Journal of Biomolecular NMR, 4 (1994) 279–299
ESCOM

J-Bio NMR 165

How well do time-averaged J-coupling restraints work?

David A. Pearlman

Vertex Pharmaceuticals Incorporated, 40 Allston Street, Cambridge, MA 02139-4211, U.S.A.

Received 29 July 1993
Accepted 13 September 1993

Keywords: Molecular dynamics; NMR refinement; Nucleoside; Adenosine; J-coupling

SUMMARY

A comparison is made of the consequences of using time-averaged and conventional vicinal ^{3}J-coupling restraints in molecular dynamics refinement of an adenosine nucleoside model system. The target values for the restraints are derived from a 3-ns unrestrained molecular dynamics simulation. A comparison of the results from the restrained refinements and the unrestrained trajectory reveals that while both restraint types (time-averaged and conventional) are capable of acceptably reproducing the averaged values of the restrained parameters, time-averaged J-coupling restraints allow a more realistic and thorough description of conformational fluctuations. The full description of conformational behavior for the sugar ring using time-averaged J-coupling restraints is in excellent agreement with the unrestrained results. J-coupling restraints can result in a localized 'heating effect' about the underlying torsion. This allows a restrained torsion to sample all low-energy rotomers separated by modest barriers in an appropriately weighted mixture that reproduces the J-restraint target value. This will generally be advantageous for experimentally derived data, though it can be misleading if all these low-energy rotomers did not contribute to the ensemble that yields the measured J-value. An analysis of how the force constant used in the restraint terms affects the refinement indicates that smaller force constants are to be preferred, and that constants in the range of $K_j \geq 0.4$ kcal s^2/mol are acceptably large to overcome the intrinsic preferences of the force field.

INTRODUCTION

Traditionally, 2D NMR structural refinement has been carried out by deriving a set of target distances and torsion angles from the collected data, and then imposing these on a model structure with a set of instantaneous restraints. For example, in conventional Molecular Dynamics (MD) refinement the restraints take a form such as (Clore et al., 1986; De Vlieg et al., 1986; Nilges et al., 1988):

$$E_I = K_I(I(t) - I_0)^2 \tag{1}$$

where E_I is the restraint penalty energy, K_I is a force constant, $I(t)$ is the value of the restrained internal (distance or torsion) in a particular MD timestep, and I_0 is the experimentally derived

0925-2738/$ 10.00 © 1994 ESCOM Science Publishers B.V.

target internal. It has now been demonstrated (Torda et al., 1989, 1990; Pearlman and Kollman, 1991a; Schmitz et al., 1992; Pearlman, 1993) that for distance restraints a better model for MD-based NMR refinement may be obtained by replacing the instantaneous internal value in Eq. 1 with its time-averaged equivalent

$$E_I = K_I\,(< I > - I_0)^2 \tag{2}$$

where I is associated with a distance r and < I > is the appropriately weighed averaged value of I over the MD simulation.

Recently, the time-averaged approach has been extended to experimentally derived vicinal ^{3}J-coupling value restraints (Torda et al., 1993), leading to the appropriate analogue of Eq. 2:

$$E_J = K_J\,(< J > - J_0)^2 \tag{3}$$

In this study, both time-averaged and conventional J-coupling refinement methods were applied to antamanide, using experimentally measured J-coupling constants as target values. A comparison of the results led to the conclusion that time-averaged J-coupling restraints allow more accurate reproduction of the experimental J-values, while imposing less artificial rigidity on the system over the course of the MD simulation.

However, because the refinements in the previous study were carried out against experimentally measured J-coupling values, the 'true' variability of these coupling constants (or of the underlying torsion angles) was not known. Thus, for example, it was not possible to determine if time-averaged J-coupling restraints resulted in a more *realistic* ensemble of conformers, or merely in one that better satisfied the modified restraint function while introducing other artifacts into the calculation. Here we attempt to probe more deeply the effects of time-averaged J-coupling refinement by applying such restraints to a set of simulated model data. Unrestrained MD is run on our system, from which we can derive a set of averaged J-coupling constants which serve as the target values in a series of MD refinement simulations on the same system. By comparison of the unrestrained and restrained trajectories, we can unambiguously determine both the relative efficacies of standard and time-averaged J-coupling refinement, and what biases these methods introduce.

This scheme for evaluating the restraints is analogous to the one we previously used in theoretically evaluating time-averaged distance restraints for DNA (Pearlman and Kollman, 1991a) and for the FK506 macrocycle (Pearlman, 1993). The advantages of refining against data obtained from free MD of a model system are that (a) we know the 'true' conformational behavior of our model system to whatever degree of accuracy we desire; and (b) any biases due to the molecular mechanics force field used during MD will cancel, since the unrestrained ('experimental') and restrained simulations use the same force field. Additionally, the J-coupling average is a simple linear average of the instantaneous values; we therefore avoid ambiguities in choosing the correct effective weighing factor to use in determining the running average – a problem that can arise when applying time-averaged distance restraints (Kessler et al., 1988).

Here, we study the effects of J-coupling restraints for the nucleic acid constituent (ribo)-adenosine nucleoside (Fig. 1). This system was chosen for its relative simplicity and because it provides a commonly occurring example of a system for which J-coupling data can be measured.

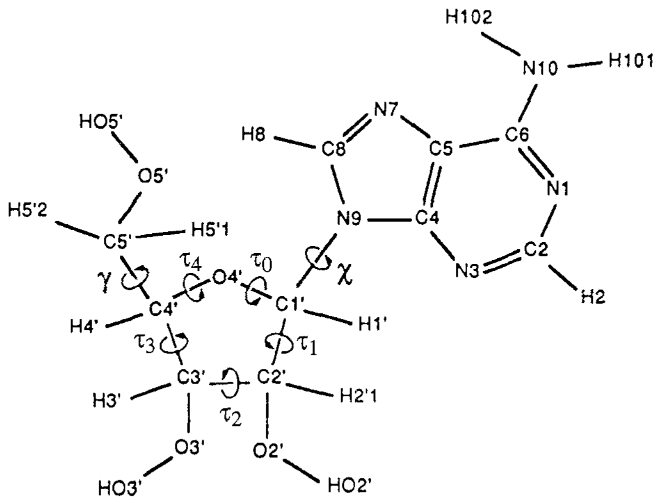

Fig. 1. Schematic representation of the adenosine nucleoside. Torsion angles γ (C3'-C4'-C5'-O5'), τ_0 (C4'-O4'-C1'-C2'), τ_1 (O4'-C1'-C2'-C3'), τ_2 (C1'-C2'-C3'-C4'), τ_3 (C2'-C3'-C4'-O4'), τ_4 (C3'-C4'-O4'-C1'), and χ (O4'-C1'-N9-C8) are indicated.

METHODS

All MD simulations were carried out on the adenosine nucleoside, using a version of the Amber/Sander MD refinement program (Pearlman et al., 1991) specially modified to incorporate time-averaged J-coupling restraints. Force field parameters were taken from the force field by Weiner et al. (1986). An all-atom model was used, and all nonbonded interactions were included. Simulations were run in vacuo, with a distance-dependent dielectric, at constant temperature (300 K).

The J-coupling values associated with five torsions of this molecule were monitored. These were $J_{H1'H2'1}$, $J_{H2'1H3'}$, $J_{H3'H4'}$, $J_{H4'H5'1}$, and $J_{H4'H5'2}$. These correspond to the five J-values that could, in principal, be measured in an appropriate NMR study. (In practice, it is frequently the case that some of these coupling constants cannot be individually determined, especially the coupling constants for the torsions centered on C4'-C5' (Van de Ven and Hilbers, 1988; Kim et al., 1992).) The J-values were calculated from the underlying torsion angle τ, using the Karplus relationship (Karplus, 1959):

$$J = A \cos^2(\tau) + B \cos(\tau) + C \tag{4}$$

with coefficients A = 10.2, B = −0.8, and C = 0.0 (Davies, 1978). Since the same coefficients are used both to derive the target J-values and during the subsequent refinement, the exact values used are not significant. An alternative, more elaborate modification of the Karplus relationship has been proposed for nucleic acids (Haasnoot et al., 1980), but we chose to use the traditional form in this study for simplicity, and because this same form is used for other systems, so the conclusions here will be more generally applicable.

The ‘experimental’ values (and standard deviations) of the five monitored coupling constants were determined from a linear average of J(t) during the final 3 ns of a 3.2-ns unrestrained MD simulation. The five J-coupling restraints were subsequently imposed, and several series of 1-ns restrained MD simulations were carried out. No restraints, other than those for the J-values, were imposed. For comparison, simulations were carried out both with conventional and time-averaged restraints. In conventional-restraint runs, the restraints were imposed using the form:

$$E_J = \begin{cases} K_J(J_{model}(t) - J_l)^2 & J_{model}(t) < J_l \\ 0 & J_l \leq J_{model}(t) \leq J_u \\ K_J(J_{model}(t) - J_u)^2 & J_u < J_{model}(t) \end{cases} \tag{5}$$

$$J_l = J_{unres} - R\sigma \tag{6a}$$

and

$$J_u = J_{unres} + R\sigma \tag{6b}$$

where J_{unres} is the averaged J-value from the unrestrained simulation, σ is the standard deviation in J_{unres} over the unrestrained simulation, and R is a constant multiplier that was fixed at a value of between 0.0 and 0.4, depending on the simulation being run. When real experimental data are used, the flat-well region accounts for errors in the measured coupling constants. For time-averaged J-coupling restraint runs, $J_{model}(t)$ in Eq. 5 was replaced by

$$< J_{model}(t) > \frac{\approx \int_0^t e^{(t'-t)/\tau} J_{model}(t')dt'}{\int_0^t e^{(t'-t)/\tau} dt'} \tag{7}$$

The exponential damping factor is included to ensure that in long simulations the most recent values $J_{model}(t)$ have a nonnegligible effect on the MD trajectory (Torda et al., 1989). For all the simulations described herein, τ = 10 ps. Previous work has indicated that this value is reasonable for time-averaged distance restraints on a small system (Pearlman, 1993). The forces dE_J/dx required for the MD integration were evaluated using the 'pseudo-force' formulation (Torda et al., 1990; Pearlman and Kollman, 1991a), which for linear averaging ensures that the effective force constant remains constant, regardless of how many data points have been collected.

Four series of simulations were carried out in parallel for both conventional and time-averaged refinement. All series consisted of nine simulations, each of which used a different value of the restraint force constant K_J. K_J varied between 0.01 and 80.0 kcal s^2/mol, depending on the simulation. The four series of simulations differed in the width of the flat well used. The flat wells for the four series reflected R = (0.0, 0.1, 0.2, 0.4) (Eqs. 5 and 6). For analysis, only the last 800 ps of each restrained simulation was used.

RESULTS

Comparing unrestrained and J-coupling restrained trajectories

The underlying torsions corresponding to the five monitored J-values are related to two conformational variables, typically used to describe DNA structure. $J_{H1'H2'1}$, $J_{H2'1H3'}$ and $J_{H3'H4'}$ are related to torsion angles that determine the conformation of the sugar, typically described by the phase angle of pseudorotation, P (Altona and Sundaralingam, 1972). $J_{H4'H5'1}$ and $J_{H4'H5'2}$ are related to torsion angle γ = C3'-C4'-C5'-O5'. Therefore, it makes sense to carry out the bulk of the analysis in these variables. The relationship between the torsions described by the measured J-coupling values and P is given approximately as (Altona and Sundaralingam, 1972):

$$P \approx \text{atan} \frac{[\tau_4 + \tau_1 - \tau_3 - \tau_0]}{[\sin(36) + \sin(72)]\tau_2} \tag{8}$$

where τ_0–τ_4 are the endocyclic torsions C4'-O4'-C1'-C2', O4'-C1'-C2'-C3', C1'-C2'-C3'-C4', C2'-C3'-C4'-O4' and C3'-C4'-O4'-C1', respectively (Fig. 1). For an unrestrained system, geometrical considerations dictate that

$$\tau_1 \approx \text{H1'-C1'-C2'-H2'1} - 120° \tag{9a}$$

$$\tau_2 \approx \text{H2'1-C2'-C3'-H3'} \tag{9b}$$

$$\tau_3 \approx \text{H2'1-C2'-C3'-H3'} + 120° \tag{9c}$$

Ring closure and energetic considerations ensure that τ_1–τ_3 are sufficient to specify the ring conformation. Similarly, the geometry about the C4'-C5' bond dictates that

$$\gamma\,(\text{C3'-C4'-C5'-O5'}) \approx \text{H4'-C4'-C5'-H5'1} + 120° \tag{10a}$$

$$\gamma\,(\text{C3'-C4'-C5'-O5'}) \approx \text{H4'-C4'-C5'-H5'2} \tag{10b}$$

Figures 2A and 3A present the conformational variabilities of P and γ over the course of the 3-ps unrestrained simulation. This represents the conformational behavior we would ideally like to reproduce in our restrained simulations. From these figures, it can be seen that P undergoes continual significant fluctuations through the range 80–200° (O4'-endo to C3'-exo). In contrast, the γ-torsion undergoes much smaller local fluctuations, but sporadically undergoes a major conformational transition between 60° (g+) and 180° (t), though for the second half of the simulation this torsion remains almost continuously in the g+ range.

Samples of the actual conformational variability observed during refinement are presented in Figs. 2B and 3B (conventional J-coupling restraints) and 2C and 3C (time-averaged J-coupling restraints). It is clear that the conventional method yields an unrealistically narrow envelope of motion, while the time-averaged method does a good job in reproducing the unrestrained conformational variability. The time-averaged phase angle of pseudorotation profile is particularly good, and is effectively indistinguishable from that for the unrestrained simulation.

In contrast, the time-averaged γ-profile is clearly different from the profile for the unrestrained run in two important respects. First, unlike its behavior during the unrestrained run, during the time-averaged simulation the γ-torsion undergoes very frequent conformational transitions. And second, the γ-torsion samples not only the g+ and t conformers, as seen in the unrestrained simulation, but also the g− (~ 300°) conformer. To understand these differences, one must consider the nature of an averaged J-coupling constant. J is calculated from a cosine function of the underlying torsion, which means that the averaged J-value cannot be simply related back to the

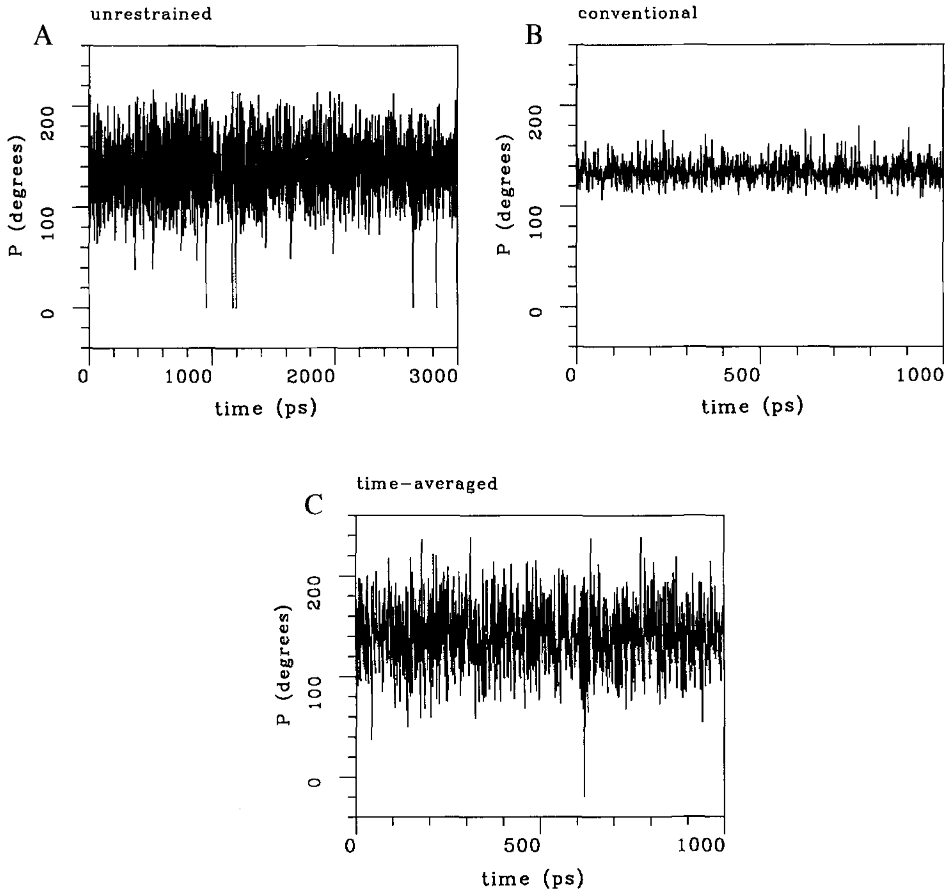

Fig. 2. The phase angle of pseudorotation P (degrees) as a function of time (ps) for (A) the 3-ns unrestrained simulation; (B) a 1-ns conventional refinement with J-coupling target values from the unrestrained simulation; and (C) a 1-ns time-averaged refinement with J-coupling target values from the unrestrained simulation. The restrained simulations correspond to a force constant $K_J = 1.0$ kcal s^2/mol and no flat-well region (R = 0.0, Eqs. 5 and 6). All simulations were run at T = 300 K. Note that the unrestrained simulation was run for 3 ns, while the restrained simulations were run for 1 ns each.

averaged torsion angle. At a very simple level, consider an angle τ that is 0° half the time and 90° half the time. The averaged value of $\cos(\tau)$ is then $(\cos(0) + \cos(90))/2 = 0.5$. But $\mathrm{acos}(0.5) = 60°$, not 45°. The implications of nonlinear averaging for J-coupling restraints is that not all the minima in the restraint profile will necessarily coincide with the minima that contributed most to the averaged J-value.

In particular, the restraint profile as a function of γ is shown in Fig. 4A, as calculated from Eqs. 4–6 and 10, and the values in Table 1. Note that, although the averaged J-values were derived from a two-state (g+, t) equilibrium, the effective restraint potential has three minima. This complex curve arises from the sum of the individual restraints on the H4'-C4'-C5'-H5'1 and H4'-C4'-C5'-H5'2 torsions (Fig. 5A,B). It is clear that the restraint profile in torsional space does not directly mirror the behavior of the data that contributed to it. Significantly, the minima in the

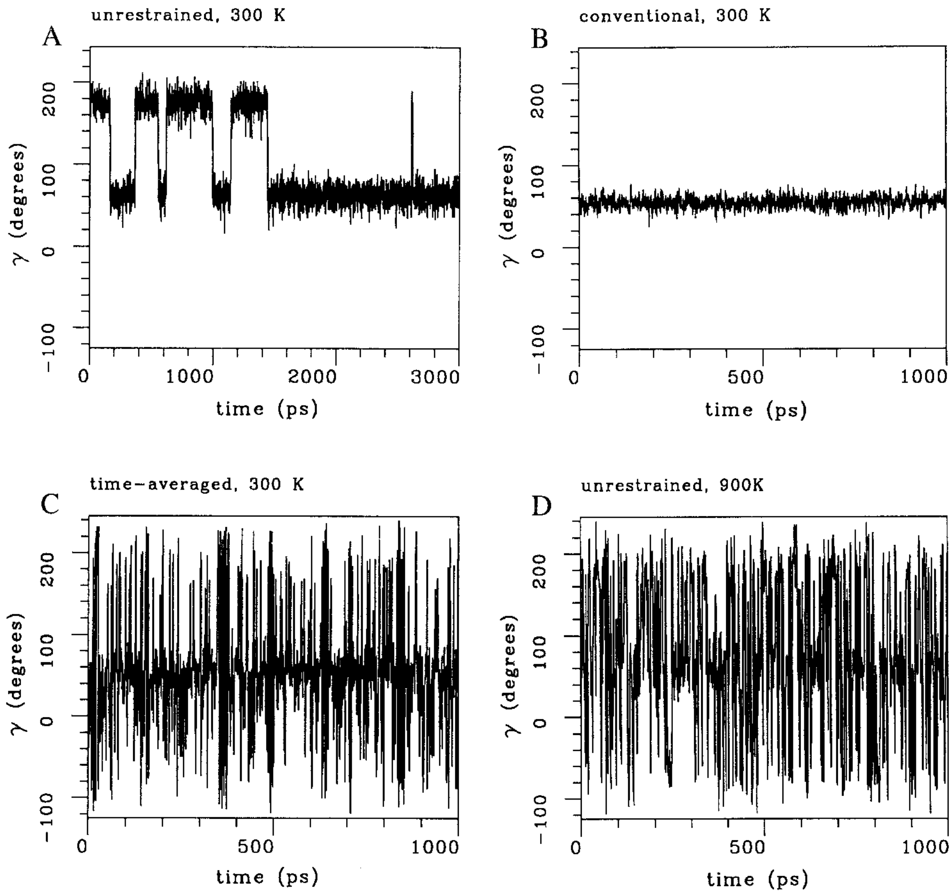

Fig. 3. The γ (C3'-C4'-C5'-O5') torsion (degrees) as a function of time (ps) for (A) the 3-ns unrestrained simulation at 300 K; (B) a 1-ns conventional refinement with J-coupling target values from the unrestrained simulation at 300 K; (C) a 1-ns time-averaged refinement with J-coupling target values from the unrestrained simulation at 300 K; and (D) a 1-ns unrestrained simulation at 900 K. The restrained simulations correspond to a force constant $K_J = 1.0$ kcal s^2/mol and no flat-well region (R = 0.0, Eqs. 5 and 6).

energetic profile for free rotation about the γ-torsion occur at the standard g+, t and g− positions (Pearlman and Kollman, 1991b). The effects of the added restraint will then be to destabilize the t and g− rotomers. This holds regardless of the width of the flat region (R in Eq. 6), at least within the reasonable range of R between 0 and 1. In the complex relationship between E_J and the torsion angle τ defined by Eqs. 4–6, an increase in the width of the flat well has the effect of flattening the peaks in $E_J(\tau)$ and altering their respective heights, but it only slightly modifies their locations.

Due to the large A (= 10.2) coefficient in the Karplus equation, the heights of the barriers in the γ-restraint profile are quite high, and cannot be overcome during a standard MD simulation at 300 K. Thus, in the conventional treatment, the net effect of the J-coupling restraints will be to force the system to remain within one of the minimum-energy wells of the γ-restraint profile. Since γ adopted the g+ rotomer in the starting conformation, it remained g+ throughout the simulation

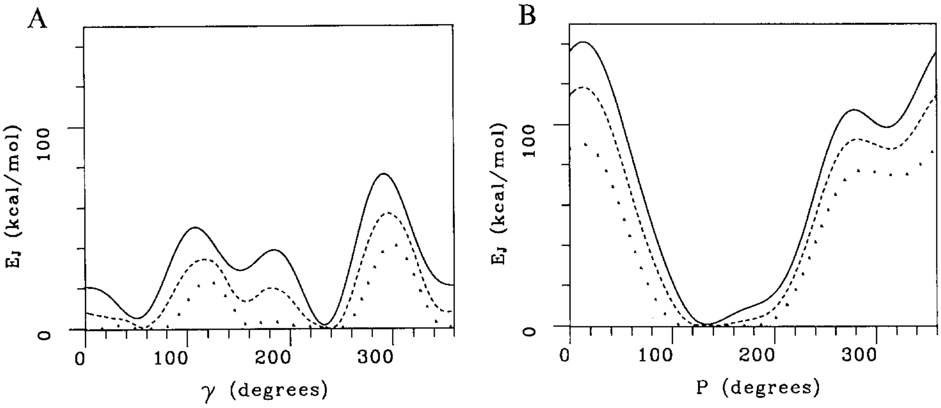

Fig. 4. (A) Profiles of the J-coupling restraint energy, E_J, as a function of the γ (C3'-C4'-C5'-O5') torsion. These profiles were calculated using the averaged J-values (and standard deviations) from the unrestrained simulation (Table 1) as target values, using Eqs. 4–6, and assuming that the relationships in Eqs. 9 and 10 strictly hold. For example, the curves in A were calculated as $E_J = E_{J(H4'\text{-}C4'\text{-}C5'\text{-}H5'1)} + E_{J(H4'\text{-}C4'\text{-}C5'\text{-}H5'2)}$. $E_{J(H4'\text{-}C4'\text{-}C5'\text{-}H5'1)}$ and $E_{J(H4'\text{-}C4'\text{-}C5'\text{-}H5'2)}$ were calculated using Eqs. 5 and 6 with J_{unres} = 2.90 and 4.86 Hz, respectively, σ = 1.69 and 4.27 Hz, respectively, and K_J = 1.0 kcal s^2/mol. Each value of γ is translated into values of J_{model} using the Karplus relationship and Eq. 10: $J_{H4'\text{-}C4'\text{-}C5'\text{-}H5'1} = 10.2\cos^2(\gamma - 120) - 0.8\cos(\gamma - 120)$ and $J_{H4'\text{-}C4'\text{-}C5'\text{-}H5'2} = 10.2\cos^2(\gamma) - 0.8\cos(\gamma)$. The curves in B were calculated similarly. The three curves in each figure correspond to differing values of the flat-well region (Eqs. 5 and 6). Solid curve: R = 0.0; dashed curve: R = 0.4; dotted curve: R = 1.0. (B) Profiles of the J-coupling restraint energy, E_J, as a function of the phase angle of pseudorotation, P. Solid curve: R = 0.0; dashed curve: R = 0.4; dotted curve: R = 1.0.

(Fig. 3B). For time-averaged refinement, on the other hand, the averaged value < J > will significantly lag the instantaneous value J(t) (e.g. see Fig. 12C). This effectively allows the system to 'tunnel' through large barriers in the restraint profile. So, for example, assume γ = g+, and that the averaged J-values also reflect the γ = g+ values of their underlying torsions. This means that the net restraint forces on γ will be small, since this corresponds to a low-energy minimum in the profile. Now, say that γ undergoes a transition to the t-state. The lag between < J > and J(t) means that this transition can occur without any initial barrier due to the restraints. After a period of time where γ = t, < J > will start to reflect J(t), resulting in a net force destabilizing the t-state. This force may push γ back towards g+, or it may push γ towards g−. This cycle – a transition between energetically stable rotomers followed by destabilization as < J > reflects J(t) – continues throughout the simulation. The end result is as seen in Fig. 3C: relatively frequent interconversion among the low-energy conformers available to the torsion.

TABLE 1
AVERAGED J-VALUES FROM AN UNRESTRAINED MD TRAJECTORY

Torsion	H1'-C1'-C2'-H2'1	H2'1-C2'-C3'-H3'	H3'-C3'-C4'-H4'	H4'-C4'-C5'-H5'1	H4'-C4'-C5'-H5'2
J	9.61(1.19)	6.79(1.53)	2.55(2.61)	2.90(1.69)	4.86(4.27)

Values are in Hz. J-values represent a linear average over the last 3 ns of an unrestrained 3.2-ns simulation. Values in parentheses are standard deviations. At every timepoint, J is calculated from the instantaneous torsion angle $\tau(t)$ using the Karplus relationship (Eq. 4) with coefficients A = 10.2, B = −0.8, C = 0.0.

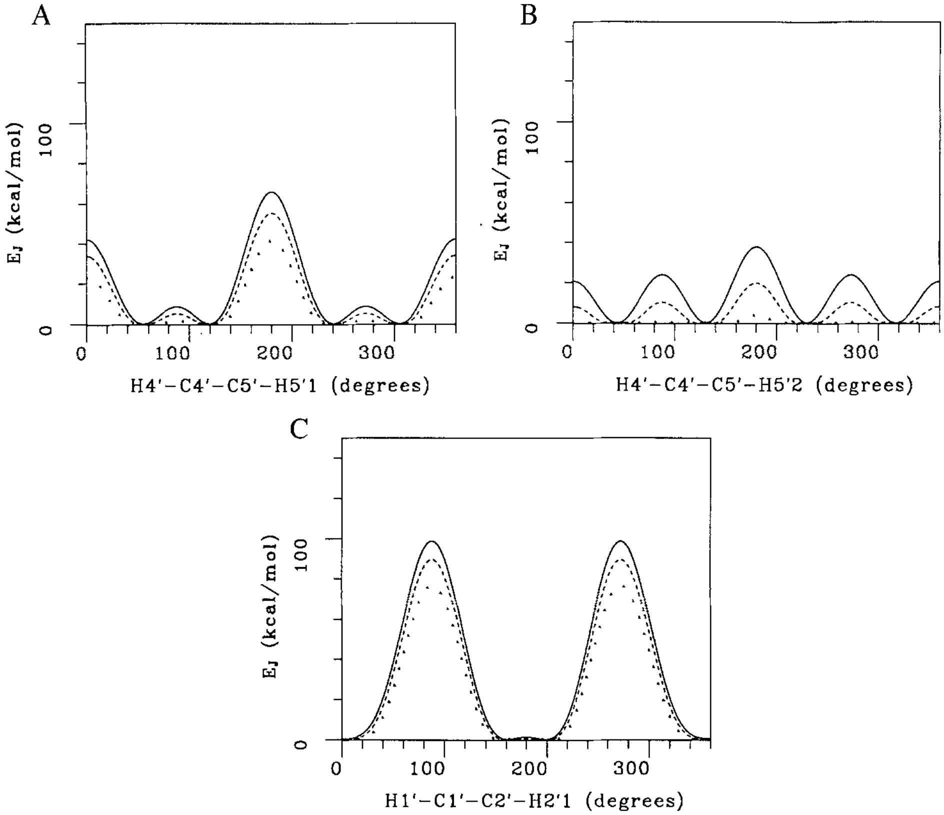

Fig. 5. Profiles of three of the J-coupling restraint energies, E_J, as a function of the underlying torsion angle. These profiles were calculated using the averaged J-value (and standard deviation) from the unrestrained simulation (Table 1), using an analogous procedure to that described in the legend to Fig. 4. In each graph, the three curves correspond to differing values of the flat-well region (Eqs. 5 and 6). Solid curve: R = 0.0; dashed curve: R = 0.4; dotted curve: R = 1.0. (A) E_J versus H4'-C4'-C5'-H5'1; (B) E_J versus H4'-C4'-C5'-H5'2; (C) E_J versus H1'-C1'-C2'-H2'1.

This is not particularly troublesome: Despite the frequent transitions, the correct averaged value of J is reproduced, and there is no reason to *expect* to be able to extract reliable temporal information out of time-averaged experimental NMR data. What *is* more questionable is the frequent sampling of the g− in the time-averaged simulation when this rotomer did not contribute to the J-values being used as restraints. Again, this is a result of the continual destabilization imposed by the time-coupling restraints, as described above. These restraints can act to push the torsion through potential energy barriers that would otherwise infrequently be transversed during a standard MD simulation at 300 K. That is, at a qualitative level they can be thought of as locally heating the system about the torsion.

In the present situation, where we are attempting to reproduce an artificial set of J-coupling values from a finite MD trajectory, this is a disadvantage, because the g− conformer was not observed in this trajectory. On the other hand, the g− conformer is of reasonably low energy. Both experimental studies (Davies, 1978; De Leeuw et al., 1980) and theoretical calculations (Pearlman and Kollman, 1991b) indicate that this rotomer is adopted roughly 10–20% of the time. Indeed, if we raise the temperature of the system and run an unrestrained simulation, we do find substantial sampling of the g− rotomer (Fig. 3D).

In essence, then, the effect of the time-averaged J-coupling restraints on the γ-torsions is to 'catalyze' rotation past modest barriers which might otherwise inhibit this interconversion on short MD time scales. Note that all the conformations sampled are reasonable in energy; the g− conformation would not be significantly sampled if it was energetically prohibited. In an alternative unrestrained trajectory, or if the trajectory were run long enough (on a time scale acceptable by experimental standards, but computationally prohibitive by simulation standards), one expects this conformer would eventually be sampled, just as it was sampled in a higher temperature unrestrained simulation. More importantly, if the force field is accurate, then when real data are used, the measured J-coupling constants *will* reflect all acceptably low-energy conformers separated by modest barriers, and the effects of the time-averaged restraints seen here would be an advantage.

Nonetheless, the results for γ here demonstrate that the form of the J-coupling restraint is relatively complex, and there can be more than one weighed average solution which appropriately reproduces the target J-coupling restraint value. The chance of such artifacts will likely be reduced in actual refinement calculations, where considerable distance restraint information will also be included.

All the calculations described above used a value of 10 ps for the exponential damping factor τ (Eq. 6). To ensure that the observations and conclusions described are not an artifact of τ, a few test simulations were run using different values for this parameter. For values of τ in the range of 25–100 ps, the frequency of conformational transitions in the γ-profile slowly decreased as τ increased (these transitions continued to occur to all conformers (g+, t, g−), just at a lower frequency). A smaller value of $\tau = 5$ ps, half of the value used above, similarly increased the rate of conformational interconversion. These findings are consistent with the above discussion, where a smaller value of τ would decrease the lag time between occurrence of a conformational transition and reflection of the transition by the average, thereby destabilizing the system.

Reproducing the appropriate means and fluctuations

It is important that we be able to reproduce the correct mean and standard deviation for the restrained variables during refinement. In Figs. 6 and 7 we plot the averages and standard deviations of P during refinement as a function of the force constant K_J used during the refinement. Figures 8 and 9 present the analogous plots for γ. In each case, four plots are presented, representing the results of refinement with a flat region (Eqs. 5 and 6) of 0, ± 0.1 σ, ± 0.2 σ, and ± 0.4 σ. In each plot, the solid line represents the results of conventional refinement, while the dashed line represents the results of time-averaged refinement. The dotted horizontal line represents the value from the unrestrained ('true') MD trajectory.

From Fig. 6, it is seen that both conventional and time-averaged refinement do an excellent job in reproducing the averaged phase angle of pseudorotation P. One exception occurs when no flat-well region is used with large values of the force constant. In this case, standard refinement apparently gets trapped in an alternative minimum centered around C3'-exo (~ 200°). The two methods are not comparable, however, in providing an appropriate description of the rms motion of the sugar ring. In Fig. 7, we see that time-averaged refinement consistently allows considerably more conformational flexibility in the sugar ring than conventional refinement, in better accord

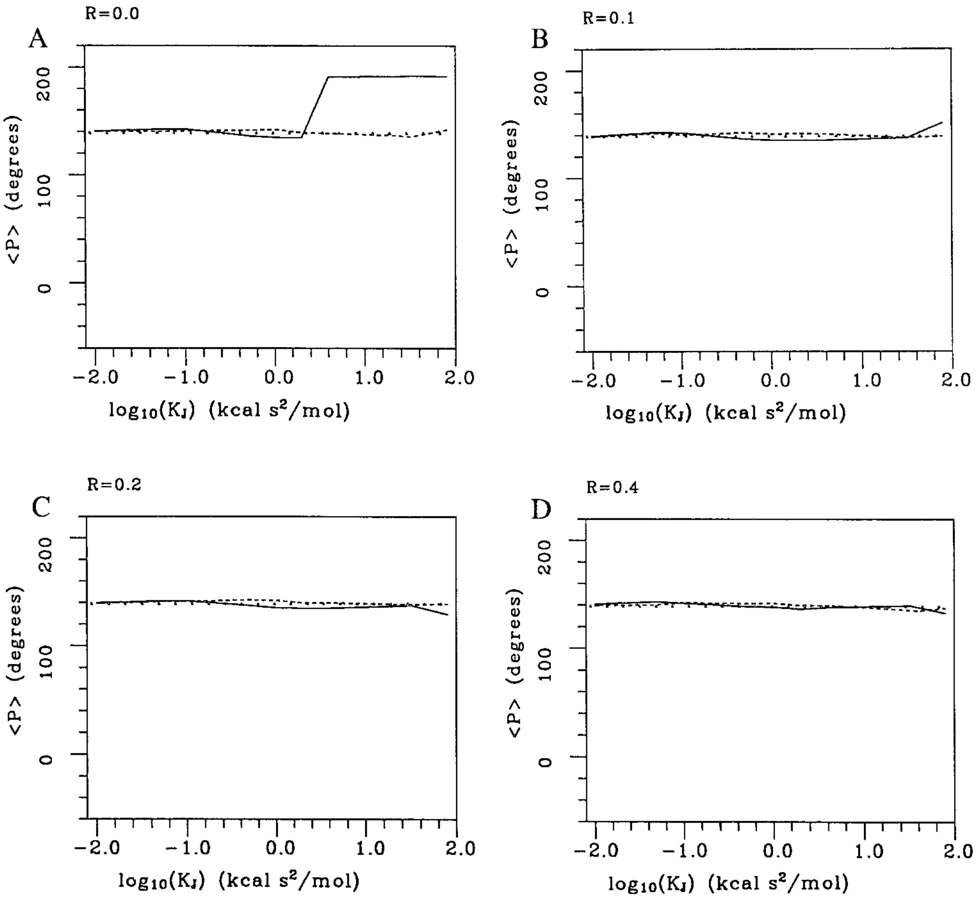

Fig. 6. The averaged value of the phase angle of pseudorotation (< P >) over the course of J-coupling refinement, as a function of $\log_{10}$ of the force constant K_J (kcal s^2/mol) used to impose the J-coupling restraints. Solid line: standard restraints; dashed line: time-averaged restraints; dotted horizontal line: the average value from the unrestrained trajectory from which the J-coupling restraints were derived. The four plots correspond to varying widths of the flat-well region (R = 0.0, 0.1, 0.2 and 0.4; Eqs. 5 and 6). Note that the dotted line representing the unrestrained average coincides very closely with the dashed curve (time-averaged restraints), and is therefore difficult to distinguish in these plots.

with the behavior for the unrestrained simulation. As the force constant is raised, the rms motion for the traditional refinement simulation gets progressively smaller, while that for the time-averaged refinement undergoes gradual increase. These are the same trends previously seen for NOE distance-based refinement. As has been noted (Pearlman, 1993), the increase in variance as the force constant increases for time-averaged refinement can be attributed to the slow variation of the restraint forces with respect to changes in the underlying restrained variables. In general, the flexibility of the refined structure is in good agreement with the flexibility of the unrestrained model for somewhat larger values of K_J when a larger flat region (larger R in Eq. 6) is used. This is not unexpected, since for the range of R-values considered here, increasing the width of the flat region effectively decreases the heights of the peaks in the restraint function (Fig. 4) – counteracting increases in K_J.

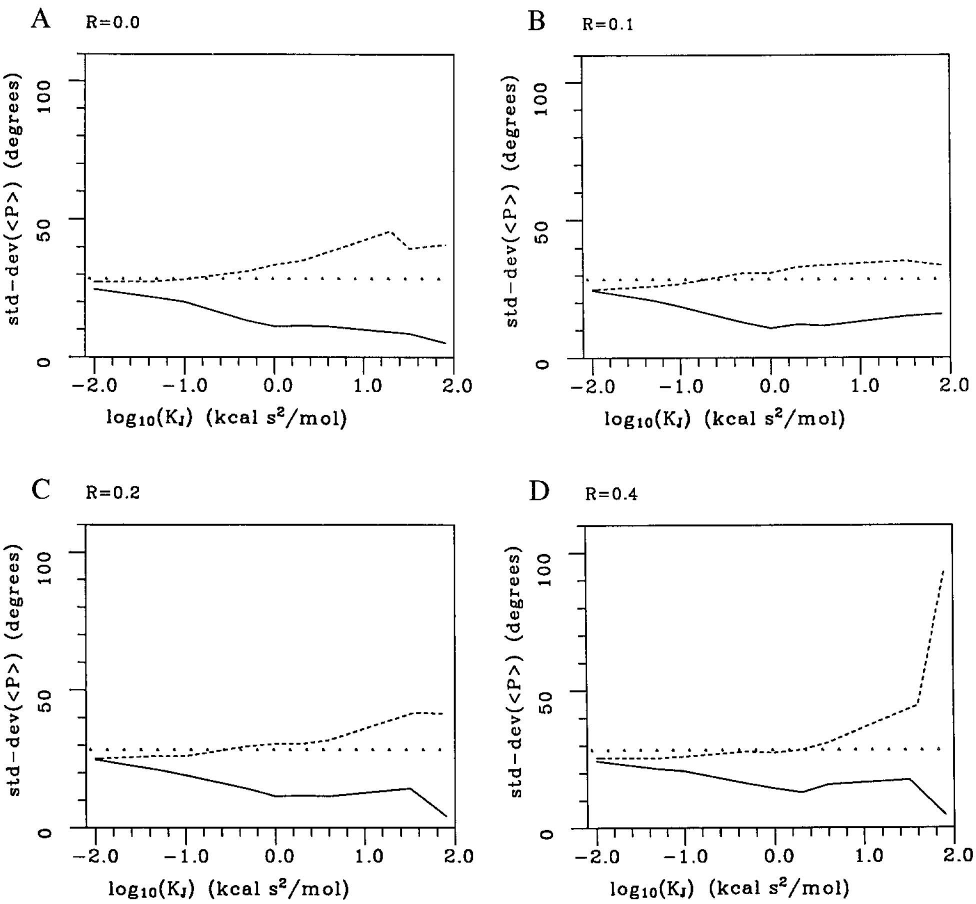

Fig. 7. The standard deviation in the averaged value of the phase angle of pseudorotation (< P >) over the course of J-coupling refinement, as a function of $\log_{10}$ of the force constant K_J (kcal s^2/mol) used to impose the J-coupling restraints. Solid line: standard restraints; dashed line: time-averaged restraints; dotted horizontal line: the average value from the unrestrained trajectory from which the J-coupling restraints were derived. The four plots correspond to varying widths of the flat-well region (R = 0.0, 0.1, 0.2 and 0.4; Eqs. 5 and 6).

In Fig. 8, it is seen that both conventional and time-averaged refinement do a poorer job of reproducing the averaged value for γ than was the case for P. Both types of refinement result in an averaged value that is 20–30° too small, though a somewhat better agreement can be achieved in some cases with high force constants and time-averaged restraints. This can be attributed to differences between the unrestrained and restrained γ-trajectories. As was the case for P, the two refinement methods reproduce approximately the same averaged value. But once again, time-averaged refinement results in rms motion about the γ-torsion that is closer to that observed during the unrestrained trajectory (Fig. 9). As for P, conventional refinement results in a severely overdamped picture of motion about the mean, and the motion decreases as the force constant increases. In the case of time-averaged refinement, the motion better estimates the unrestrained motion for all force constants, but overestimates the motion for large force constants. At a qualitative level, this can be attributed to the localized 'heating effect', arising from the differing

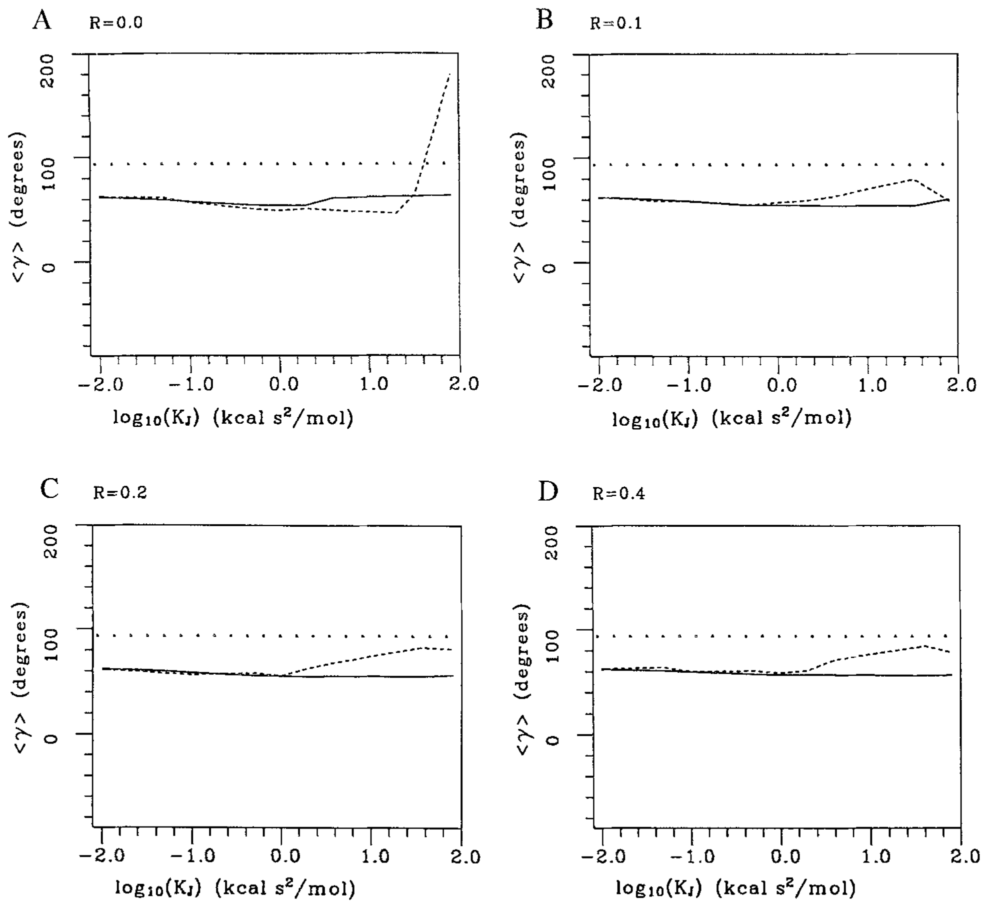

Fig. 8. The averaged value of the γ-torsion (C3'-C4'-C5'-O5') over the course of J-coupling refinement, as a function of $\log_{10}$ of the force constant K_J (kcal s^2/mol) used to impose the J-coupling restraints. Solid line: standard restraints; dashed line: time-averaged restraints; dotted horizontal line: the average value from the unrestrained trajectory from which the J-coupling restraints were derived. The four plots correspond to varying widths of the flat-well region (R = 0.0, 0.1, 0.2 and 0.4; Eqs. 5 and 6).

minima in the potential-energy and restraint curves, as discussed earlier, and from sampling the g− rotomer, which was not seen in the unrestrained simulation. As was seen for P, the degree of overestimation for the same force constant K_J decreases somewhat as the width of the flat well increases, consistent with decreases in the restraint barrier heights (counteracting increases in K_J) as these flat wells are made wider (Figs. 4 and 5).

The superior ability of the time-averaged simulations to reflect the fluctuations of the unrestrained simulation also leads to lower restraint violations for these time-averaged runs. Figure 10 plots the rms deviation between the five target J-values and the final J(t) (conventional restraints = solid line) or $< J >$ (time-averaged restraints = dashed line), as a function of $\log_{10}$ (K_J). For every value of K_J, the use of time-averaged restraints better reproduces the target values. As expected, the rms deviations are inversely proportional to the force constant used, although little improvement occurs for K_J values larger than about 0.4 kcal s^2/mol.

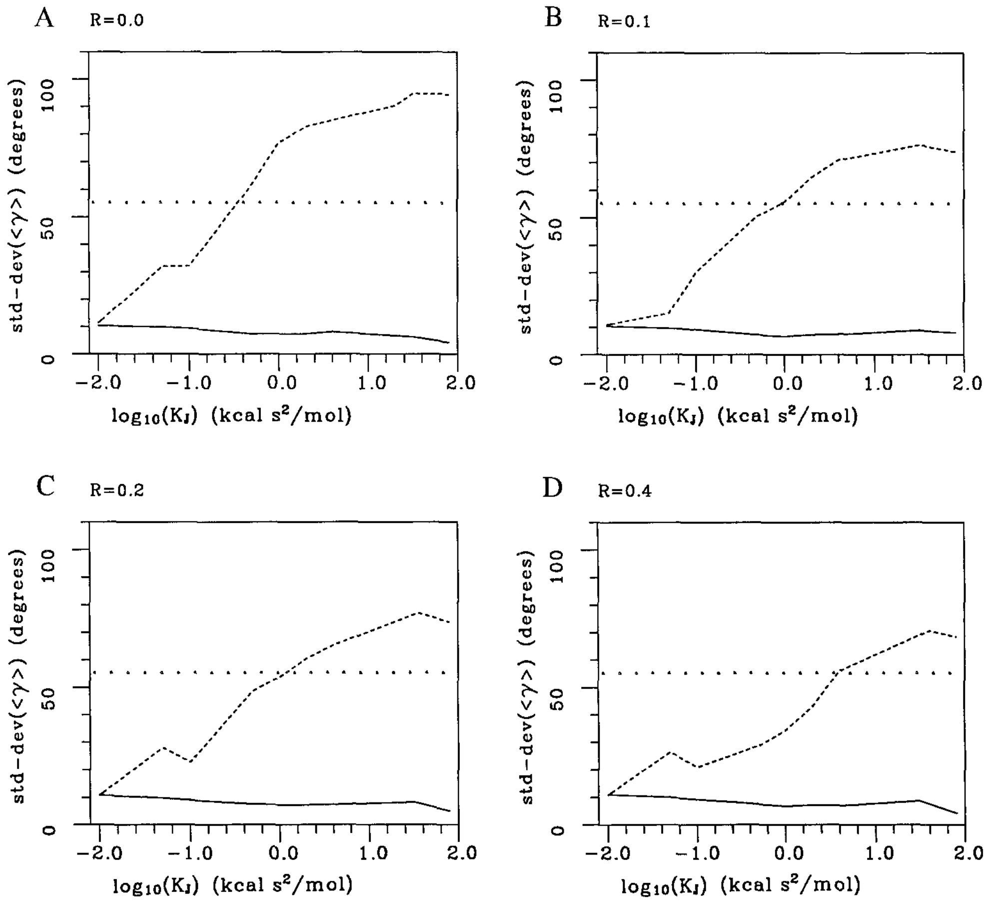

Fig. 9. The standard deviation in the averaged value of the γ-torsion (C3'-C4'-C5'-O5') over the course of J-coupling refinement, as a function of $\log_{10}$ of the force constant K_J (kcal s^2/mol) used to impose the J-coupling restraints. Solid line: standard restraints; dashed line: time-averaged restraints; dotted horizontal line: the average value from the unrestrained trajectory from which the J-coupling restraints were derived. The four plots correspond to varying widths of the flat-well region (R = 0.0, 0.1, 0.2 and 0.4; Eqs. 5 and 6).

From Figs. 7, 9 and 10, it appears that the use of time-averaged restraints with a moderate force constant of around $10^0 = 1.0$ kcal s^2/mol results in the most appropriate description of average structure and rms motion for this system.

How large a force constant is required to induce change?

Although it is important that a moderate force constant such as this can be used to reproduce the unrestrained ('true') conformational behavior, it is also necessary that such a force constant be capable of inducing conformational changes. This is because, in actual refinement, one needs the experimentally derived restraints to be able to counteract the intrinsic preferences of the force field (where the two disagree). Figure 10 shows that force constants of $K_J \geq 0.4$ kcal s^2/mol are large enough when the restraint target values are consistent with the unrestrained preferences of the force field. To test the efficacy of small force constants with time-averaged restraints in cases

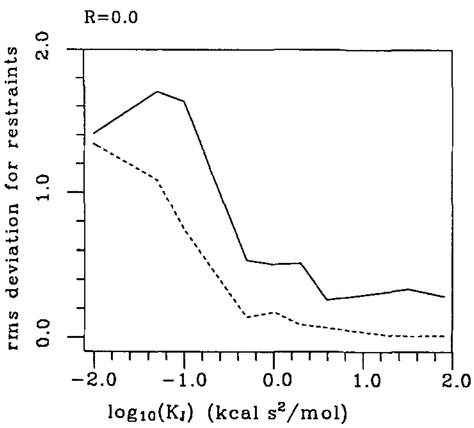

Fig. 10. The rms deviation between the five target J-values, J_0, and the values adopted during refinement, as a function of $\log_{10}$ of the force constant K_J (kcal s^2/mol). All plotted data are for refinements run with no flat-well region (R = 0.0). Data for other flat-well widths are similar. Solid line: conventional refinement. Rms deviation calculated between J_0 and the final J(t); dashed line: time-averaged refinement. Rms deviation calculated between J_0 and the averaged values $< J >$.

where the target J-values are different from those that would be determined in an unrestrained simulation, a series of refinements was run where target J-values corresponding to a pseudorotation angle P = 20° (C3'-endo) were used. These were estimated from the Karplus equation, Eq. 4, Eqs. 8 and 9, and the approximate relationship (Altona and Sundaralingam, 1972):

$$\tau = 38.7° \cos([i - 2] \cdot 144 + P) \tag{11}$$

The J-values used were $J_{H1'H2'1} = 0.322$, $J_{H2'1H3'} = 5.970$ and $J_{H3'H4'} = 9.406$ Hz. The J-coupling restraints on the C4'-C5' torsions remained as before. P = 20° (C3'-endo) was chosen as an energetically reasonable conformation (Pearlman and Kim, 1985), but it was infrequently sampled in the unrestrained simulation (Fig. 2A).

Figure 11 presents the results of these refinements. It can be seen that while very small force constants ($K_J \leq 0.05$ kcal s^2/mol) are not sufficient to overcome the intrinsic preferences of the force field, a force constant of only 0.4 kcal s^2/mol *is* adequate. This is similar to the implications of the rms deviations in Fig. 10, where the restraints were derived from the unrestrained simulation. The agreement between the refined averaged value and the target value improves only very slightly as the force constant is further increased. Thus, a force constant of around 1.0 kcal s^2/mol, which appeared to optimize the agreement between the refined model and 'experiment', should also be sufficiently large to effect the appropriate conformational changes in real refinements where force field bias will be an issue.

False minima

Although nearly all the simulations presented above were able to correctly reproduce the appropriate unrestrained average value of the phase angle of pseudorotation, when time-averaged restraints were used it was possible to get trapped in a higher energy, physically improbable, incorrect minimum. To understand how this is possible, it is necessary to again take note of the increased complexity of the penalty versus torsion-angle profile, due to the Karplus relationship

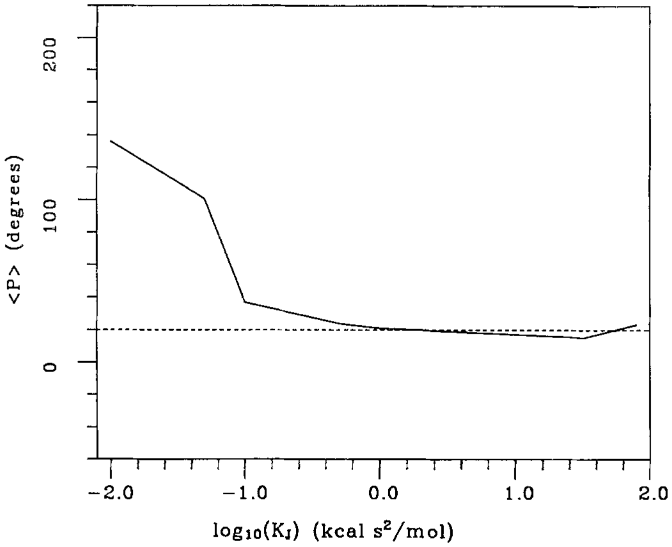

Fig. 11. The averaged value of the phase angle of pseudorotation (< P >) over the course of time-averaged J-coupling refinement, as a function of $\log_{10}$ of the force constant K_J (kcal s^2/mol). For the refinements represented in this plot, J-coupling restraints corresponding to 20° (horizontal dashed line) were used (see text).

relating the angle to the J-coupling constant. We have already seen how this affects the profile for the γ-torsion (Figs. 4 and 5). Referring again to Fig. 5, note that each E_J versus restrained angle profile consists of three maxima, and that these can be very large, even with $K_J = 1.0$ kcal s^2/mol.

In Fig. 4B, the penalty function versus the phase angle of pseudorotation is plotted. In creating this plot, we assumed that the relationships in Eqs. 9 and 11 strictly hold. The target values for the J-coupling constants in Table 1 were used. The three plots correspond to differing flat-well widths. The complexities of the three individual J-coupling restraints combine to create an even more complex surface for the pseudorotation angle. Keep in mind that it is not this profile which actually dictates the MD trajectory. The restraints are applied to the individual angles via the J-constants, such as those in Fig. 5. This profile simply shows how the restraints combine to influence P.

The crucial detail of this plot is the appearance of two minima, the first corresponding to the global minimum and true unrestrained average of around 130°, and the second corresponding to P = 320°. For a system with standard internal coordinates (i.e. where the relationships of Eqs. 9 and 11 hold), this second minimum is too high to be of consequence at 300 K. However, this is the minimum in which the system sometimes got trapped during time-averaged refinement.

How this is possible can be discerned from Fig. 12. In Fig. 12A, the pseudorotation angle is plotted versus time for a time-averaged refinement simulation which got trapped in the false minimum. At around 90 ps the system undergoes a transition, and the pseudorotation angle moves from a range about 130° to a range about 320° (the secondary minimum in Fig. 4B). In Fig. 12B, the corresponding transition in the H1'-C1'-C2'-H2'1 torsion is shown by the solid line. This torsion angle converts from a range about 160° to a range about 10°. But by reference to Eq. 9, this second range would imply

$$\tau_1 \approx \text{H1'-C1'-C2'-H2'1} - 120° \approx -110° \quad (12)$$

which is far outside the range of roughly [−40°, +40°] implied by Eq. 11. How is this possible? Recall that Eqs. 9 and 11 are derived for standard ring geometries. But if, say, the valence angles about the ring undergo significant distortion, then these relationships will no longer hold. In fact, that is what happens. For example, the value of the O4'-C1'-H1' valence angle is plotted versus time as a dashed line in Fig. 12B. At the same time P and the H1'-C1'-C2'-H2'1 torsion interconvert, the value of this angle goes from ~105° to a nonphysical value of ~130°. Other valence angles in the ring undergo similar transitions. The force constants for the aliphatic valence angles are typically around 35 kcal/mol rad^2 (Weiner et al., 1986), so the energetic cost of these transitions is relatively small ($\Delta\theta$ = 25° costs 6.7 kcal/mol), compared to the compensating reduction in the restraint energy. The net result is that the shallow high-energy minimum in the restraint profile of Fig. 4B is moved much lower, and the system can get trapped in this improbable state.

This is not generally a problem for conventional restraints. The barrier between the two minima in Fig. 4B is so large that one would likely never sample the high-energy minimum – and

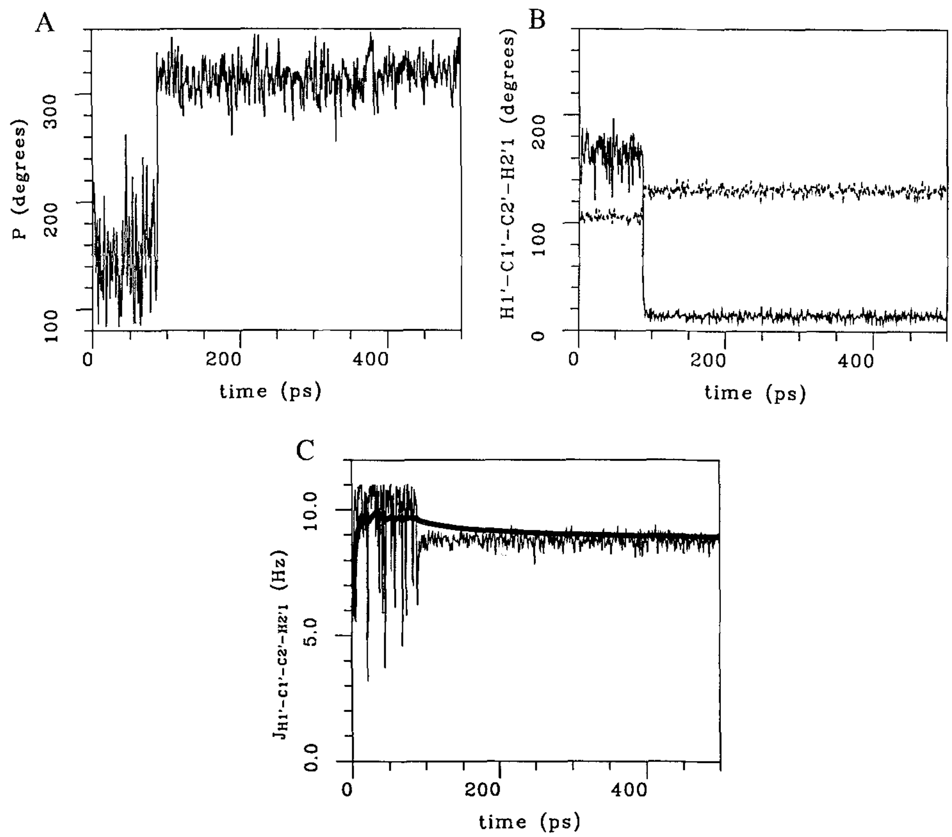

Fig. 12. Trajectories of internal coordinates and J-values as a function of time, for a time-averaged J-coupling refinement simulation which got trapped in a false minimum. For this run, K_J = 4.0 kcal s^2/mol, and R = 0.0 (no flat well). (A) The phase angle of pseudorotation, P, as a function of time. (B) Solid line: the torsion angle H1'-C2'-C2'-H2'1 as a function of time; dashed line: the valence angle O4'-C1'-H1' as a function of time. (C) Thin line: the coupling constant $J_{H1'-C1'-C2'-H2'1}(t)$ as a function of time; thick line: the running linear average of $J_{H1'-C1'-C2'-H2'1}(t)$ as a function of time.

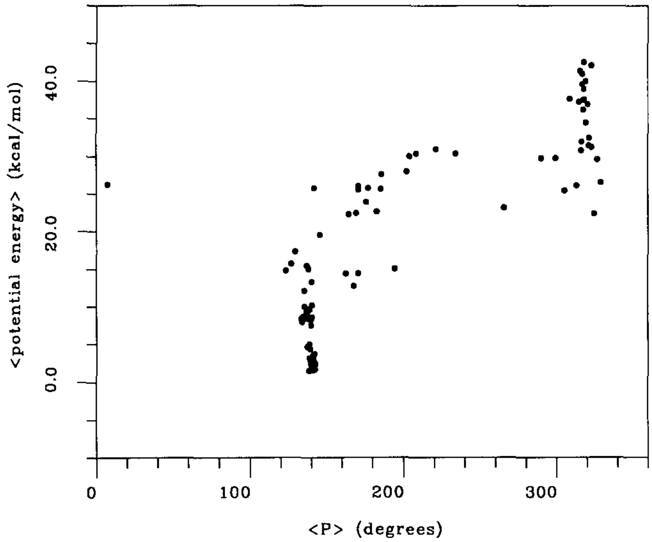

Fig. 13. The averaged potential energy for time-averaged J-coupling restraint simulations, plotted against the averaged value of the phase angle of pseudorotation, < P >, for the simulation. Note that the simulations where < P > reflects the true minimum (around 130°) include all of the lowest potential energies.

so never induce the system to undergo the improbable valence-angle transition. Equivalently, by reference to Fig. 5C, one can see that in standard refinement one would never be able to surmount the ~100 kcal/mol barrier required to allow the H1'-C1'-C2'-H2'1 torsion to convert. But with time-averaged restraints the running average lags behind the instantaneous average, which effectively allows tunneling through the high-energy barriers. In Fig. 12C, the instantaneous value of $J_{H1'H2'1}$ is plotted versus time. Overlaid on this plot, in a heavy solid line, is the running average value of $J_{H1'H2'1}$. When the system undergoes the transition in P, $J_{H1'H2'1}$ undergoes a concurrent change. But $< J_{H1'H2'1} >$, which is the value used to determine the restraint energy, changes only very slowly. Hence, the system never actually 'sees' the large barriers in Fig. 5C as the transition occurs.

As might be expected, the energies of the structures that get stuck in the alternative (incorrect) minimum are of higher energy than the best refined structure. This can be seen in Fig. 13. There, the molecular mechanics potential energy is plotted versus the averaged pseudorotation angle. All of the lowest energy structures correspond to the appropriate minimum, while all of the incorrect minimum structures are among the highest in potential energy. The general implication of this plot is that if one is imposing time-averaged J-coupling restraints, one should run a number of simulations, and that if the results diverge into more than one conformational class the potential energies should be compared. Alternatively, in this case it would have been sufficient to observe that the valence angles were far from ideal.

Four additional notes: (1) The tendency to get trapped in the false secondary minimum appeared to decrease as the force constant K_J used was lowered; (2) For a limited set of test calculations, varying the exponential damping factor τ (Eq. 7) in the range 5–100 ps had no discernible effect on the propensity for adopting the false minimum; (3) This problem seems to arise only from the rather complex relationship between the J-coupling restraints and the closed five-membered sugar ring; similar problems were not observed for time-averaged restraints applied to γ; and (4) In a real refinement, any J-coupling restraints will be in addition to a considerably larger number of distance restraints. The substantially increased amount of conformational information imposed by the combination of the two may well be sufficient to avoid nonphysical minima of the sort seen here.

DISCUSSION

We have applied and compared conventional and time-averaged J-coupling restraints during MD refinement of the adenosine nucleoside. From these results, it is apparent that time-averaged refinement is much better able to reproduce the appropriate conformational variability in the refined envelope of structures. Conventional refinement gives a very tight envelope of structures and a misleadingly low estimate of rms fluctuation. While time-averaged refinement can somewhat overestimate the rms mobility when large force constants are used, the results from these refinements are still consistently closer to 'reality' (the unrestrained MD trajectory from which the J-coupling restraints were extracted) than for conventional refinement. In the present case, both methods are capable of reasonably reproducing the correct averaged value of the underlying torsion. It is quite possible, though, that this will not be the case when refinement of actual experimental data sets (including both J-coupling values and NOE distances) is performed. In that case, it is likely that the more realistic flexibility of the time-averaged model will better allow the full restraint set to be satisfied, using an envelope of low-energy structures. This has already been shown to be true for time-averaged distance restraints (Torda et al., 1990; Pearlman and Kollman, 1991a; Schmitz et al., 1992; Pearlman, 1993).

While the time-averaged restrained trajectory reproduced the unrestrained behavior of the sugar ring quite well, detailed agreement between the unrestrained and restrained behavior for the γ-torsion was less good. As we have shown, this can be attributed to the nonlinear relationship between J and the underlying torsion angle, as given by the Karplus equation. This leads to an effective torsional restraint with multiple minima, and the locations of these minima do not necessarily coincide with the positions of the important minima in the potential-energy function that dictates the experimental behavior. The result is a 'heating effect' about the restrained torsion that promotes frequent conformational transitions among minima in the potential energy profile, and that 'catalyzes' transitions over the barriers between these minima. For real experimental data, this is probably advantageous: relatively large energy barriers can be traversed on the time scale of the experiment, and it is to our advantage to use a refinement method which can expedite these transitions on the shorter time scale of the MD simulation, while still appropriately reproducing the correct averaged statistics. On the other hand, it foiled our attempts to appropriately reproduce the unrestrained trajectory that led to the averaged J-coupling constant used in the time-averaged refinements here. The unrestrained coupling constant used as a target value in the time-averaged refinement reflected only g+ and t rotomers. The restrained trajectory frequently sampled the energetically feasible (but less likely) g− rotomer as well, while still satisfying the J-coupling restraints. The g− conformer, while not sampled during the unrestrained 300-K MD trajectory here, corresponds to an experimentally (Davies, 1978; De Leeuw et al., 1980) and theoretically (Pearlman and Kollman, 1991b) accessible minimum that could be sampled in a chaotically different or higher temperature (Fig. 3D) simulation.

The key point here is that there are multiple population solutions that will reproduce the measured averaged J-coupling restraint, and time-averaged restraints will tend to provide the solution based on sampling all low-energy conformers separated by modest barriers about the restrained torsion. It is expected that this will provide good results in any case where (a) all such conformers are accessible on the NMR time scale; and (b) the potential-energy force field used is acceptably accurate for rotation about the underlying torsion. Note that this is not equivalent to unrestrained high-temperature sampling: the value of the target J-coupling restraint will control the relative proportions of the rotomers.

For the pseudorotation angle, three J-coupling restraints combine to produce a restraint profile with only one significant minimum, which coincides with a low-energy conformation given by the potential-energy force field. Thus, the time-averaged refinement has no problem reproducing the same distribution of states (and averaged J-values) observed in the unrestrained simulation.

The results here are consistent with those from previous studies (Pearlman, 1993), and suggest that use of smaller force constants is to be preferred. In particular, these results indicate that for time-averaged J-coupling restraints, force constants in the range of 1 kcal s^2/mol are both reasonable and large enough to overcome the intrinsic preferences of the force field. Use of force constants that are too large increases the likelihood that the rms fluctuations will be overestimated, due to the effective localized torsional 'heating effect' discussed above. Bear in mind that if the conformations implied by the J-coupling restraints correspond to very unfavorable potential energies (as calculated with the force field used in the simulation), a larger force constant could be necessary. For a reasonably parameterized force field and accurate target restraint values, this should not typically be the case. Note that the use of larger flat-well regions in the restraint function acts to reduce the peak heights in the restraint potential, and so effectively counteracts the effects of larger force constants.

The lag between the averaged value of J used in the restraint expression and the current value J(t) can effectively allow 'tunneling' through rather large restraint energy barriers. For the complex interaction between J-coupling restraints and the five-membered sugar ring in the nucleoside, this can result in a structure getting trapped in a false minimum, created by a balance between the restraints and physically improbable deformations in the valence angles. Fortunately, such structures are easily separated from those refined to the appropriate minimum, either by examination of the energies of a group of refined structures, or else by examination of the internal coordinates (especially valence angles) of the resulting structures. It is also expected that in normal refinement such false minima will generally be precluded by the use of substantially larger numbers of restraints (including distance restraints). This false-minimum problem was not observed for the restrained γ-torsion, which is not part of a closed ring.

It is quite important to note that the work here assumes one can reliably determine experimental values of the J-coupling constants. In fact, accurate experimental J-coupling values can often be difficult to determine (Van de Ven and Hilbers, 1988; Kim et al., 1992; Harbison, 1993). In some cases, the J-coupling information derived is too crude to be useful in specific restraints of the type described here, although it can be used to make a gross decision as to, e.g., the quadrant in which a torsion lies (Clore and Gronenborn, 1989). This study makes no claims about the potential for actually determining quantitatively useful J-values in a real experiment. Rather, the point here is that *if* one is going to include detailed J-coupling information during refinement, then time-averaged restraints appear preferable to their conventional counterparts.

The potential advantages of using time-averaged J-coupling restraints, relative to alternative schemes that have been proposed for J-coupling refinement, have been discussed (Torda et al., 1993). This study has shown that if the experimental J-values can be accurately measured, time-averaged restraints to these values can, in fact, do a better job in reproducing the underlying conformational phenomena than conventional restraints. In conjunction with time-averaged distance refinement, whose advantages over conventional refinement have now also been documented, it should be possible to derive from NMR refinement considerably more and better information than in the past. In particular, we should now be able to generate an ensemble of structures that better reflects the true conformational variability of a molecule. Though hitherto little used in modeling studies, such an ensemble is critical if we are to use molecular structures as a basis for drug design work and other modeling studies.

ACKNOWLEDGEMENTS

We thank Chris Lepre and David Cheng for helpful discussions and for critical reading of this manuscript.

REFERENCES

Altona, C. and Sundaralingam, M. (1972) *J. Am. Chem. Soc.*, **94**, 8205–8212.
Clore, G.M. and Gronenborn, A.M. (1989) *Crit. Rev. Biochem. Mol. Biol.*, **24**, 479–564.
Clore, G.M., Nilges, M., Sukumaran, D.K., Brünger, A.T., Karplus, M. and Gronenborn, A.M. (1986) *EMBO J.*, **5**, 2729–2735.
Davies, D.B. (1978) *Prog. NMR Spectrosc.*, **12**, 135–225.
De Leeuw, H.P.M., Haasnoot, C.A.G. and Altona, C. (1980) *Isr. J. Chem.*, **20**, 108–126.
De Vlieg, J., Boelens, R., Scheek, R.M., Kaptein, R. and Van Gunsteren, W.F. (1986) *Isr. J. Chem.*, **27**, 181–188.
Haasnoot, C.A.G., De Leeuw, F.A.A.M. and Altona, C. (1980) *Tetrahedron*, **36**, 2783–2792.
Harbison, G.S. (1993) *J. Am. Chem. Soc.*, **115**, 3026–3027.
Karplus, M. (1959) *J. Chem. Phys.*, **30**, 11–15.
Kessler, H., Griesinger, C., Lautz, J., Muller, A., Van Gunsteren, W.F. and Berendsen, H.J.C. (1988) *J. Am. Chem. Soc.*, **110**, 3393–3396.
Kim, S.-G., Lin, L.-J. and Reid, B.R. (1992) *Biochemistry*, **31**, 3564–3574.
Nilges, M., Clore, G.M. and Gronenborn, A.M. (1988) *FEBS Lett.*, **229**, 317–324.
Pearlman, D.A. (1994) *J. Biomol. NMR*, **4**, 1–16.
Pearlman, D.A. and Kim, S.-H. (1985) *J. Biomol. Struct. Dyn.*, **3**, 99–125.
Pearlman, D.A. and Kollman, P.A. (1991a) *J. Mol. Biol.*, **220**, 457–479.
Pearlman, D.A. and Kollman, P.A. (1991b) *J. Am. Chem. Soc.*, **113**, 7167–7177.
Pearlman, D.A., Case, D.A., Caldwell, J.C., Seibel, G.L., Singh, U.C., Weiner, P. and Kollman, P.A. (1991) *AMBER 4.0*, University of California, San Francisco, CA.
Schmitz, U., Kumar, A. and James, T.L. (1992) *J. Am. Chem. Soc.*, **114**, 10654–10656.
Torda, A.E., Scheek, R.M. and Van Gunsteren, W.F. (1989) *Chem. Phys. Lett.*, **157**, 289–294.
Torda, A.E., Scheek, R.M. and Van Gunsteren, W.F. (1990) *J. Mol. Biol.*, **214**, 223–235.
Torda, A.E., Brunne, R.M., Huber, T., Kessler, H. and Van Gunsteren, W.F. (1993) *J. Biomol. NMR*, **3**, 55–66.
Van de Ven, F.J.M. and Hilbers, C.W. (1988) *Eur. J. Biochem.*, **178**, 1–38.
Weiner, S.J., Kollman, P.A., Nguyen, D.T. and Case, D.A. (1986) *J. Comput. Chem.*, **7**, 230–252.