3D reconstruction and comparison of shapes of DNA minicircles observed by cryo-electron microscopy

Arnaud Amzallag*, Cédric Vaillant, Mathews Jacob[1], Michael Unser[1], Jan Bednar[2], Jason D. Kahn[3], Jacques Dubochet[4], Andrzej Stasiak[4] and John H. Maddocks*

Laboratory for Computation and Visualization in Mathematics and Mechanics, EPFL FSB IMB, Ecole Polytechnique Fédérale de Lausanne CH-1015 Lausanne, Switzerland, [1]Biomedical Imaging Group, EPFL LIB, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, [2]Laboratoire de Spectrometrie Physique, UMR 5588, CNRS, 140 Av. de la Physique, BP 87, 38402 St Martin d'Heres Cedex, France, [3]Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20742-2021, USA and [4]Laboratory of Ultrastructural Analysis, University of Lausanne, 1015 Lausanne, Switzerland

ABSTRACT

We use cryo-electron microscopy to compare 3D shapes of 158 bp long DNA minicircles that differ only in the sequence within an 18 bp block containing either a TATA box or a catabolite activator protein binding site. We present a sorting algorithm that correlates the reconstructed shapes and groups them into distinct categories. We conclude that the presence of the TATA box sequence, which is believed to be easily bent, does not significantly affect the observed shapes.

INTRODUCTION

DNA structure in biological mechanisms

DNA base pair sequence is believed to influence the structure and deformability of the DNA double-helix (1,2), and thereby affects biological processes, such as DNA packaging (3), DNA loops in prokaryote regulatory complexes (4,5), nucleosome positioning (6–8) and DNA–protein interactions (9–11).

In the nucleosome, 147 bp of DNA wraps for almost two complete turns around histones [see Refs (12,13) for a crystal structure]. The radius of curvature of the double-helix axis is between 4 and 5 nm, whereas the DNA thickness is 2 nm. In this highly bent regime, the DNA sequence strongly affects the nucleosome position along the DNA molecule (6–8), and it is believed that this is due to DNA deformability.

In DNA–protein binding, the contribution of the DNA sequence to the binding affinity is frequently indirect, by opposition to direct recognition where nucleotide bases interact with amino acids through hydrogen bonds. For instance, the Bovine Papilloma Virus BPV-1 E2 binds two sites spaced by 4 bp (10). The DNA spacer does not contact the protein, but its sequence affects the stability of the DNA–protein complex (11). The role of the spacer in the stability of the DNA–protein complex has also been shown in the case of the cyclic AMP receptor protein (CRP) (14,15) also known as catabolite activator protein (CAP) (16).

The TATA box is a short sequence in the promoter region of genes that binds protein complexes and initiates transcription. Cyclization experiments (17–19) showed that free DNA in solution containing a TATA box sequence exhibits greatly enhanced J-factors that can be attributed to strong bends and high flexibility (9). It was shown that prebending of DNA enhances its interaction with TATA box binding protein (TBP) (20). As there are at most a few direct hydrogen bonds between DNA and the TBP (21,22), it is thought that the mechanical properties of the TATA box are probably very important for its function (9). This hypothesis is supported by a recent all-atom computation that predicts mostly indirect recognition between TBP and the TATA box (23).

Studies of DNA cyclization using DNA between 147 and 163 bp in length provided strong indications that the TATA box is highly flexible (9). However, the exact nature of this flexibility is not known. The sequence could behave as a kink and permit high local bending. On the other hand, the bending abilities of the TATA box could be approximately limited to the curvature expected for a 158 bp long DNA circle. A flexible kink should perturb the shape of the minicircle and be easily visible on cryo-electron microscopy (cryo-EM) images. However, if the bending is limited to a curvature similar to the one in the relaxed minicircle, it might not affect

the shape of observed minicircles. In this study, we observe and compare shapes of DNA minicircles of length 158 bp in which an 18 bp fragment contains either a TATA box or a CAP (CRP) site.

Cryo-electron microscopy of DNA minicircles

Cryo-EM allows the observation of DNA molecules in nearly physiological conditions; thin aqueous layers containing suspended DNA molecules are rapidly cooled and cryo-vitrified at such speed that ice crystals do not form (24–26). The frozen sample can be tilted, and one can obtain micrographs of individual DNA molecules visible from two different angles of view. This method has been used to reconstruct the 3D path of individual DNA molecules (27), and to determine DNA persistence length (28).

To observe how DNA shape is influenced by its sequence, it is advantageous to minimize variations due to thermal fluctuations and to visualize molecules as close as possible to their minimal energy shape. DNA minicircles seem to be best suited for this purpose. Because of their short length (close to the persistence length) the closure constraint of minicircles effectively limits the range of possible fluctuations.

On the other hand, the small size of the minicircles (∼17 nm in diameter) implies that even nanometer-size errors in the 3D reconstruction procedure significantly affect the reconstructed shapes. It is therefore desirable to use specialized software that can reconstruct the filaments with sub-pixel resolution (29).

As we wish to study sequence-dependent effects, it would be advantageous to know which point in our reconstructed center lines corresponds to which base pair of the sequence, which would require the use of a molecular marker. However, the intrinsic DNA shape can be altered by protein–DNA binding or by binding of specific chemicals that can be used to map specific sequences. For this reason we did not attempt such an approach, and instead visualize totally naked DNA.

MATERIALS AND METHODS

DNA constructs

Two DNA minicircles constructs are analyzed: t11T15 and c11T15 (9,19). The two minicircles are 158 bp long, and their sequences differ by 14 bp (Figure 1). The TATA box site in t11T15 is replaced by a CAP (CRP) binding site in c11T15 (16). Kahn and collaborators measured the cyclization rate of t11T15 and c11T15 sequences and determined their J-factors (9,19). The J-factor can be interpreted as a measure of the effective concentration of one DNA end in the vicinity of the other end, with orientations and helical twist that allows minicircle closure (30). The J-factor of the t11T15 minicircle is ∼3500 nM whereas the J-factor of c11T15 is 95 nM. The t11T15 and c11T15 fragments will be referred to as TATA and CAP, respectively, throughout the text.

Cryo-electron microscopy

The DNA minicircles were immobilized in a 50 nm layer of vitreous ice, at a temperature of −170°C. Images were taken

```
c11T15 (158 bp, J=95 nM):
GATGAATTCACGGATCCGGTTTTTTGCCCGTTTTTTTGCCGTTTTTTGCCC
GTTTTTTGCCGTTTTTTTGCCCGTTTTTTTCCGGATCCGTACAGGAATTCTA
GACCTAGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCC
ATGGAATC


t11T15 (158  bp, J~3500 nM):
GATGAATTCACGGATCCGGTTTTTTGCCCGTTTTTTTGCCGTTTTTTGCCC
GTTTTTTGCCGTTTTTTTGCCCGTTTTTTTCCGGATCCGTACAGGAATTCTA
GACCTAGGGTGCCTAATGAGTGCCCTTTTATAGCTTAAACGCGTTGCGCC
ATGGAATC
```

<div align="right">

Restriction site ClaI :
AT/CGAT
TATA box : CTTTTATA
A-tracts :   TTTTTT
</div>

Figure 1. The sequences of the studied minicircles. The differences between the two sequences are highlighted in red.

at the magnification of ×53 000 and registered on Kodak EM negative plates. The negatives were scanned at 1800 dpi, 8 bit gray-scale. The first image is taken with the sample tilted by −15° and the second at +15°. The tilt axis is vertical in both images presented in Figure 2.

3D reconstruction of DNA

We used the software package developed by Jacob *et al.* (29) to reconstruct the DNA minicircles shapes from the cryo-electron micrographs. For each minicircle the user traces an initial approximation of the visible DNA path on the two images. A smoothing filter of the images aids in this initial tracing. Our study is blind in the sense that the user does not know the sequence (TATA or CAP) of the minicircle in order to avoid bias in initial path tracing. Given this initial estimate, the program then performs the reconstructions by assuming a 3D curve model. The shape of the curve is optimized such that its 2D projections onto the micrograph planes match with the signals in the two images. The reconstructed curves are output in a list of points expressed in 3D Euclidean space. We re-sampled (using the spline function of Matlab) the output curves with a cubic spline to have 200 points per minicircle, equally spaced within one curve. We then analyzed the shape of 64 reconstructions of TATA and 31 of CAP.

Shape analysis and visualization

The main part of the code for data analysis was written in Matlab, with some Python scripts. Methods are described together with results in the next section. 3D pictures were produced with VMD (31).

RESULTS AND DISCUSSION

Curvature

Given three consecutive points A, B and C on a discrete curve, the curvature at B can be approximated by the inverse
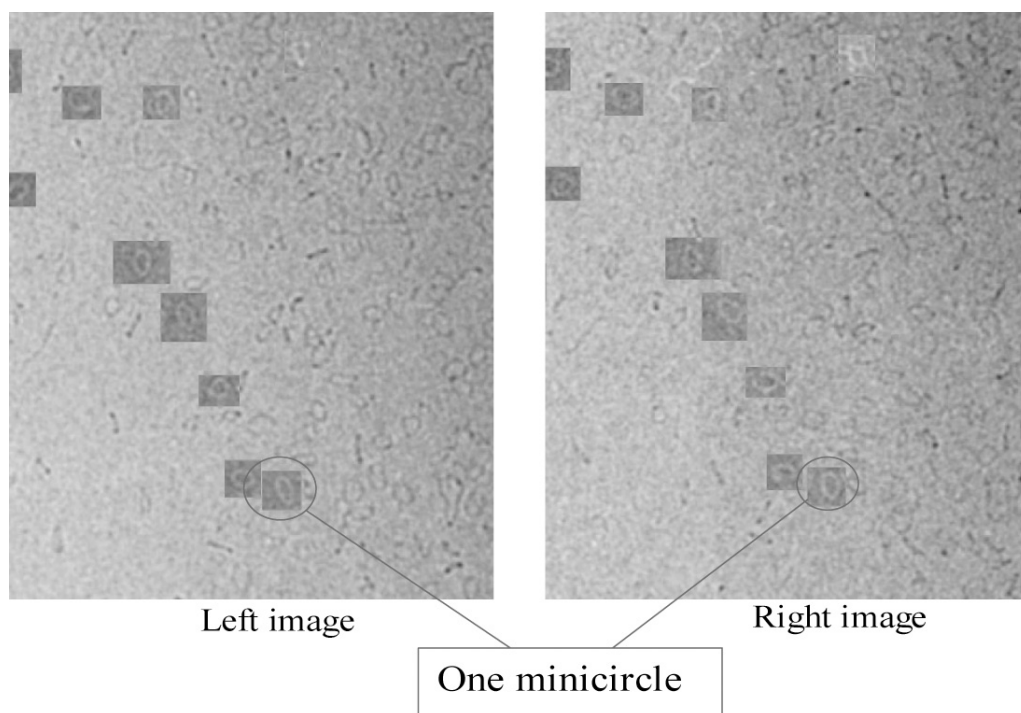
Figure 2. Regions of a pair of stereo micrographs. Stereo pairs of minicircles used for reconstruction are indicated with negative color frames.

of the radius of the circle that goes through A, B and C. A kink should induce high curvature in a short portion of the minicircle double-helix axis. We analyzed the distribution of such curvature values in the reconstructed minicircles. Each minicircle provided 200 entries for the curvature measured at each of the 200 indexed points. The curvature distributions of the points belonging to the TATA circles and to the CAP circles are computed separately and compared in Figure 3. For both sequences, the curvature distribution is peaked; the maximum corresponds to the curvature of a 158 bp prefect circle (0.12 nm^{-1}). The distribution of curvature is very similar for both TATA and CAP (Figure 3). Note that the shape data have no reference that indicates the location of the TATA box or CAP (CRP) site sequences.

Superposition of DNA minicircles shapes along their principal axes of inertia

Figure 4 shows axial paths of reconstructed minicircles that have been translated and rotated so that their center of mass (assuming uniform mass density), and their principal axes of inertia coincide. Such a presentation allows us to visually compare many minicircle shapes at the same time. The resulting picture does not show a clear difference between the shapes of TATA (red) and CAP (blue) minicircles.

The shape-distance for curves: minimum RMSD over all rigid-body motions, index shifts and curve orientations

Although curvature analysis and visualization did not reveal the presence of a kink in TATA in comparison to CAP minicircles, there may be a more subtle sequence-dependent shape
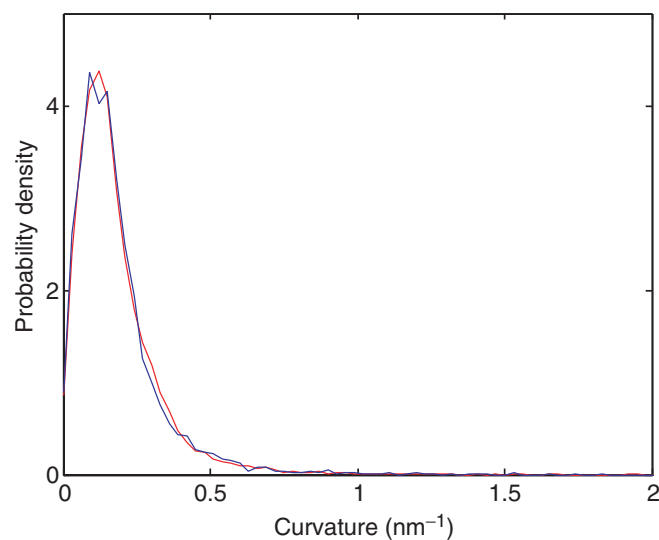

Figure 3. Probability density function of curvature in reconstructed TATA and CAP minicircles (the corresponding profiles are red and blue, respectively). The function is approximated by a normalized histogram counting curvature values within intervals of 0.03 nm^{-1}.

pattern. Therefore, rather than looking for a particular shape, we designed a method to identify groups of similar shapes, and looked whether the sequence correlates with the groups or not. We first chose a distance for the determination of shape similarity, then we clustered the shapes according to their mutual similarities measured in terms of this distance.

Because we do not know the correspondence between the sequence and the curve in each image, in order to estimate the
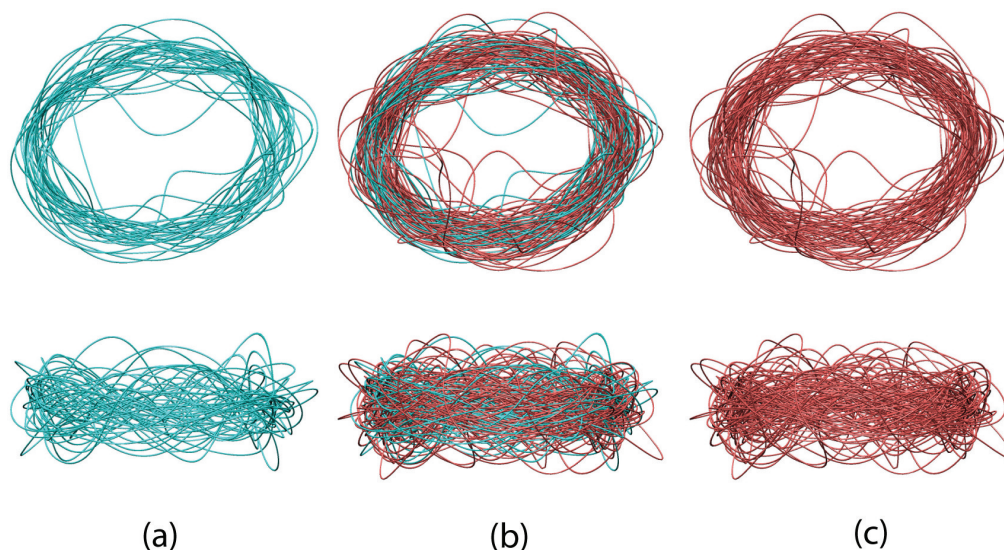
Figure 4. All the reconstructed shapes of DNA are aligned by superposition of their principal axes of inertia. The upper and lower views differ by a rotation of 90° around the horizontal axis: (**a**) all the 31 CAP minicircles (blue), (**b**) all the 95 minicircles and (**c**) all the 64 TATA minicircles (red).

similarity between two minicircle shapes we need to adapt the standard root mean square deviation (RMSD) minimization procedure that is often used to compare the geometries of two solid objects. The standard method is as follows: for two ordered sets of N points $\mathbf{x}$ and $\mathbf{y}$, RMSD is the square root of the sum over i of the squares of the Euclidean distances between two corresponding points $\mathbf{x}_i$ and $\mathbf{y}_i$. Then, to eliminate rigid-body motions, one computes a 3×3 rotation matrix $\mathcal{U}$ and a translation vector $\mathbf{r}$ which, when applied to $\mathbf{x}$, minimizes the RMSD function defined in Equation 1, producing the best superposition of the two structures:

$$\min_{\mathcal{U}, \mathbf{r}} \left[\frac{1}{N} \sum_{i=1}^{N} |\mathcal{U}\mathbf{x}_i + \mathbf{r} - \mathbf{y}_i|^2 \right]^{1/2}. \qquad \mathbf{1}$$

A Fortran 95 code given in (32) was used to compute this minimum RMSD.

Our shape-distance function is then defined in Equation 2 via minimization over all possible rigid-body rotations and translations in 3D, plus further minimizations in all shifts of an index (the variable δ), and two curve orientations, clockwise or counter-clockwise (the variable α):

$$\min_{\mathcal{U}, \mathbf{r}, \delta, \alpha} \left[\frac{1}{N} \sum_{i=1}^{N} |\mathcal{U}\mathbf{x}_i + \mathbf{r} - \mathbf{y}_{(-1)^{\alpha}i+\delta}|^2 \right]^{1/2}, \qquad \mathbf{2}$$

$$\delta \in \{1, \ldots, N\}, \quad \alpha \in \{1, 2\}.$$

The additional minimization over δ is necessary in our case because we do not know which point of the discretized curve $\mathbf{y}$ should correspond to the first point of the curve $\mathbf{x}$. However, if there is a common pattern between shapes of minicircles, a particular mapping of $\mathbf{x}$ onto $\mathbf{y}$ should give a minimal RMSD. The minimization over δ in Equation 2 allows all possible phasing differences in index to compete in the fit. Minimization over α recognizes that a given curve can be discretized with two distinct orientations.

Except for particular symmetrical shapes, identical curves that happen to be discretized with opposite orientations cannot be perfectly superposed by standard RMSD.

As a matter of implementation the additional minimizations in Equation 2 were achieved by calling the RMSD function given in (32) inside a Matlab loop for all possible shifts δ ($\delta = 1, \ldots, 200$ in our data with the index of $\mathbf{y}$ to be understood *modulo* 200), and the two choices of α. The smallest RMSD value found in the loop defines the distance between the two shapes.

Error of reconstruction measurements

To measure similarity or dissimilarity between different reconstructed minicircles it is important to determine the error of reconstruction and to see how much this error could affect the comparison between different reconstructed minicircles. In order to estimate the reconstruction error, we applied our distance function to two reconstructed shapes coming from the same image pair, but obtained by two different users of the reconstruction program. We computed the user error for six image pairs (Figure 5). We find that the average error is 0.9 nm, with SD 0.3 nm.

Analysis of shape-distances with respect to TATA and CAP sequences

We analyzed a set of 95 distinct minicircles (64 TATA, 31 CAP) all reconstructed by the same user. We therefore have a set of 4465 (or 95 ∗ 94/2) pairwise shape-distances. Figure 6 gives the normalized histograms, i.e. probability distributions of pairwise distances in three groups: TATA to TATA, CAP to CAP and TATA to CAP. The average shape-distances are 2.03 nm for TATA–TATA (SD 0.57 nm), 1.96 nm for CAP–CAP (SD 0.52 nm) and 1.98 nm for TATA–CAP (SD 0.55 nm). TATA–TATA and CAP–CAP shape-distances are not significantly smaller than TATA–CAP distances. Therefore, we do not observe
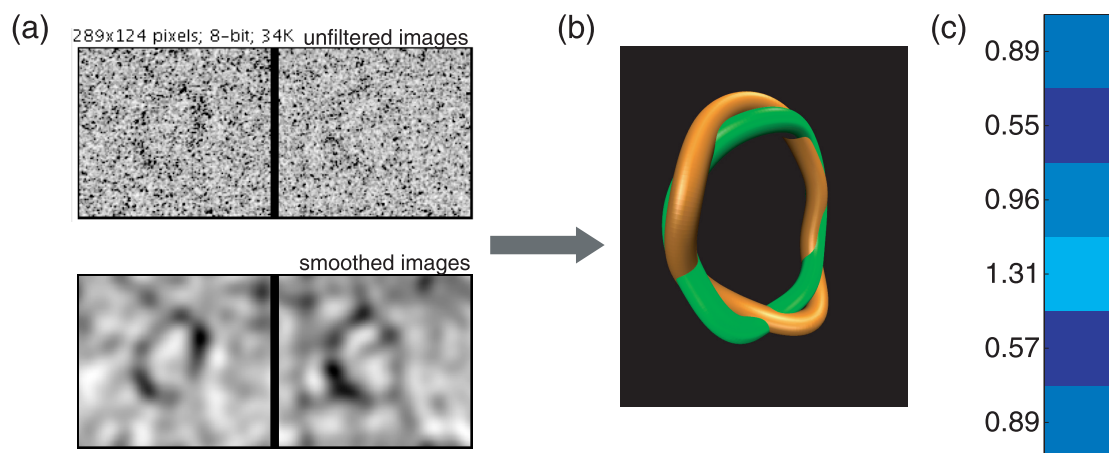
Figure 5. Estimation of the error of reconstruction. (**a**) The same DNA minicircle is shown from two different angles. In the right image, the sample is rotated by 30° around the vertical axis with respect to the left image. (**b**) Two reconstructions from the stereo-pair in (a) starting from two different user initializations. The shape-distance between the two reconstructions is 0.89 nm. (**c**) Error values, i.e. shape-distances between two reconstructions of each of six stereo-pairs. To allow comparison with shape-distances between different data, the error values are expressed with the color code of Figure 7.

increased shape similarity between minicircles with the same sequence.

Shape clustering

We cannot use classical methods for clustering our shapes, as we do not have a sensible way to represent them as vectors in a multidimensional space. We also do not have reference shapes to build clusters. Accordingly we adopt the reference-free SPIN algorithm (33) that is capable of ordering elements of a set using only their pairwise distances. For an ordered list of shapes and a shape-distance function, there exists a unique shape-distance matrix defined as follows: each element (i, j) of the matrix is the shape-distance between minicircles i and j. By definition, the matrix is symmetric and the elements on the diagonal vanish; the i-th line (or column) is a list of the distances between minicircle i and all others. SPIN finds a permutation of an initial ordered list of shapes that minimizes the elements near the diagonal. If the resulting matrix has a block of low (dark blue) values near the diagonal, with comparatively higher values above and below (and therefore necessarily by symmetry to right and left), the shapes in the block can be considered as clusters. A SPIN sorted shape-distance matrix and the corresponding clusters are represented in Figure 7. Three columns were added on the left of the matrix. They show some properties of the shapes. Each line and each column of the matrix correspond to a minicircle. For each line i of the matrix, the corresponding element i of the column 'Minicircle type' shows whether the corresponding minicircle i is of type TATA (gray) or CAP (white). It is clear that the TATA and CAP minicircles are spread throughout each cluster. Similarly, the i-th element of the column 'Circle' (respectively 'Ellipse') shows the distance between the minicircle i and a circle (respectively an ellipse). The circle diameter is 17.1 nm (corresponding to a perimeter of 158 bp). The longer ellipse axis is also 17.1 nm while the shorter axis is 13.7 nm. These two columns and the lower part of Figure 7 suggest that the method was able to identify clusters of circular and
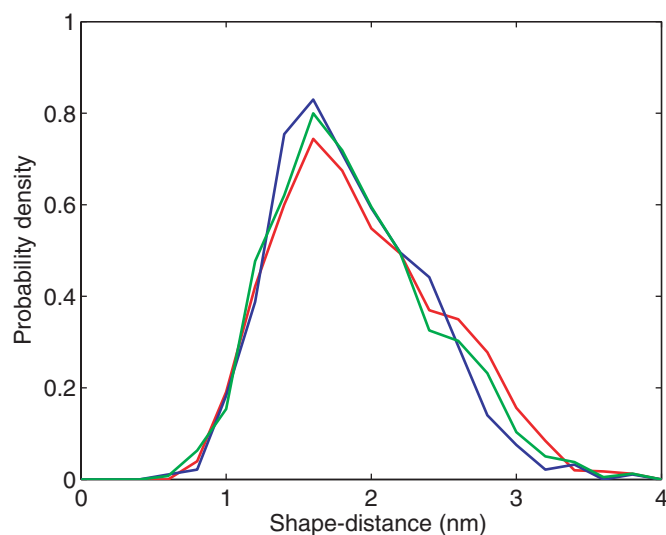

Figure 6. Normalized histograms (i.e. probability density) of the shape-distance values between any two TATA minicircles reconstructed shapes (red), any two CAP shapes (blue) and between one TATA shape and one CAP shape (green).

ellipsoid shapes, and to find another non-planar cluster. Stereo images of the cluster 7–15 are presented in Figure 8.

Interestingly, the distance matrix apparently reveals presence of multiple clusters of shapes. It is known that DNA circles with non-uniform sequence have multiple local energy minima (34). For this reason, we believe that our clustering analysis detected sampling of at least two and possibly more energy wells in the configuration space. However, the small difference between the majority of the clusters (comparable with the error of the reconstruction method) warns against over-interpretation of the distance matrix data. Importantly, each detected cluster contains both TATA and CAP minicircles, so that the different clusters seem to be associated with the sequence-dependent features that are shared between the two sequences, e.g. the six phased A-tracts,
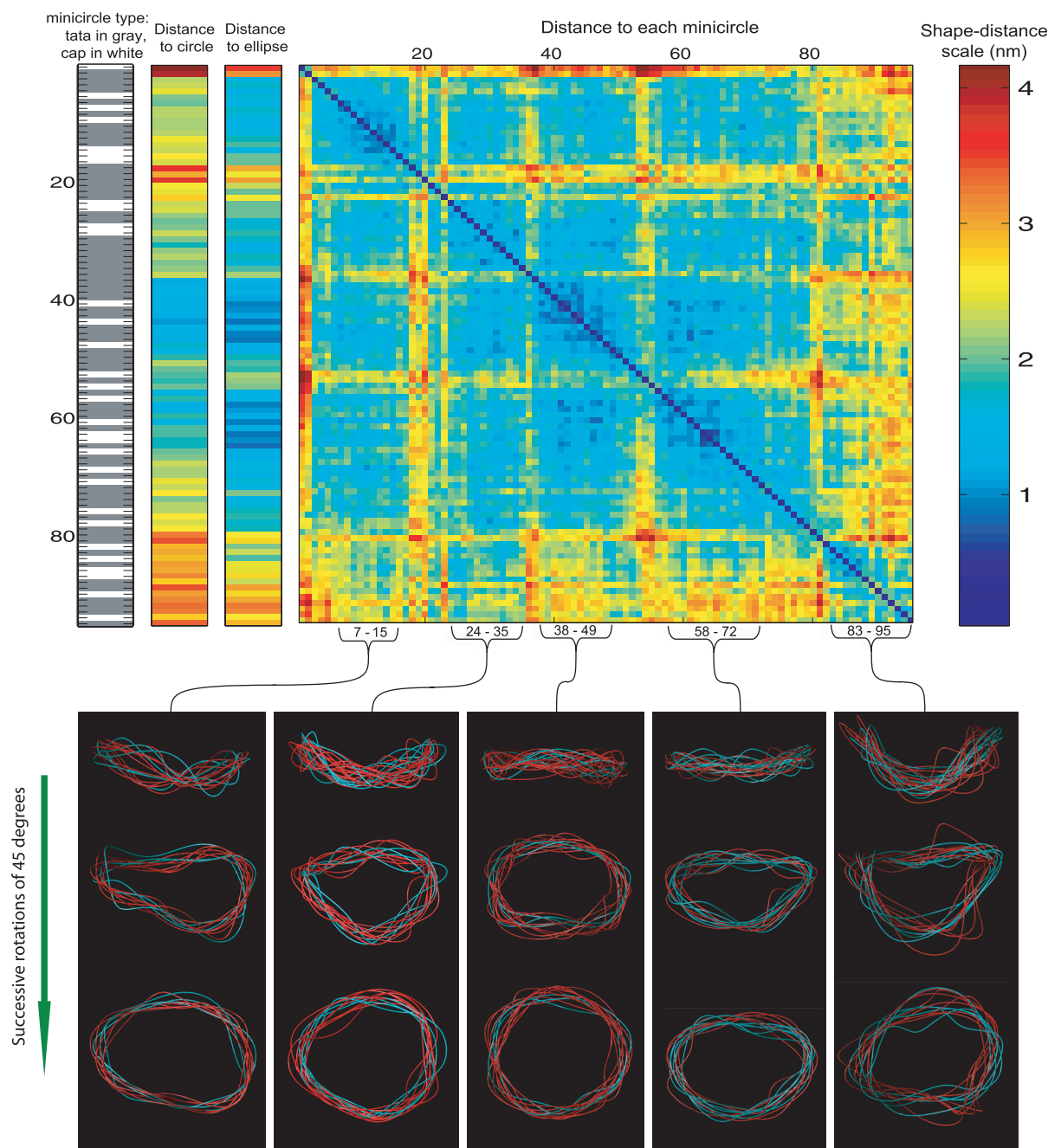
Figure 7. (Upper panel) The shape-distance matrix after clustering. Each line of the figure contains information about one minicircle reconstructed shape: the sequence type (first column), the shape-distance between the shape and a perfect circle (second column), between the shape and an ellipse (third column), and between the shape and every other reconstructed shape in the set (matrix). Shape-distance values are represented by colors with scale in nanometers shown on the right. (Lower panel) Different views of the clusters are produced by successive rotations of 45° around the horizontal axis. The indices of the minicircles shown are indicated with the brackets below the matrix. TATA (resp. CAP) minicircles are colored in red (resp. blue).

rather than the differences between TATA and CAP sequences. We therefore conclude that TATA and CAP sequences produce minicircles with similar 3D shapes.

CONCLUSION

Using cryo-EM we have investigated the effect on the 3D shape of 158 bp long DNA minicircles with identical

sequences except for the interchange of TATA and CAP boxes. Although, the TATA minicircles cyclize two orders of magnitude more efficiently than CAP in ligation experiments, we did not detect significant differences in the observed 3D shapes. Analysis of the reconstruction errors revealed that the average user error (0.9 nm) was two times smaller than the average shape-distance between two minicircles (2 nm). We conclude, therefore, that thermal
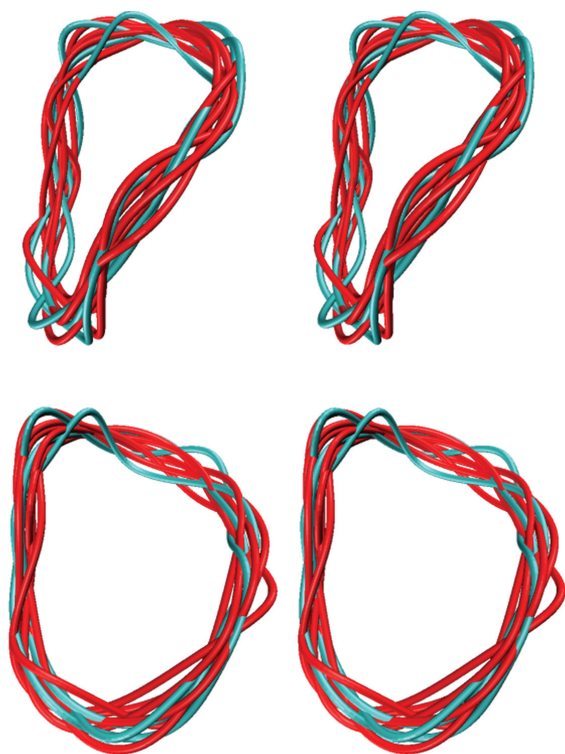
Figure 8. Stereo images of the cluster 7–15. Images are presented in 'side by side' stereo mode.

fluctuations 'blur' the possible differences in 3D shapes of DNA minicircles induced by the presence of CAP or TATA sequences.

ACKNOWLEDGEMENTS

REFERENCES

1. 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.
2. Olson,W.K., Gorin,A.A., Lu,X.J., Hock,L.M. and Zhurkin,V.B. (1998) DNA sequence-dependent deformability deduced from protein–DNA crystal complexes. *Proc. Natl Acad. Sci. USA*, **95**, 11163–11168.
3. Tolstorukov,M.Y., Virnik,K.M., Adhya,S. and Zhurkin,V.B. (2005) A-tract clusters may facilitate DNA packaging in bacterial nucleoid. *Nucleic Acids Res.*, **33**, 3907–3918.
4. Schleif,R. (1992) DNA looping. *Annu. Rev. Biochem.*, **61**, 199–223.
5. Vilar,J.M.G. and Leibler,S. (2003) DNA looping and physical constraints on transcription regulation. *J. Mol. Biol.*, **331**, 981–989.
6. Widlund,H.R., Cao,H., Simonsson,S., Magnusson,E., Simonsson,T., Nielsen,P.E., Kahn,J.D., Crothers,D.M. and Kubista,M. (1997) Identification and characterization of genomic nucleosome-positioning sequences. *J. Mol. Biol.*, **267**, 807–817.
7. Anderson,J.D. and Widom,J. (2000) Sequence and position-dependence of the equilibrium accessibility of nucleosomal DNA target sites. *J. Mol. Biol.*, **296**, 979–987.
8. Virstedt,J., Berge,T., Henderson,R.M., Waring,M.J. and Travers,A.A. (2004) The influence of DNA stiffness upon nucleosome formation. *J. Struct. Biol.*, **148**, 66–85.
9. Davis,N.A., Majee,S.S. and Kahn,J.D. (1999) TATA box DNA deformation with and without the TATA box-binding protein. *J. Mol. Biol.*, **291**, 249–265.
10. Hegde,R.S., Grossman,S.R., Laimins,L.A. and Sigler,P.B. (1992) Crystal structure at 1.7 Å of the bovine papillomavirus-1 E2 DNA-binding domain bound to its DNA target. *Nature*, **359**, 505–512.
11. Zhang,Y., Xi,Z., Hegde,R.S., Shakked,Z. and Crothers,D.M. (2004) Predicting indirect readout effects in protein–DNA interactions. *Proc. Natl Acad. Sci. USA*, **101**, 8337–8341.
12. Luger,K., Mader,A.W., Richmond,R.K., Sargent,D.F. and Richmond,T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, **389**, 251–260.
13. Richmond,T.J. and Davey,C.A. (2003) The structure of DNA in the nucleosome core. *Nature*, **423**, 145–150.
14. Ivanov,V.I., Minchenkova,L.E., Chernov,B.K., McPhie,P., Ryu,S., Garges,S., Barber,A.M., Zhurkin,V.B. and Adhya,S. (1995) CRP-DNA complexes: inducing the A-like form in the binding sites with an extended central spacer. *J. Mol. Biol.*, **245**, 228–240.
15. Emmer,M., deCrombrugghe,B., Pastan,I. and Perlman,R. (1970) Cyclic AMP receptor protein of *E.coli*: its role in the synthesis of inducible enzymes. *Proc. Natl Acad. Sci. USA*, **66**, 480–487.
16. Gartenberg,M.R. and Crothers,D.M. (1988) DNA sequence determinants of CAP-induced bending and protein binding affinity. *Nature*, **333**, 824–829.
17. Shore,D., Langowski,J. and Baldwin,R.L. (1981) DNA flexibility studied by covalent closure of short fragments into circles. *Proc. Natl Acad. Sci. USA*, **78**, 4833–4837.
18. Shore,D. and Baldwin,R.L. (1983) Energetics of DNA twisting. I. Relation between twist and cyclization probability. *J. Mol. Biol.*, **170**, 957–981.
19. Kahn,J.D. and Crothers,D.M. (1992) Protein-induced bending and DNA cyclization. *Proc. Natl Acad. Sci. USA*, **89**, 6343–6347.
20. Parvin,J.D., McCormick,R.J., Sharp,P.A. and Fisher,D.E. (1995) Pre-bending of a promoter sequence enhances affinity for the TATA-binding factor. *Nature*, **373**, 724–727.
21. Juo,Z.S., Chiu,T.K., Leiberman,P.M., Baikalov,I., Berk,A.J. and Dickerson,R.E. (1996) How proteins recognize the TATA box. *J. Mol. Biol.*, **261**, 239–254.
22. Kim,Y., Geiger,J.H., Hahn,S. and Sigler,P.B. (1993) Crystal structure of a yeast TBP/TATA-box complex. *Nature*, **365**, 512–520.
23. Paillard,G. and Lavery,R. (2004) Analyzing protein–DNA recognition mechanisms. *Structure (Camb.)*, **12**, 113–122.
24. Dubochet,J., Adrian,M., Chang,J.J., Homo,J.C., Lepault,J., McDowall,A.W. and Schultz,P. (1988) Cryo-electron microscopy of vitrified specimens. *Q. Rev. Biophys.*, **21**, 129–228.
25. Dubochet,J., Adrian,M., Dustin,I., Furrer,P. and Stasiak,A. (1992) Cryoelectron microscopy of DNA molecules in solution. *Meth. Enzymol.*, **211**, 507–518.
26. Adrian,M., tenHeggeler-Bordier,B., Wahli,W., Stasiak,A.Z., Stasiak,A. and Dubochet,J. (1990) Direct visualization of supercoiled DNA molecules in solution. *EMBO J.*, **9**, 4551–4554.
27. Dustin,I., Furrer,P., Stasiak,A., Dubochet,J., Langowski,J. and Egelman,E. (1991) Spatial visualization of DNA in solution. *J. Struct. Biol.*, **107**, 15–21.
28. Bednar,J., Furrer,P., Katritch,V., Stasiak,A.Z., Dubochet,J. and Stasiak,A. (1995) Determination of DNA persistence length by cryo-electron microscopy. Separation of the static and dynamic contributions to the apparent persistence length of DNA. *J. Mol. Biol.*, **254**, 579–594.

29. Jacob,M., Blu,T., Vaillant,C., Maddocks,J. and Unser,M. (2006) 3D shape estimation of DNA molecules from stereo cryo-electron micro-graphs using a projection-steerable snake. *IEEE Trans. Image Process.*, **15**, 214–227.

30. Crothers,D.M., Drak,J., Kahn,J.D. and Levene,S.D. (1992) DNA bending, flexibility, and helical repeat by cyclization kinetics. *Methods Enzymol.*, **212**, 3–29.

31. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD—visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38.

32. Coutsias,E.A., Seok,C. and Dill,K.A. (2004) Using quaternions to calculate RMSD. *J. Comput. Chem.*, **25**, 1849–1857.

33. Tsafrir,D., Tsafrir,I., Ein-Dor,L., Zuk,O., Notterman,D.A. and Domany,E. (2005) Sorting points into neighborhoods (SPIN): data analysis and visualization by ordering distance matrices. *Bioinformatics*, **21**, 2301–2308.

34. Furrer,P.B., Manning,R.S. and Maddocks,J.H. (2000) DNA rings with multiple energy minima. *Biophys. J.*, **79**, 116–136.