DNA structures from phosphate chemical shifts

Joséphine Abi-Ghanem¹, Brahim Heddi², Nicolas Foloppe³,* and Brigitte Hartmann¹,²,*

¹INTS, INSERM S-665, 6 rue Alexandre Cabanel, Paris 75015, ²IBPC, CNRS UPR 9080, 13 rue Pierre et Marie Curie, Paris 75005, France and ³51 Natal Road, Cambridge CB1 3NY, UK

Received June 30, 2009; Revised October 14, 2009; Accepted November 1, 2009

ABSTRACT

For B-DNA, the strong linear correlation observed by nuclear magnetic resonance (NMR) between the ³¹P chemical shifts (δP) and three recurrent internucleotide distances demonstrates the tight coupling between phosphate motions and helicoidal parameters. It allows to translate δP into distance restraints directly exploitable in structural refinement. It even provides a new method for refining DNA oligomers with restraints exclusively inferred from δP. Combined with molecular dynamics in explicit solvent, these restraints lead to a structural and dynamical view of the DNA as detailed as that obtained with conventional and more extensive restraints. Tests with the Jun-Fos oligomer show that this δP-based strategy can provide a simple and straightforward method to capture DNA properties in solution, from routine NMR experiments on unlabeled samples.

INTRODUCTION

Nuclear magnetic resonance (NMR) is potentially the most powerful experimental method to investigate the structure and dynamics of macromolecules in solution. Traditionally, interproton distances extracted from nuclear Overhauser effect (NOE) measurements and scalar couplings are used as distance and torsion angle restraints in refinement. These conventional short-range restraints are now often supplemented by long-range information inferred from residual dipolar coupling (RDC) measurements that require oriented labeled molecules (1–5).

New methods using directly the chemical shifts have recently emerged, in particular, in the protein refinement context. These approaches were undertaken to model proteins that cannot be studied by conventional NMR but also to circumvent the long and tedious task to collect and analyze the NOE data. Indeed, the chemical shifts are particularly sensitive to the electronic environment and very accurately measured. However the difficulty is to interpret and translate these experimental observables into properties that can guide structural refinement protocols.

Thus, ¹³C₂ chemical shifts, sensitive to protein secondary structure, have been proposed to improve protein models (6,7). Chemical shifts of Cα, Cβ, Hα, Hβ, HN and Cγ are collectively used to define torsion angles (8,9). Reliable three-dimensional structures were obtained with NOE-derived distances supplemented by using ¹³C₂ chemical shifts, measured and computed with a density functional approach, as a function of all backbone and side-chain torsional angles (10–13). However, the quantum calculations can be quite time consuming when applied to large molecules and simplifying the level of theory could degrade the accuracy of the resulting structures. Another approach, free of NOE restraints and minimizing the dependence on ab initio calculations, was recently proposed by Vendruscolo and co-authors (14,15). In their CHESHIRE protocol, the experimental chemical shifts are compared to an extensive structural database for determining analogs of short protein fragments, and predicting the most compatible dihedral restraints; the assembly of selected fragments from the database results in models then refined by molecular mechanics in conjunction with back-calculations of chemical shifts. Two approaches emanate from this strategy, either employing a Monte–Carlo algorithm for the reconstitution of the entire protein (9,16,17) or, for the CS-ROSETTA method, an additional Monte–Carlo sampling coupled with an all-atom force field for the refinement (9,17). Chemical-shift-generated structures can be now generated via a web server (18). These strategies applied to high-resolution refinements are still limited to a few cases of relatively small proteins. Nevertheless, they appear very promising and open to further development.

With nucleic acids, the ¹³C chemical shifts, earlier used to assess ϵ angles in the deoxyribonucleotide d(TpA) (19), enabled to determine sugar puckers and α/γ exocyclic
angles in ribo and deoxy-nucleotide units (20) and in RNA (21). In addition, they allowed to qualitatively detect perturbations in base-pairing and base-stacking (22) in RNA. $^1$H chemical shifts associated to ab initio calculations were proposed for helping in the sequential assignment of resonances (23). $^3$P chemical shifts (δP) and $^3$J scalar dipolar couplings have been shown to improve the precision in refinement of large RNA molecules, by constraining dihedral angles (24). The δP anisotropy collected on labeled nucleic acids oriented in liquid crystals were also used to constrain the orientation of the phosphodiester groups relative to the molecular alignment tensor (25,26). However, it is not sure that constraining the B-DNA phosphodiester linkages in one orientation is pertinent, given that they are submitted to conformational transitions, as explained below. To our knowledge, no exploitation of chemical shifts for overall structure determination was reported for nucleic acids.

In this context, we propose a new and convenient strategy to refine nonlabeled B-DNA oligomers exclusively on the basis of their δP. In B-DNA in solution, some phosphate groups oscillate between two conformations, BI and BII (27–32). The BI↔BII transition corresponds to a crankshaft motion of the strongly correlated torsions ε and ζ, used to describe the two conformations, i.e. BI (ε = trans, ζ = g–, with $\varepsilon - \zeta = -90^\circ$) and BII (ε = g–, ζ = trans, with $\varepsilon - \zeta = +90^\circ$) (Figure 1). These two states, initially identified from crystallographic studies (33), were then detected in NMR by measurements of $^3$J$_{H2O,P}$ spin–spin coupling constants and/or δP (28,34–36). Because the ε,ζ crankshaft motions in solution are fast (35), the δP is a continuous function of the fraction of the BI and BII states. Theoretical studies (37,38), statistical analysis of X-ray structures (39,40) and recent NMR studies (41) showed separated energy minima for BI and BII, with transition barriers depending on the local dinucleotide sequence.

The BI/BII ratio, inferred from δP, is primarily controlled at the dinucleotide level (29), confirming that the propensity to undergo the BI/BII transition is sequence dependent. Crucially, these phosphate states, the twist, the roll and the base displacement are intimately coupled (39,42–46). These relationships reflect the B-DNA intrinsic mechanics and can be captured in NMR by a marked linear correlation between δP and three internucleotide distances, $H2^\prime_i-H6^1_i+1$ [ds(H2$^\prime$)], $H2^\prime_i-H6^1_i+1$ [ds(H2$^\prime$)] and $H6^1_i-H6^1_i+1$ [ds(H6/8)] (29,47) (Figure 1). These correlations were used to translate the δP in terms of BI/BII ratios (29). Most importantly, they also allow to interpret δP in terms easily integrated in a structural refinement.

This approach was incipient in our recent refinement of the Jun-Fos oligomer double-stranded DNA has the 14-bp sequence $5^\prime$-d(G1 C2 A3 T4 T5 C6 T7 G8 A9 G10 T11 C12 A13 G14)A4–3$\cdot$S$^\prime$-d(C15 T16 G17 A18 C19 T20 C21 A22 G23 A24 A25 T26 G27 C28)-3$. This system was previously characterized experimentally and structurally in details in solution (29,47,48). Thus, it was selected to test a new method for the structural refinement of DNA structures, which exploits NMR phosphate chemical shifts combined with MD simulations.

**Phosphate chemical shifts interpreted in terms of distance restraints**

In previous studies (29,47), the δP (Supplementary Table S1) were found linearly correlated to the three internucleotide distances $H2^\prime_i-H6^1_i+1$ [ds(H2$^\prime$)], $H2^\prime_i-H6^1_i+1$ [ds(H2$^\prime$)] and $H6^1_i-H6^1_i+1$ [ds(H6/8)], extracted from NOESY cross-peaks with particular care regarding to spin diffusion. This relation was also retrieved on X-ray structures (29), where the three internucleotide distances [ds(H2$^\prime$)], [ds(H2$^\prime$)] and [ds(H6/8)] correlated with (ε – ζ) which represents the BI and BII states. This ensures that no systematic NMR biases occur in our calibration of the relation between δP and the internucleotide distances of interest. These distances were extrapolated from δP (referenced to phosphoric acid) measured for the 22 phosphate groups of the

**MATERIALS AND METHODS**

**DNA sequence**

The Jun-Fos oligomer double-stranded DNA has the 14-bp sequence $5^\prime$-d(G1 C2 A3 T4 T5 C6 T7 G8 A9 G10 T11 C12 A13 G14)A4–3$\cdot$S$^\prime$-d(C15 T16 G17 A18 C19 T20 C21 A22 G23 A24 A25 T26 G27 C28)-3$$. This system was previously characterized experimentally and structurally in details in solution (29,47,48). Thus, it was selected to test a new method for the structural refinement of DNA structures, which exploits NMR phosphate chemical shifts combined with MD simulations.
Jun-Fos oligomer (29), following the equations established from the correlations:

\[
\begin{align*}
\text{ds(H}^2\text{)} &= \delta P / 0.34 + 4.70 \\
\text{ds(H}^4\text{)} &= \delta P / 0.47 + 3.94 \\
\text{ds(H}^6/8\text{)} &= \delta P / 0.29 + 5.86
\end{align*}
\]

**NMR distance restraints**

MD simulations were carried out under two sets of NMR internucleotide distance restraints termed Res\_total and Res\_δP. The 5′- and 3′-terminal dinucleotides were unrestrained. The 106 internucleotide distances collected from NOESY cross-peaks were not restrained since they were spontaneously respected in all the MDs.

The Res\_total set contained 100 internucleotide distances (Supplementary Table S2). These distances were used as restraints in a previous structural refinement (29): 48 H2−H6/8i−1, H2′−H6/8i−1 and H6/8i−1-H6/8i+1, and 34 various distances (H1′-H6/8i−1, H1'-H4'i+1, H2-H1'i+1 and H2'/2H5-H5'1). This pool was supplemented by 18 H2′−H6/8i−1, H2″−H6/8i−1 and H6/8i−1-H6/8i+1 distances, not directly measured but inferred from measured distances or from the corresponding δP. The experimental error was estimated to ±10% of the considered distance.

The second, new, set of restraints Res\_δP consisted of 66 ds(H2′), ds(H2′0) and ds(H6/8) internucleotide distances (three distances per dinucleotide step, defined earlier), all extrapolated from the corresponding δP.

To estimate the agreement between experimental NMR (d\_exp) and theoretical MD (d\_theor) distances, we calculated two descriptors of the fit between d\_exp and d\_theor. First, the overall similarity, comparing the profiles of d\_exp and d\_theor along the sequence, was assessed with the correlation coefficient R (calculated by linear regression) between d\_exp and the averaged d\_theor values, for different sets of relevant distances. Second, the average difference between d\_exp and d\_theor (△d) was calculated. In addition, we consider that an individual distance violation occurs when the intervals d\_exp ± 10% (± experimental error) and d\_theor ± standard deviation (± SD) do not overlap.

**MD simulations**

Simulations were performed using the AMBER 8 program (49), with the Parm98 (50) force field. We previously found (48) that restrained simulations carried out with Parm98 led to a more realistic representation of the Jun-Fos oligomer than Parmbsc0 (51). The Jun-Fos oligomer in an initial AMBER standard B-DNA conformation was neutralized with 26 Na⁺ ions and hydrated with 6770 TIP3P water molecules (52,53) in a truncated octahedron. Simulations were performed with periodic boundary conditions at constant temperature (300 K) and pressure (1 bar) using the Berendsen algorithm (54). The integration time-step was 2 fs and covalent bonds involving hydrogens were restrained using SHAKE (55). Long-range electrostatic interactions were treated using the particle mesh Ewald (PME) approach (56) with a 9-Å direct space cut-off. The Lennard-Jones interactions were cut off at a distance of 9 Å. The non-bonded pair-list was updated heuristically and the center-of-mass motion removed every 10 ps.

The water molecules and counterions were energy-minimized and equilibrated at 100 K around the fixed DNA for 100 ps in the NVT (at constant volume and temperature) ensemble; the entire system was then heated from 100 to 300 K in 10 ps by 5-K increments with harmonic positional restraints on the solute atoms (force-constant of 5.0 kcal/mol/Å²). The simulation was continued in NPT (at constant pressure and temperature). The positional restraints were gradually removed over 250 ps and followed by 1 ns of unrestrained simulations for further equilibration.

The free MD of 1 ns yielded the starting point for the other, restrained, MD protocols. These restrained MDs were run in presence of NMR distance restraints described above, i.e. either Res\_total (MD\_ref) or Res\_δP (MD\_δP), for 15 ns (MD\_ref) or 35 ns (MD\_δP). These restraints were applied either instantaneously (MD\_δP) or in time-averaged manner (MD\_ref), via a mixed parabolic (for d\_exp - 10%) and hyperbolic (for d\_exp + 10%) potential around a central flat-bottomed well covering the experimental range of the distances, including experimental errors (d\_exp ± 10%). The time-averaged restraints on property R were applied with the following equation:

\[
R = (1/C) \int \exp[(t - t)/τ] r(t')^{-1} dt'^{-1/5}
\]

where i is the current time, r(t') the internal coordinate at time t' and C a normalization factor. The damping constant τ was set to 10 ps and, following previous tests (48), the best results were for i = 1 used here. For the flat-bottomed harmonic potential, force constants of 5 (MD\_ref and MD\_δP) and 10 kcal/mol/Å² (MD\_δP), were tested.

Convergence of the MDs with respect to the DNA structure was achieved. The root mean square deviation (RMSD) between snapshots and either the starting or the average structures were very stable after 1 ns of the restrained simulations. Statistics for the DNA descriptors (sugar and backbone conformations and inter-base pair parameters) give virtually identical results when they are extracted from different 10-ns blocks of the trajectories. Such analysis on MD\_δP trajectory extended to 35 ns confirms that 15 ns are sufficient to reach convergence.

**Crystallographic data**

The crystal structures used in section ‘Test 3: coupling between backbone states and helical parameters’ and Figure 3 were comprised of 19 B-DNA oligomers with a resolution <2Å (PDB codes: 431D, 436D, 460D, 461D, 463D, 476D, 1DPN, 1EN3, 1EN8, 1EN9, 1ENE, 1ENN, 1EI4, 1FQ2, 5DNB, 1D23, 1D49 and 355D).
Structural descriptors
DNA structures and helicoidal parameters were analyzed with the Curves 6.1 algorithm (57,58), following the Cambridge convention (59). To avoid end effects, only the 12 central base-pairs (11 dinucleotides) were analyzed. The first nanosecond of each MD with NMR restraints was discarded.

RESULTS
Overview of the restraint sets and simulations
To present the different restraint sets pertaining to the present work, we summarize the experimental data obtained from our previous NMR study of the Jun-Fos oligomer (29). For the 11 central dinucleotides, the NMR spectra enabled to extract 106 intranucleotide distances, 82 internucleotide distances from NOE cross-peaks, 97 torsion angles from 3JH1-P, H2-P, H5-P, H8-P, and 3JH4-P scalar dipolar couplings, and all the BI/BII ratio from δP. 18 additional internucleotide distances were extrapolated from δP. The internucleotide distances are sensitive to the torsions and to the sugar conformations, and the 3JH1-P, H2-P, H5-P, and 3JH4-P scalar dipolar couplings, reflect the sugar conformations and the backbone angles β and γ. These intranucleotide distances do not vary significantly within a B-DNA devoid of disrupting features such as mismatches. Indeed, all of these observables were spontaneously respected in all the MDs carried out on the Jun-Fos oligomer (29,47,48), comprising unrestrained MDs (48). In contrast, neither the experimental internucleotide distances [mainly sensitive to the twist-and-roll parameters (60)], nor the experimental BI/BII ratios could be reproduced in unrestrained MD (48). Thus, they must be restrained during any refinement. 3JH1-P values could be expressed in terms of δ angle restraints; however, it would not allow for the dynamics nature of the BI->BII equilibrium. Still, due to the tight relationship between backbone motions and inter-base-pair parameters, the internucleotide distance restraints enable to indirectly control the BI/BII ratios, twist, roll and base displacement (48). In other words, the restraints crucial for B-oligomer refinement in general are the internucleotide distances.

Here, two set of restraints were used, Res_total and Res_δP. Res_total included a total of 100 internucleotide distance restraints (Supplementary Table S2), mainly of a conventional type, i.e. directly derived from NOESY cross-peaks (‘NOE distances’ hereafter). These experimental data correspond to 4.6 distances per dinucleotide step on average and represent ~67% of all theoretically possible internucleotide distances that could be routinely observed in the Jun-Fos oligomer. Restraints set Res_δP contained three types of internucleotide distances (Figure 1), ds(H2'), ds(H2") and ds(H6/8), all inferred (‘Materials and Methods’ section) from the 22 δP previously measured (29); so the set Res_δP contains 66 (3 × 22) restraints. In Res_total, 48 of these distances were obtained from NOEs, and were compared to their counterpart in Res_δP inferred from δP. The distances derived from NOEs and δP differ by only 0.3 Å on average, present similar profiles along the sequence (Supplementary Figure S1), and are closely correlated (correlation coefficient of 0.92, slope of regression line of 0.9). Thus, Res_δP is congruent with Res_total.

The reference refinement (MD_ref) relied on a protocol that has already been shown to yield a detailed dynamical structure of the Jun-Fos oligomer in solution (48). MD_ref used Res_total time-averaged restraints because such restraints were slightly more realistic than when applied instantaneously (48). The δP-based protocol using Res_δP was tested with four variant simulations, presented in Supplementary Table S3. The best fit between experimental and simulated data was obtained with Res_δP distances applied instantaneously with a force constant of 10 kcal/mol/Å² for the parabolic potential (Supplementary Tables S4 and S5). This refinement, called MD_δP, is thus compared to MD_ref, the main objective being to assess if the new δP-based method (MD_δP) performs as well as the ‘enhanced conventional’ approach (MD_ref).

In each MD, the average RMSD between the snapshots and canonical B and A-DNA were <2.9 Å and >5.3 Å, respectively. These RMSD values show that the overall
simulated structures were stable in the B form, consistent with the NMR data.

Test 1: backbone conformation and directly coupled distances

This section compares the distances modeled in simulations MD₆P to their NOE counterparts, used in MD_ref. MD_ref is thus also used for comparison to MD₆P. It is then natural to extend this comparison to the related BI/BII populations, between MD₆P, MD_ref and their experimental references.

The 48 NOE distances ds(H₂₀), ds(H₂₀₀) and ds(H₆/₈) are well reproduced in MD₆P (Table 1). This satisfying correspondence is illustrated in Figure 2 that compares the experimental and simulated ds(H₂₀), ds(H₂₀₀) and ds(H₆/₈) for MD_ref and MD₆P. Introducing the distances ds(H₂₀), ds(H₂₀₀) and ds(H₆/₈) as restraints acts on the backbone behavior, i.e. the BI→BII equilibrium (Figure 1) (48). The δP, collected for all the Jun-Fos oligomer phosphate groups, reflect the diversity of local flexibilities along this B-DNA sequence, from 0 (δP of −0.70) to 85% (δP of 0.00) of BII conformers (Supplementary Table S1 and Figure S2). These NMR

<table>
<thead>
<tr>
<th></th>
<th>MD_ref</th>
<th>MD₆P</th>
</tr>
</thead>
<tbody>
<tr>
<td>ds(H₂₀)</td>
<td>0.90</td>
<td>0.93</td>
</tr>
<tr>
<td>ds(H₂₀₀)</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>ds(H₆/₈)</td>
<td>0.35</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Experimental NMR observables for the Jun-Fos oligomer include internucleotide distances measured from NOE spectra and the percentage of BI conformation for every phosphate. These observables are compared to their simulated counterpart, whether obtained with a conventional refinement protocol (MD_ref) or with the δP-based method (MD₆P). The NOE distances are divided into a group of distances restrained in the two MDs [ds(H₂₀), ds(H₂₀₀) and ds(H₆/₈)] and the distances (‘Other’ distances) restrained in MD_ref but not in MD₆P. The correlation coefficients (R) and the average differences (δ in Å) between NMR and simulated distances are given for each distance category, across MDs. The RMSD between the experimental and simulated BII percentages was calculated as: RMSD = [Σ(%BII_theo − %BII_sim)² / N]½, with N = 22, number of phosphate. The best and worst values of the RMSD would be 0 and 100, respectively.

![Figure 2](https://example.com/figure2.png)
data show that the steps conductive to BII are TpG (26–52% of BII), their complementary CpA (69–85% BII) and GpA (32–52% BII), all located in the TGA stretches of the Jun-Fos oligomer (29,47). The other phosphate groups of the oligomer remain essentially trapped in BI. Besides, unusual \( \alpha/\gamma \) conformations are observed neither in NMR (29) nor in the present simulations, suggesting that the corresponding force-field artifact (29,51,61,62) may be circumvented under restraints. The experimental BI/BII percentages are especially well retrieved in MD\(_P\) (Table 1, Figure 2 and Supplementary Figure S2). This proper representation of backbone motions is clearly the result of the well-respected restraints \( Res_\delta P \) having a desirable impact, as unrestrained MDs performed poorly in the same comparison with experiment (48).

### Test 2: distances not explicitly restrained in MD\(_P\)

We now compare to experiment (when available) the distances unrestrained in the MD\(_P\) refinement, which is crucial to validate the \( \delta P \)-based approach. Indeed, it is of great interest to see if \( Res_\delta P \) is enough to maintain all aspects of the DNA structure within the experimental regime. The global characteristics of the B-DNA form (\( \chi \) angle and predominance of south sugar conformations) are well respected in unrestrained MD (48). Moreover, the simulated internucleotide distances, sensitive to these parameters, are highly correlated to the corresponding 106 distances extracted from NOESY cross-peaks (correlation coefficients \( \geq 0.9 \)) in MD\(_P\) and MD\(_\delta P\). Therefore, the internucleotide distances are properly treated by the force field alone in absence of restraints.

We now turn to the 34 internucleotide NOE distances restrained in MD\(_P\) but not in MD\(_\delta P\). Most of them were not spontaneously respected in the unrestrained MD (48). In Table 1 and Figure 2 these distances are in the category ‘Other’ distances.

In MD\(_\delta P\), these unrestrained distances are reasonably close to experiment (Table 2, Figure 2). Some of these distances are characteristic of adenines (H\(_2\)-H\(_{1+1}\)) and cytosines (H\(_2^{2+1}\)-H\(_{5+1}\)). The analysis of very-high-resolution X-ray DNA structures reveals that these distances are in fact coupled with ds\(_2\), ds\(_{2+1}\) and ds\(_{H/6}\) (correlation coefficients of 0.8–0.9) and thus respond indirectly to the restraints in MD\(_P\). The distances H\(_{1+1}\)-H\(_{6/8}\) and H\(_{1+1}\)-H\(_{4+1}\) are sensitive to the conformations of two successive sugars (60). In all the simulations the unrestrained sugars were mainly in south, with south percentages higher for purines (>90%) than pyrimidines (>70%), in agreement with NMR (25,29,63,64). This concordance is sufficient to avoid severe violations on H\(_{1+1}\)-H\(_{6/8}\) and H\(_{1+1}\)-H\(_{4+1}\) in MD\(_\delta P\). However, the correlation between experimental and simulated distances is better in MD\(_P\) than in MD\(_\delta P\) (Table 1). Yet, this does not degrade the representation of the helicoidal parameters in MD\(_\delta P\), as examined in the next section.

### Table 2. Conformational families for the Jun-Fos oligomer

<table>
<thead>
<tr>
<th>Conformational family</th>
<th>BI/BII configurations</th>
<th>NMR</th>
<th>MD(_P)</th>
<th>MD(_\delta P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TpGpA ApCpT</td>
<td>( \geq 0.6 ) 0.6</td>
<td>28</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>TpGpA ApCpT</td>
<td>( \geq 0.6 ) 0.6</td>
<td>28</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>TpGpA ApCpT</td>
<td>( \geq 0.6 ) 0.6</td>
<td>28</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>TpGpA ApCpT</td>
<td>( \geq 0.6 ) 0.6</td>
<td>28</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>TpGpA ApCpT</td>
<td>( \geq 0.6 ) 0.6</td>
<td>28</td>
<td>14</td>
</tr>
</tbody>
</table>

The five conformational families of the Jun-Fos oligomer and their populations (percentage) were previously deduced from NMR (48). The phosphate configurations (BII in bold italic) observed on the TpGpA regions are a distinctive feature of each family (48). For MD\(_P\) and MD\(_\delta P\), the population of each family is compared with the experimental populations.

### Test 3: coupling between backbone states and helical parameters

A common method to check the reliability of simulated DNA structures is to compare their average inter-base helical parameters values with those extracted from X-ray structures. A more detailed approach is to analyze whether the simulations satisfactorily reproduced the well-documented couplings between the helical parameters, twist, roll and base displacement (X-disp), and the backbone BI/BII states (29,39,42,45,46,48,65).

The average values of the roll and twist were extracted for the three possible combinations BI•BI, BI•BII and BI•BII encountered on complementary dinucleotides, from the MD snapshots and from a set of high-resolution X-ray structures of free B-DNA. We focus on CpA•TpG steps because they populate the three different facing phosphate combinations. Overall, the MD models reflect the general trend observed in X-ray structures: in both MD\(_P\) and MD\(_\delta P\), the greater the BII character of facing phosphates in a complementary dinucleotide, the higher the twist and the more negative the roll (Figure 3).

In addition, several proximal phosphates in BII are typically accompanied by a displacement of bases towards the major groove (more positive X-disp) that propagates to neighboring bases to maintain sufficient stacking (37,46). Therefore, the global X-disp values are sensitive to the density of BII steps, i.e. the fraction of phosphates observed simultaneously in BII. In high-resolution X-ray structures the global X-disp of a purely BII oligomer is \(-1.4 \) Å, but it is null or positive with more than 25% of BII steps. These trends are equally well represented in MD\(_P\) and MD\(_\delta P\), with X-disp of \(-1.7 \) Å for a pure BII configuration and X-disp of \(-0.6 \) with 32% of BII phosphates (7 BII and 15 BI), for both MD\(_P\) and MD\(_\delta P\).

In sum, the MD\(_\delta P\) protocol represents correctly the intrinsic mechanical couplings of B-DNA, providing a sound basis to address the conformational dynamics of the Jun-Fos oligomer.

### Test 4: the dynamical structure of the Jun-Fos oligomer

Arguably, the most stringent test of the \( \delta P \)-based method is whether it yields the overall correct structural dynamics of DNA in solution. Having a precise representation of
The main objective of this work is to propose a new method for refining DNA structure in solution with NMR δP as the sole experimental input, combined with MD simulation techniques. The Jun-Fos oligomer is an appropriate system to test the new method, as much information has accumulated about its structure and dynamics in solution (29,47,48). The Jun-Fos structure in solution cannot be obtained directly from unrestrained MD simulations with current force-field limitations (48), and appropriate experimental input is required.

At its core, this new δP-based method simply relies on the strong linear correlations in B-DNA between δP and the internucleotide distances ds(H2'/2), ds(H2'/2') and ds(H6'/8) (29,47). These correlations allow to translate any δP in terms of three distance restraints, following simple equations previously established (29) (reiterated in ‘Materials and Methods’ section).

Having constituted a set of restraints exclusively inferred from δP, we showed that all the distances extracted from the corresponding MD_δP refinement reproduced well the experimental data. Cross-correlations between restrained and unrestrained distances, together with the adequate treatment of the generic B-DNA features by the force field, largely explain why the δP-based method yields a good agreement between measured and simulated DNA distances. In addition, MD_δP correctly represents the BI/BII backbone states, accounts for the DNA intrinsic mechanics, i.e. the relationships between BI/BII ratio and the helicoidal parameters (twist, roll, base displacement) and allows the characterization of the Jun-Fos conformational families in solution. In sum, compared to the reference conventional refinement, the δP-based refinement can generate credible representations of DNA structure and dynamics in solution.

The δP-based method does not require labeled DNA or NOE measurements. Furthermore, it circumvents imprecisions in distances restraints inferred from NOEs, which can be introduced via spin diffusion and require very careful treatment of data. One limitation concerns the problem of repeated or large (>20 bp) sequences generating considerable overlaps and/or anisotropic overall motions that prevent the assignment of all the 31P and 1H resonances. This restriction is not particular

Figure 3. Impact of the conformation of facing phosphate groups on key helical descriptors. The average values of twist (°) and roll (°) of the Jun-Fos oligomer TpG•CpA steps are plotted as a function of the three possible conformational combinations of facing phosphates, BI•BI, BI•BII and BII•BII. The data were extracted from a set of high-resolution X-ray structures (green), from MD_ref (blue) and MD_δP (black).
Figure 4. Structure and dynamics of the Jun-Fos oligomer in terms of conformational families. Conformational families 1–5 (from top to bottom, defined in Table 2) are characterized in terms of twist and roll profiles along the oligomer sequence. Twist (−) and roll (−) average values were extracted from MD_ref (blue) and MD_dP (black). The standard deviations of twists and rolls are 5°. The DNA sequence is represented by its first strand in the X-axis, but both strands were included in the calculations.
to the δP-based method, and these DNAs require specific labeling and RDC measurements in oriented medium to be reliably refined (3).

To our knowledge, the δP-based method demonstrates for the first time how to exploit chemical shifts as the main experimental basis to refine B-DNA structures. This is an alternative to the classical method based on numerous NOE measurements and data treatments. This could significantly increase the throughput of the structural characterization of DNA sequences for systematic structural biology analyses, for instance the study of protein target sites and their mutants that are typically B-DNA of 10–15 bp. This is indeed a pressing need, considering the biology analyses, for instance the study of protein target sites and their mutants that are typically B-DNA of 10–15 bp. This is indeed a pressing need, considering the biological significance of the DNA sequence and its interaction with proteins.

The DNA curvature, very significant increase the throughput of the structural characterization of DNA sequences for systematic structural biology analyses, for instance the study of protein target sites and their mutants that are typically B-DNA of 10–15 bp. This is indeed a pressing need, considering the biology analyses, for instance the study of protein target sites and their mutants that are typically B-DNA of 10–15 bp. This is indeed a pressing need, considering the biological significance of the DNA sequence and its interaction with proteins.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**FUNDING**

Funding for open access charge: Institut National de la Transfusion Sanguine (INTS), France.

Conflict of interest statement. None declared.

**REFERENCES**


