Usability of the kink parameters for nucleic acid structure in
Database

Satoshi Fujii
School of Pharmaceutical Sciences, University of Sizuoka, 52-1 Yada, Suruga-ku, Shizuoka-shi, Shizuoka 422-8526, Japan

ABSTRACT

DNA/RNA molecules with the specific three-dimensional structure, express the specific structural and biological functions. The knowledge integration based on three dimensional structures can be used to explain biochemical observations, to predict biological functions and to design drugs specific to a given complex system. The database cataloging the interaction motifs of nucleic acid moieties has been developed. Skew matrix, a kind of kink parameter, affords the structural description between the adjacent moieties. The proper values in skew matrix have the good properties, i.e., impregnable and flexible presentation. The geometrical parameters about hydrogen bond and base stacking would also provide the tolerant aspect for stereochemical bioinformatics. The tentative database including these parameters with the species and physical properties of the surrounding nucleic acid components and amino acids has been constructed.

RESULTS AND DISCUSSION

Coordinates and other database: The initial coordinates have been obtained from the Protein Data Bank. The NDB (Nucleic Acid Database) provides the structural data about nucleic acid (2). This project also makes the software available to produce reports describing any of the stored properties of any subset of the structures in the database. CIF is an acronym for the X-ray crystallographic information file and the CIF format is suitable for archiving, in any order, all types of text and numerical data. The mmCIF was developed for the macromolecular crystallographic experiment by NDB project.

Calculation: The two geometrical parameters, helical parameters and kink parameters were calculated by program BIOCON. For the kink parameter, the skew matrix (3x3) and the location of grass root of matrix for adjacent function groups are calculated. The helical parameters which have been defined in the Cambridge convention (3) are derived from the spatial location of the bases, while the sugar phosphate backbone is not taken into account. The quantitative analysis of bending of the helical axis is done and orientational and shifting parameters are calculated.

Comparison of geometrical parameters: Skew matrix, affords the structural description between the adjacent groups in any irregular geometry. The relative rotational and positional features are represented and can be converted to the other coordinate system. Essential angular components of base stacking, tilt, roll and twist angles may be a preferable information to cite the curved or bend phenomenon in not only the nucleic acid but also the interaction with protein molecules.

Model building: To test and estimate the confidence in inverse process, to render the structure feature, model buildings using with two kinds of parameters were carried. The conformational flexibility of torsion angle in nucleic
acid is great and its domain of the permitted values would be virtually infinite. The limitations or simplifications of typical conformation for inter- and intranucleotide torsion angles by taking only representative values, were considered to generate the structure. However, the torsion angle is not so convenient to generate the model, because the several combinations of torsion angle afford the quite similar global structure and the small value change of torsion angle causes reversely the serious structure change and it is difficult to image the global structure from the torsion angles. In stead of torsion angle, the helical parameter and kink parameter are elected as the input term to generate the first global structure. It is easy to image the global structure from these parameters and to obtain the values from the conventional model such as HGS model or literature. New conventional model building program, name of which is HLX6, can be successfully applied to the DNA/RNA duplex with the nobel G:A base pairing and triplex with the modified sugar moiety. (These results would be shown in Symposium session.)

The building performance in our procedure is threefold. (i) Coarse connectivity check. Input terms are basic geometrical information such as base pairing scheme and also helical parameters or kink parameters. The glyosyl torsion angle and sugar ring puckering mode also are treated as independent values. These parameters create the consecutive bending array of helix axis. And then the generation of the stranded structure is accomplished in accordance with allowed helical parameters. The resulting O3'...C4' distance, of which O3' and C4' atoms are belonged to 3'-link and 5'-link moiety, respectively, is used as selection determinant. Algorithm to reduce the search region from initial glossary definitions is considered. (ii) The further conformer space search is carried with fine step among the above selected domain of variable values. Algorithm Metropolis is constructed to reduce computer time in this procedure. (iii) Energy optimization is carried by the convenient calculations to manipulate the structure maturation.

Database: The calculated kink parameters are written in mmcIF format using data description language. The program BIOCON also lists the hydrogen bonding between base moiety. In the case of non-canonical base pairing, the hydrogen bonding mode, such as atom name of hydrogen donor or acceptor, is written in mmcIF format. The hydrogen bonds, salt bridges and hydrophobic interactions play key roles in nucleic acids packing and/or recognition formations. These specific interactions might be identified by the similar mode and can be classified to several patterns. The geometrical parameters about hydrogen bond and base stacking would provide the tolerant aspect for multiple helical packing and recognition motifs. The NDB provides the structural information including torsional and helical parameters (4). It provides the Query search engine on categories in several fields. The mmcIF shows the search menu and allows a variety of kink parameter combination, is written using CIFlibx2 software package.

The tentative database including these geometrical parameters with the species and physical properties of the surrounding amino acids has been architected. The database, structural information of assemblage and distribution of bases, sugars and amino acids, can be a useful tool. There are many difficulties, such as classification of pairing and visualizations. The performance in our procedure is in progress.

CONCLUSION

The tentative database including kink geometrical parameters with the species and physical properties has been constructed. The model building program HLX6 for helices with the provided kink parameters was developed as a useful tool for understanding of the biological roles of these helices. In stead of torsion angle and helical parameter, kink parameter is elected as the input information to generate the starting global structure.

REFERENCES

2. http://ndbserver.rutgers.edu/NDB/
4. http://ndbserver.rutgers.edu/NDB/mmCIF/