weights describing various combinations of residue orientations. Individual parameters can be varied in small fixed angular increments, generating smooth variations of structure that can be monitored graphically. The residues to be moved and the direction of angular change are chosen by random number techniques. The approach, unfortunately, is feasible only for small molecules in view of the time it takes to sample the range of low-energy states available to individual chain units. Other approaches that sample broader regions of torsion angle space must be employed in the study of long chains. Local conformations can be chosen, for example, on the basis of individual statistical weights with the crossing of high-energy barriers avoided. The smooth folding of the macromolecule is, unfortunately, lost in such an approach. The Monte Carlo sample is merely a collection of arbitrary, unrelated 3D structures. Sequential configurations of the Monte Carlo sample are random snapshots of overall chain movement, offering no clues to the transitional pathways that link them. Nevertheless, information concerning chain flexibility can be extracted from catalogs of conformational data accumulated during the sampling process. Structures can be organized on the basis of some criterion and distributions of the relevant parameters accumulated. The distributions can then be compared against ideal models or with those of related polymers.

One such probe of macromolecular flexibility is the spatial density distribution $\mathbf{W}_0(\mathbf{r})$. This quantity is a 3D function describing the probability of finding the terminus of a chain molecule at vectorial location $\mathbf{r}$ relative to a reference frame embedded at its origin (0). The location of the chain terminus in this manner is analogous to the use of distribution functions in describing the electron density of a molecular orbital. The characteristic shape of $\mathbf{W}_0(\mathbf{r})$ is tied to local chain properties, just as the shape of a molecular orbital is linked to the quantum states of its electrons. Because of the constraints of chemical bonding and the restrictions on local bending and twisting of chain residues, the distributions are skewed in short chains to shapes determined by the polymer architecture. The shapes can be correlated with observed measures of chain extension and flexibility, and the densities can be used to estimate the likelihood of polymer cyclization and looping as a function of chain length and sequence.

The conformation and properties of double-helical DNA are intimately tied to the linear sequence of its heterocyclic base side groups. A number of models have been offered to account for the subtle irregularities of local conformation in crystalline oligomers and the observed twisting and bending of the chain in solution. Computer programs have also been developed to translate the primary sequence of bases into 3D models on the basis of these rules. The flexible nature of the double helix is generally ignored in these representations with individual repeating residues of the chain described by fixed local geometries. Adjacent base pairs are found in experimental and theoretical studies, however, to adopt a broad range of accessible conformations rather than a single narrowly defined minimum energy state. The conformation of the DNA as a whole is more aptly described by a Monte Carlo computer monitored by the distribution of $\mathbf{W}_0(\mathbf{r})$.

We have taken advantage of color graphics techniques to study the conformation and mobility of selected DNA sequences. We have studied three short fragments of kinetoplast DNA from *Crithidia fasciculata* that exhibit dramatically different behavior on nondenaturing polyacrylamide gels. We find characteristic differences in the distributions of conformations between curved and rod-like sequences and in the overall flexibility of A·T and G·C rich regions. We employ a series of recent potential energy estimates of the local flexibility of adjacent nucleic acid base pairs to generate static representations and Monte Carlo samples of the double helix. We monitor the Monte Carlo chain flexibility with $\mathbf{W}_0(\mathbf{r})$, distinguishing regions of high and low probability density on the basis of color. These distributions incorporate a vast quantity of data that cannot be comprehended at the molecular level. We also color-code the DNA to examine effects of chain sequence on overall structure and flexibility. We additionally superimpose selected trajectories and various static representations of the double helix on the density distributions in an effort to understand the flexibility of the DNA as a whole.

Molecular Modeling of Two Regulatory Proteins, fix K and fn R, Homologous to CAP

D. Kahn, J.F. Gibrat and J. Garnier
Laboratory of Physical Biochemistry, INRA, University of Paris-Sud, 91405 Orsay, France

Two regulatory proteins — fix K of nitrogen fixation in the *Rhizobium* and Fn R of *E. coli* — are homologous to cAmp protein of *E. coli* or CAP. However, the homology is weak, about 25% of identical residues. In order to ascertain the alignment of the amino-acid sequences, secondary structure prediction was performed with the homolog program[1] with a specific weighting factor taking into account the homology. This procedure allows us to detect precisely the insertions or deletions to be made for further molecular modeling.

Subsequently, the conserved residues were assigned the corresponding coordinates of the CAP. These constitute the core of the two proteins, which have been energy minimized by quasi Newton or simplex methods. For the variable loops a specific method has been developed to avoid as much as possible the nearest minima. This is achieved essentially by simulated annealing, which takes care of the observed repartition of residue conformations in proteins of known structure. Tentative residues are proposed to interact with nucleotide bases of the putative recognition sequence.

1 Levin, J., Robson, B., and Garnier, J. *FEBS. Lets.* 1986, **205**, 303–308