Dynamics of Macromolecules and Nuclear Magnetic Relaxation: Application of Mode-coupling diffusion theory to DNA, proteins and their complexes.

Simone Fausti, Giovanni La Penna, Angelo Perico, Roberto Pratolongo

National Research Council, Institute of Physico-chemical Studies of Synthetic and Natural Macromolecules, via De Marini 6, 16149 Genova, Italy

SUMMARY: A mode-coupling approach has been used to solve the Smoluchowski diffusion equation describing the correlation functions relevant in nuclear magnetic relaxation experiments on biological molecules like protein and DNA. A good reproduction of relaxation patterns for molecules that are fluctuating about three-dimensional structures imposed by secondary motives has been found. Important differences can be explained with the poor statistics collected by numerical simulations that do not contain fluctuations contributing to rates in the ns time-scale, therefore comparable with rotational rates.

Introduction

Nuclear Magnetic Resonance (NMR) based techniques are sources of an increasing amount of information on macromolecular dynamics[1)]. Different experiments, combined with chemical synthesis of enriched species, can select magnetic relaxation pathways, therefore splitting such information between different frequencies and time-scales in the molecule. In order to translate all the experimental relaxation parameters in term of molecular motions (that are statistically represented by 2-rank autocorrelation functions of given vectors in the molecule) several approaches based on diffusive dynamics have been used: model functions have been derived and the related parameters have been optimized to fit experimental data[2)].

In this contribution, we show that once the statistics of the macromolecule has been described by some numerical approach (like Molecular Dynamics, MD, or Monte Carlo, MC, simulations) it is possible to solve the diffusion equation of selected configurational variables at the Smoluchowski level and compute the required 2-rank autocorrelation functions and the related spectral densities in the NMR frequency domain. We represent, by expanding the conditional probability in a suitable basis set, the diffusion equation in an eigenvector problem.

 CCC 1022-1360/99/$ 17.50+.50/0

The basis set is designed to give almost the exact result for rigid molecules, but it can be extended to describe the diffusive dynamics of fluctuating molecules, as the biological macromolecules can be modelled in a first approximation.

We are interested in the description of the dynamics involved in the stability of protein-DNA complexes, in particular the DNA-NK2 homeodomain complex involved in the first stage of DNA transcription, a crucial step for several genetic diseases. Therefore, we focus the application of diffusive equation to the NK2 77 residues homeodomain protein, studied by ^{15}N-NMR [3], and on model DNA oligomers ([d(TpCpGpCpG)$_2$], DNA5 hereafter, and [d(CpGpCpApApApTpTpTpGpCpG)$_2$], DNA12 hereafter) that have been extensively studied by ^{13}C-NMR [4-5].

Method

The statistics have been computed by multiple time-step MD simulations in the NVT (NK2 and DNA12) and NVE (DNA5) statistical ensembles using the ORAC program[6] and the AMBER 4.1 force-field[7] with the Smooth Particle Mesh Ewald algorithm for long-range electrostatic and with a 9 Å cut-off. The deterministic time-evolution contained in such 1-1.5 ns MD trajectories is never used and MD is required only to cope with correlated movements. Any MC technique suitable for macromolecules can be used.

To write the diffusive equation e representation of the macromolecule in terms of beads with given friction is required. It is possible to start from an all-atom representation for small molecules (DNA5), almost consistent with the atomistic force-field used in the MD simulation, continuing with coarse-grained representations of groups up to one bead for each amminoacid in proteins or nucleotide in DNA. According to observations in the area of rigid molecules[8], we found that 3-4 beads for amminoacid or nucleotide give a converged description of relaxation patterns. The build-up of group-beads can be done by representing the group-bead with the accessible surface area of the group and the equivalent Stokes' radius (Zimm rule). The choice of bead representation affects mainly the strength of hydrodynamic interaction (HI) that must be set by matching the low frequency NMR results in the rigid molecular portions.

The autocorrelation functions involved in NMR are:

$$P_2^{(f)}(t) = \left\langle P_2(\cos(\beta^{(f)}(t)) \right\rangle = \sum_{M=-2}^{2} \left\langle D_{M,0}^{(2)*}(\Omega^{(f)}(t)) D_{M,0}^{(2)}(\Omega^{(f)}(0)) \right\rangle \qquad (1)$$

where f is the bond (N-H or C-H) the fluctuation of which governs the relaxation parameter. Here we shall focus on the spin-lattice T_1 relaxation time. The conditional probability is the solution of the diffusion equation (see reference 9 for details): solving the diffusion equation using an expansion in term of 2-rank functions with 2nd and 4th powers of the first modes that one gets out with a bead-to-bond representation, one finally comes out with a simple multi-exponential representation of the autocorrelation function:

$$P_2^{(f)}(t) = \sum_i < f \psi_i >^2 \exp(-\lambda_i t) \qquad (2)$$

where ψ are the eigenstates represented in term of the 2-rank basis set and within brackets are the scalar contraction of such tensors. The number of rates λ in the conditional probability is equal to the number of elements in the basis set.

For rigid molecules, this approach, using the three modes obtained in terms of the virtual bonds to build-up 6 2-rank 2nd powers and 36 2-rank 4th powers (42 2-rank functions) gives the exact result for the above correlation function[9].

Results

In Fig. 1 we show the ^{13}C-NMR T_1 pattern computed with an all-atom bead model (except hydrogens that are incorporated in the heavy atoms) with Oseen HI tensor ($\alpha \approx 0.3$) and 5 modes used in the build-up of the 2-rank basis set, in the case DNA5.

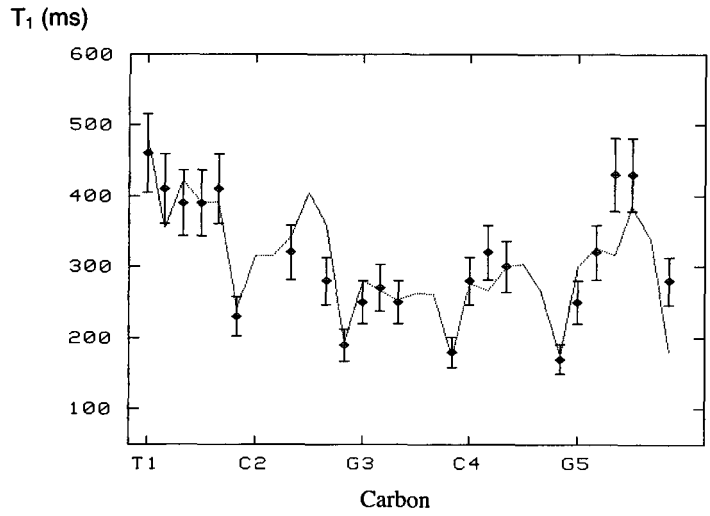

Figure 1: ^{13}C NMR T_1 125 MHz relaxation pattern for different nuclei in the DNA5 molecule. Nuclei are in the order C5', C4', C3', C2', C1' and C6 for C/T or C8 for G/A. Experiments (points with errorbars) and calculated values (dotted line).

Three frequencies are available in the experiments, but only the 125 MHz data are shown being the results similar at the other lower frequencies. It is well captured most of the mobility pattern, the difference between head and tail (the head T is not involved in a base-pair therefore it is much more disordered) and the difference of the aromatic core with respect to the sugar and the backbone, due to the rigidity related to the base-pair interaction.

In Table I are the results for the DNA12 molecule with 3-4 beads per nucleotide and Rotne-Prager HI ($\alpha \approx 0.7$). Here the experiments are less rich in information, but again the different relaxation behaviour of backbone and sugar nuclei is captured and translated in terms of mobility (correlation times).

Table 1: Correlation times (τ) in ns and ^{13}C-NMR T_1^{-1} in s for the DNA12 molecule. Experimental (Exp.) and calculated (DD) results.

Carbon	Exp. T_1^{-1} (s)	DD T_1^{-1} (s)	DD τ (ns)
C1'	2.42	2.38	2.80
C5'	2.19	2.19	3.41
C6-C8		3.91	3.22

The application of a similar approach to the NK2 homeodomain protein give good results in the low frequency case (34.5 MHz), where the pattern is again dominated by the molecular wobbling around the 3-helix bundle structure. On the other hand, the high frequency experimental data (60.8 MHz) contain information about internal modes with rates of about 1 ns. The evidence of this comes out by an *ad hoc* matching of the computed figures onto the experiments, by scaling the internal rates for a factor 1/100, a reduction that brings such rates into a range 1/10 of the rotational rates (10 ns). To fill the large gap between internal and rotational rates it is necessary to sample, by MD or whatever technique, the many wells in the energy landscape, since this kind of diffusion through many low-energy barriers is likely to contribute to rates that are 100 times smaller than the wobbling rates usually sampled by MD of such systems. Alternatively a master equation kinetic approach can be used when such barriers become higher.

References

1. Abragam A., *Principles of Nuclear Magnetism*, Clarendon Press, Oxford, U.K., 1986.
2. Lipari G., Szabo A.; *J. Am. Chem. Soc.* **104**, 4546-4559 (1984).
3. La Penna G., Mormino M., Pioli F., Perico A., Fioravanti R., Gruschus J.M., Ferretti J.A.; *Biopolymers* **49**, 235-254 (1999).
4. Borer P.N., LaPlante S.R., Kumar A., Zanatta N., Martin A., Hakkinen A., Levy G.C.; *Biochemistry* **33**, 2441-2450 (1994).
5. Gaudin F., Chanteloup L., Thuong N.T., Lancelot G.; *Magn. Res. Chem.* **35**, 561-565 (1997).
6. Procacci P., Darden T., Marchi, M.; *J. Phys. Chem.* **100**, 10464-10468 (1996); Procacci P., Paci E., Darden T., Marchi M.; *J. Comp. Chemistry* **18**, 1848-1862 (1996).
7. Cornell W.D., Cieplak P., Bayly C.I., Gould I.R., Merz K.M. jr., Ferguson D.M., Spellmeyer D.C., Fox T., Caldwell J. W., Kollman P.A.; *J. Am. Chem. Soc.* **117**, 5179-5197 (1995).
8. Venable R.M., Pastor R.W.; *Biopolymers* **27**, 1001-1014 (1988)
9. La Penna G., Pratolongo R., Perico A.; *Macromolecules* **32**, 506-513 (1999).