Schematic Diagrams of Nucleic Acids and Protein–Nucleic Acid Complexes

BY VICTOR I. LESK
2 Field Way, Cambridge CB1 4RW, England

AND ARTHUR M. LESK
2 Field Way, Cambridge CB1 4RW, England, MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England, and European Molecular Biology Laboratory, Meyerhofstrasse 1, Postfach 1022.09, D-6900 Heidelberg, Federal Republic of Germany

(Received 25 May 1989; accepted 7 July 1989)

Abstract
Simplified representations of components of nucleic acids have been designed and implemented as programs integrated with other software that draws schematic diagrams of proteins. Examples illustrating the structures of oligonucleotides, tRNA and a protein–nucleic acid complex indicate the utility of these representations for making intelligible illustrations of complex structures containing nucleic acids.

Introduction
One role of computer graphics in molecular biology is to provide intelligible depictions of complicated structures. For proteins, schematic diagrams or ‘cartoons’ have been especially useful in showing the overall folds of proteins, or details of selected portions of a structure in the context of a simplified representation of the ‘Gestalt’. Such schematic diagrams, originally executed by hand, were first devised by Rossmann, Liljas (Holbrook, Liljas, Steindel & Rossmann, 1975) and Furugren (Brändén, Jörnvall, Eklund & Furugren, 1975), and further developed by Richardson (1981). Since then, computer programs have been coded to produce them. The main advantage of the programs is the ease of changing the orientation or viewpoint, and of producing stereo pairs of pictures.

In recent years, many nucleic-acid structures have been determined to high resolution by X-ray crystallography, separately or in combination with proteins. Some of these, such as transfer RNAs, are of size comparable to proteins, and present much the same difficulties of illustration. Nucleic acid–protein complexes are also complicated structures. We therefore felt that it would be useful to devise simplified representations of nucleic acids, and to integrate facilities for representing nucleic acids into programs that draw pictures of proteins.

Design of icons for components of nucleic acids
We simplify the backbone to either (a) line segments joining the phosphates, with small spheres at the phosphate positions, or (b) a curved ribbon in the mean plane of the backbone atoms, the width of which can be specified: such a representation has been useful for proteins as well (Lesk & Hardman, 1982, 1985; Feldman, Brooks & Lee, 1986; Carson & Bugg, 1986; Carson, 1987; Evans, 1987; Priestle, 1988).

The ring systems of purines and pyrimidines are shown as polygons. Optionally, the polygons may be thickened; the extraannular substituents may also be added (see Fig. 1).

We simplify sugar rings to pentagons or pentagonal prisms. There is no explicit representation of the sugar pucker, except for the overall thickness of the prism. In order to distinguish sugar pentagons from the pentagons representing the imidazole moieties of purine rings, we provide the option of rounding off the corners of the sugar pentagons. A parameter varying between 0 and 1 controls the shape: 0 = regular pentagon, 1 = circle, intermediate values produce line segments connected by circular arcs (see Fig. 1).

Examples
The size and complexity of the structure govern the degree of simplification required for intelligibility.

Fig. 2 shows an oligonucleotide, solved by Nelson, Finch, Luisi & Klug (1987). The inclusion of atom

Fig. 1. Representation of a single guanine residue of a nucleic acid. The purine ring is represented as the fused hexagonal and pentagonal prism, and the N2 and O6 substituents are shown in conventional ball-and-stick representations. The sugar ring is shown as a thicker pentagonal prism with rounded corners. Attached to the sugar are a shaded disc corresponding to the C5' and an open disc corresponding to the phosphorus atom.
C5' in the backbone eliminates potential ambiguity about the direction of the chain.

Fig. 3 shows a complex of an oligonucleotide with the drug distamycin (Coll, Frederick, Wang & Rich, 1987). The drug is shown in a space-filling representation. Note that it is useful to be able to show the oligonucleotide and the drug in contrasting representations.

Fig. 4 shows yeast tRNA$^{\text{Phe}}$ (Robertus, Ladner, Finch, Rhodes, Clark & Klug, 1974). In this case it is the overall folding of the chain that it is desired to represent.

Fig. 5 shows yeast tRNA$^{\text{Phe}}$ with a ribbon through the backbone and the bases, but not the sugars.

Figs. 6 and 7 show the model of the $\lambda$-cro-DNA complex (Anderson, Ohlendorf, Takeda & Matthews, 1981). In Fig. 6 note the distinguishable schematic representations of protein and nucleic acid.

In conclusion, several schematic representations of nucleic acids are useful adjuncts to the conventional schematic diagrams of proteins in studies of biological macromolecules for which detailed atomic structures are available.

We thank Drs B. W. Matthews and H. Nelson for coordinate sets.
References


