**, 681–691.
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E. (2000) The Protein Data Bank. *Nucleic Acids Res.*, **28**, 235–242.
- 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) The Nucleic Acid Database: a comprehensive relational database of three-dimensional structures of nucleic acids. *Biophys. J.*, **63**, 751–759.
- Dickerson, R.E., Bansal, M., Calladine, C.R., Diekmann, S., Hunter, W.N., Kennard, O., von Kitzing, E., Lavery, R., Nelson, H.C.M., Olson, W.K. *et al.* (1989) Definitions and nomenclature of nucleic acid structure parameters. *J. Mol. Biol.*, **205**, 781–791.

11. Olson, W.K., Bansal, M., Burley, S.K., Dickerson, R.E., Gerstein, M., Harvey, S.C., Heinemann, U., Lu, X.-J., Neidle, S., Shakked, Z. *et al.* (2001) A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.*, **313**, 229–237.
12. Babcock, M.S. and Olson, W.K. (1994) The effect of mathematics and coordinate system on comparability and 'dependencies' of nucleic acid structure parameters. *J. Mol. Biol.*, **237**, 98–124.
13. Babcock, M.S., Pednault, E.P.D. and Olson, W.K. (1994) Nucleic acid structure analysis. Mathematics for local Cartesian and helical structure parameters that are truly comparable between structures. *J. Mol. Biol.*, **237**, 125–156.
14. Xin, Y. and Olson, W.K. (2009) BPS: a database of RNA base-pair structures. *Nucleic Acids Res.*, **37**, D83–D88.
15. Watson, J.D. and Crick, F.H.C. (1953) A structure for deoxyribose nucleic acid. *Nature*, **171**, 737–738.
16. Hoogsteen, K. (1963) The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallogr.*, **16**, 907–916.
17. Lu, X.-J., Shakked, Z. and Olson, W.K. (2000) A-form conformational motifs in ligand-bound DNA structures. *J. Mol. Biol.*, **300**, 819–840.
18. Nishikawa, T., Okamura, H., Nagadoi, A., König, P., Rhodes, D. and Nishimura, Y. (2001) Solution structure of a telomeric DNA complex of human TRF1. *Structure*, **9**, 1237–1251.
19. Chandrasekaran, R. and Arnott, S. (1989) In Saenger, W. (ed.), *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology, Group VII/1b, Nucleic Acids*. Springer, Berlin, pp. 31–170.
20. Arnott, S. (1999) In Neidle, S. (ed.), *Oxford Handbook of Nucleic Acid Structure*. Oxford University Press, Oxford, UK, pp. 1–38.
21. Alexeev, D.G., Lipanov, A.A. and Skuratovskii, I.Y. (1987) The structure of poly(dA)-poly(dT) as revealed by an X-ray fibre diffraction. *J. Biomol. Struct. Dynam.*, **4**, 989–1011.
22. van Dam, L. and Levitt, M.H. (2000) BII nucleotides in the B and C forms of natural-sequence polymeric DNA: a new model for the C form of DNA with 40° helical twist. *J. Mol. Biol.*, **304**, 541–561.
23. Premilat, S. and Albiser, G. (1983) Conformations of A-DNA and B-DNA in agreement with fiber X-ray and infrared dichroism. *Nucleic Acids Res.*, **11**, 1897–1908.
24. Premilat, S. and Albiser, G. (1984) Conformations of C-DNA in agreement with fiber X-ray and infrared dichroism. *J. Biomol. Struct. Dynam.*, **2**, 607–613.
25. Premilat, S. and Albiser, G. (1986) DNA models for A, B, C and D conformations related to fiber X-ray, infrared and NMR measurements. *J. Biomol. Struct. Dynam.*, **3**, 1033–1043.
26. Premilat, S. and Albiser, G. (1997) X-ray fibre diffraction study of an elevated temperature structure of poly(dA)-poly(dT). *J. Mol. Biol.*, **274**, 64–71.
27. Premilat, S. and Albiser, G. (1999) Helix-helix transitions in DNA: fibre X-ray study of the particular cases poly(dG-dC) and poly(dA) 2poly(dT). *Eur. Biophys. J.*, **28**, 574–582.
28. Premilat, S. and Albiser, G. (2001) A new D-DNA form of poly(dA-dT)-poly(dA-dT): an A-DNA type structure with reversed Hoogsteen pairing. *Eur. Biophys. J.*, **30**, 404–410.
29. Swinger, K.K., Lemberg, K.M., Zhang, Y. and Rice, P.A. (2003) Flexible DNA bending in HU-DNA cocrystal structures. *EMBO J.*, **22**, 3749–3760.
30. Walther, D. (1997) WebMol—a Java-based PDB viewer. *Trends Biochem. Sci.*, **22**, 274–275.
31. Davey, C.A., Sargent, D.F., Luger, K., Mäder, A.W. and Richmond, T.J. (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.*, **319**, 1097–1113.
32. Lee, B.M., Xu, J., Clarkson, B.K., Martinez-Yamout, M.A., Dyson, H.J., Case, D.A., Gottesfeld, J.M. and Wright, P.E. (2006) Induced fit and 'lock and key' recognition of 5 S RNA by zinc fingers of transcription factor IIIA. *J. Mol. Biol.*, **357**, 275–291.