J Biomol NMR (2008) 40:189–202
DOI 10.1007/s10858-008-9223-9

ARTICLE

Comparison of multiple crystal structures with NMR data for engrailed homeodomain

Tomasz L. Religa

Received: 29 October 2007 / Accepted: 24 January 2008 / Published online: 15 February 2008

Abstract Two methods are currently available to solve high resolution protein structures—X-ray crystallography and nuclear magnetic resonance (NMR). Both methods usually produce highly similar structures, but small differences between both solutions are always observed. Here the raw NMR data as well as the solved NMR structure were compared to the multiple crystal structures solved for the WT 60 residue three helix bundle engrailed homeodomain (EnHD) and single point mutants. There was excellent agreement between TALOS-predicted and crystal structure-observed dihedral angles and a good agreement for the $^{3}J(H^{N}H^{\alpha})$ couplings for the multiple crystal structures. Around 1% of NOEs were violated for any crystal structure, but no NOE was inconsistent with all of the crystal structures. Violations usually occurred for surface residues or for residues for which multiple discreet conformations were observed between the crystal structures. Comparison of the disorder shown in the multiple crystal structures shows little correlation with dynamics under native conditions for this protein.

Keywords Backbone dynamics · Crystallography · Engrailed homeodomain · NMR · Side-chain dynamics · Structure

Introduction

Two methods are currently available for obtaining high resolution protein structures and studying their dynamics: X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy (Brunger 1997). Both methods work on large ensembles of molecules thus reporting on time and spatial averages of molecular properties of the ensemble. In addition to this "average" data, both methods can provide information about the heterogeneity ("standard deviation") of the atomic positions, i.e. the amplitude and the timescale of motions present in the ensemble. Further, under favourable conditions, it is possible to determine the structures of some of the different conformational states present in the ensemble. For example, in crystallography, multiple electron densities are sometimes present for some of the amino acids (Smith et al. 1984, 1986), whereas in NMR, for conformers in a slow exchange on an NMR chemical shift scale, multiple peaks may be observed and NOEs arising from multiple conformations can be separated and assigned according to their populations (Santiveri et al. 2004).

Despite these obvious similarities in the type of structural data provided by both methods, there are many differences in the sample preparation, data acquisition, data analysis and, consequently, in the nature of the reported "average" and "standard deviation" end products. X-ray crystallography reports on the ensemble average present in a crystal lattice at low temperatures, with all the atomic properties (their position and dynamics) inferred from the X-ray diffraction pattern by a Fourier transformation. In such an analysis the quality of the final structure is, except at low resolution, directly dependent on the quality of the primary diffraction data and can be easily translated into the accuracy of the atomic positions, albeit disregarding structural heterogeneity during such structure calculation might overestimate the accuracy of the structure (DePristo et al. 2004). NMR provides specific structural information about macromolecules, such as distance between atoms

T. L. Religa (✉)
MRC Centre for Protein Engineering, Hills Road,
Cambridge CB2 0QH, UK
e-mail: tlr25@mrc-lmb.cam.ac.uk

(via NOEs), presence of hydrogen bonds (via H/D exchange or long range HNCO experiment (Cordier and Grzesiek 1999)) dihedral angles (via chemical shifts (Cornilescu et al. 1999) or J-couplings (Karplus 1959)) or orientation of bond vectors within the molecular frame (via residual dipolar couplings (RDCs) (Tjandra and Bax 1997)). The combination of these data in a simulated annealing program leads to a single structure, which satisfies all of the constraints with multiple, equally valid solutions usually being reported (leading to "NMR ensemble of structures"). For some datasets it is obvious that the data cannot be satisfied by a single structure and then it is possible to calculate structures of the different members of the ensemble (Clore and Schwieters 2006; Schwieters and Clore 2007; Zhang et al. 2007).

Comparisons of X-ray and NMR structures of identical or homologous proteins have revealed a high degree of similarity between them with the larger differences usually localised in the mobile N-, C-termini, loop regions and side-chains of surface residues (Billeter 1992; Wagner et al. 1992). Numerous reasons have been given for this discrepancy, such as: intrinsic differences between the structures in the solution and crystal, inaccuracy of the NMR data or intrinsic subjectivity involved in the structure determination process (Liu et al. 1992), lower precision of the NMR structures (Snyder et al. 2005), crystal contact artifacts (Zhang et al. 1995), or lower, than it is believed, accuracy of the crystal structures (DePristo et al. 2004). It is also possible that the intrinsic differences between multiple X-ray structures or NMR structures are due to the intrinsic dynamics of the molecule and provide a representative sampling of the heterogeneity of protein native states in the solution (as sampled by NMR) (Best et al. 2006).

Drosophila melanogaster Engrailed homeodomain (EnHD) is a 61-residue three-helix bundle protein (Fig. 1), which has been a formidable model system for studying protein—DNA interactions. Consequently, over 15 crystal structures of the protein have been determined in the DNA bound and free forms for the WT protein, or its single points mutants (Clarke et al. 1994; Fraenkel et al. 1998; Grant et al. 2000; Kissinger et al. 1990; Stollar et al. 2003; Tucker-Kellogg et al. 1997). Small variations are observed in the position of the backbone and side-chain atoms between different models, which are partly ascribed to networks of correlated conformational adjustments (Stollar et al. 2003). Bearing in mind these differences and their origin—either due to intrinsic protein dynamics (Best et al. 2006), inaccuracy of crystallographic data (DePristo et al. 2004) or crystallographic artefacts—they were compared here to structural and dynamic data obtained in the solution for EnHD (unlike what has been done mostly before—comparison of structural features). This included standard parameters measured during structure determination, such as chemical shifts, $^{3}J(H^{N}H^{\alpha})$ couplings, NOEs as well as the relaxation parameters—backbone R_1, R_2 and $\{^{1}H\}$–^{15}N NOEs and methyl group ^{2}H relaxation. In addition, the dynamics of the protein were simulated in explicit water to predict the possible influence of mobility on the measured values.

Materials and methods

Sample preparation

EnHD was expressed and purified as described previously (Mayor et al. 2000) with ^{2}H, ^{13}C, ^{15}N sample produced in K-MOPS minimal medium (Neidhard et al. 1974) in 70% D_2O, supplemented with ^{13}C glucose and $^{15}NH_4Cl$ as the sole carbon and nitrogen sources, respectively. All the experiments were performed on a 500 μM protein sample, appropriately isotope labelled, in 50 mM deuterated acetate, 100 mM NaCl, pH 5.7, 7% D_2O on a Bruker DRX500 spectrometer equipped with single axis gradient probe. A 2D NOESY spectrum (150 ms mixing time) was acquired on a Bruker DRX600 spectrometer equipped with triple axis gradient probe.

Assignments and structure determination

The backbone and side-chain assignments were performed using standard triple resonance methods (Bax et al. 1991; Wuthrich 1986) including stereo-specific assignments of methyl groups (Neri et al. 1989) with the spectra referenced using TSP (^{1}H), DSS (^{13}C) or indirectly, via gyromagnetic ratios (^{15}N). For structure determination, the volumes of the peaks from the 2D NOESY spectrum were converted into distances based on sequential NOEs from regular secondary structure elements and put into 3 categories: 1.8–2.7; 1.8–3.6; 1.8–5.5 Å. The peaks for which we were unable to determine the volume accurately were automatically put into the bin for the weakest signals. The inter-residue NOEs were not included, since the NOE peak picking was performed on an overlaid NOESY/TOCSY spectrum. Sum averaging without multiplicity correction was used for NOE constraints (Fletcher et al. 1996). In addition to these restraints, $^{3}J(H^{N}H^{\alpha})$ couplings were measured using a quantitative J-coupling method (Vuister and Bax 1993) (T_{1zz} for residues 10–56 was measured to 99 ± 13 ms, which results in a multiplication factor of 1.11), hydrogen bond restraints were obtained from H/D exchange experiments (Mayor et al. 2003) and dihedral angle restraints from TALOS (Cornilescu et al. 1999). Neither EnHD nor any other homeodomain was found in the TALOS database used.

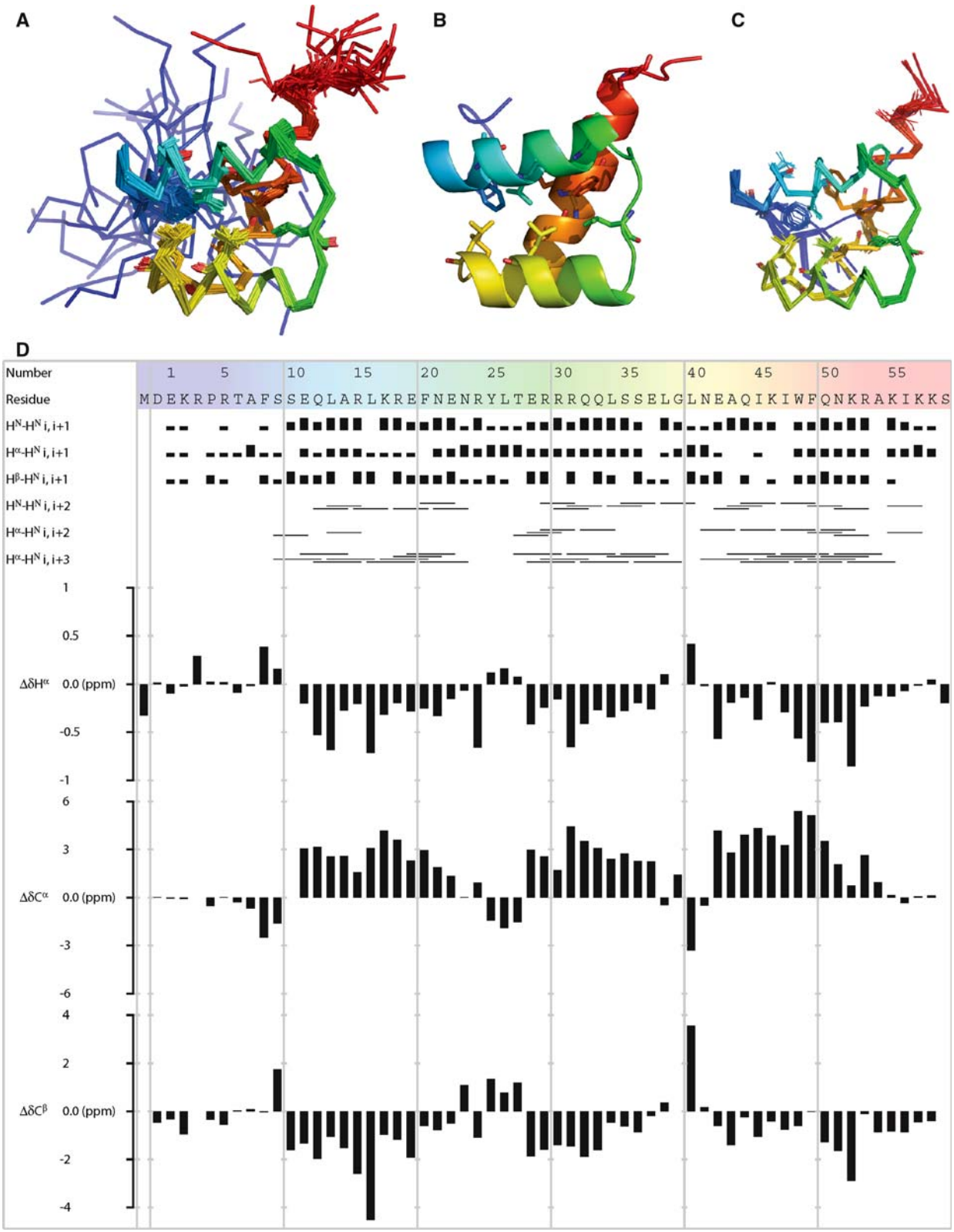

Fig. 1 Structure of EnHD (pdbID: 2JWT) and summary of sequential NOEs and secondary chemical shift values. (**a**) Bundle of the 25 lowest-energy conformers with backbone atoms N, C^{α} and C of residues 10–55 superimposed for a minimal RMSD. (**b**) Lowest energy structure. (**c**) Crystal structure ensemble obtained from 15 different crystal structures and superimposed on the backbone atoms as the NMR structures. (**d**) Sequential NOEs and H^{α}, C^{α} and C^{β} chemical shift deviations from random coil values. Only hydrophobic residues are shown in the structural models

Structures were calculated on Redhat Linux 9.0 with CNS 1.1 (Brunger et al. 1998) using an anneal.inp script with: 2,000 steps of high temperature annealing stage, 9,000 steps of torsion angle slow-cool annealing stage and 9,000 steps of second slow-cool cartesian annealing. Twenty five (out of 50) lowest energy structures were taken as the representative of the ensemble. The quality of structures was analysed using the program PROCHECK (Laskowski et al. 1996).

Chemical shifts, distance restraints as well as relaxation parameters were assigned BMRB database accession number 15536. Structure was deposited with PDB ID code 2JWT.

Residual dipolar couplings

The H^N–N RDCs were measured in 7% acrylamide gels (37:1 bisacrylamide:acrylamide ratio), radially compressed using gel stretching apparatus (Chou et al. 2001). Both isotropic and anistropic H^N–N RDC values were measured with the DSSE (Cordier et al. 1999) pulse program and processed with an NMRPipe script. The accurate chemical shifts of individual peaks were determined using the integration routine in Sparky. The RMSD between back-to-back measurement of RDC was <0.2 Hz. Comparison of the predicted and determined RDCs was carried out with PALES (Zweckstetter and Bax 2000).

Backbone dynamics

The ^{15}N T_1, T_2 and {^{1}H}–^{15}N NOE backbone H^N–N experiments (Farrow et al. 1995; Mandel et al. 1995) were acquired previously (Religa et al. 2005) on a Bruker DRX500 spectrometer and repeated here on a ^{2}H, ^{13}C, ^{15}N sample, which was also used for side-chain dynamics. For T_1 experiments 12 delays were used, set to 4, 8, 12, 20, 30, 50, 70, 90, 120, 150, 200 and 400 ms. For T_2 experiments the relaxation delays were set to: 14, 28, 42, 56, 70, 84, 119, 161, 203, 245, 294 and 350 ms. A recycle delay of 3 s was used in both cases. Heteronuclear steady-state NOE enhancement ({^{1}H}–^{15}N NOE) was calculated as peak height ratio for the spectra obtained with and without presaturation of amide proton resonances. Presaturation was achieved by applying 120° pulses for 3 s at 6 ms intervals with field strength of 25 kHz, with the reference spectrum having just a delay of the same length. The recycle delay was set to 7 s (Renner et al. 2002) in both experiments. The peak heights were fitted in Sparky 3.106 (Goddard and Kneller); errors in the measurement were determined from RMSD in the peak heights from the lower intensity spectra.

The relaxation data was analysed using Sparky 3.106, using non-linear least-squares Levenberg-Marquardt fitting algorithm and 100 steps of Monte-Carlo analysis for error estimation. The relaxation data were fitted to five motional models of increasing complexity (1:S^2; 2:S^2, τ_e; 3:S^2, R_{ex}; 4:S^2, τ_e, R_{ex}; 5:S^2_f, S^2_s, τ_e) using Tensor 2 software (Dosset et al. 2000). Briefly, the program allows estimation of the tumbling time from R_2/R_1 ratio (which was calculated here to be 5.01 ± 0.05 ns) and the suitability of the motional model by using the *F*-statistics.

Side-chain dynamics

The side-chain relaxation parameters were recorded using the published pulse sequences (Millet et al. 2002; Muhandiram et al. 1995), with a modification to improve suppression of the water signal via a WATERGATE element (Piotto et al. 1992) during the final INEPT. The delay times for the experiments were (in this order): 2, 270, 10, 210, 20, 160, 40, 110 and 70 ms (I_zC_z); 2, 57.7, 4.5, 45, 9.5, 36, 15, 28, 21 ms for D_z, D_+, $3D_Z^2$ - 2, $D_+D_z - D_zD_+$, D_+^2. The data was analysed using custom scripts in R 2.3.1. The data quality was confirmed using inequality and consistency relationships (Millet et al. 2002). The spectral density values at 0, 76.7 and 153.5 MHz were obtained from fitting the relaxation equation (Skrynnikov et al. 2002) using glm.fit function with inverse of variances as weights in the fitting process. These values were then used to extract side-chain order parameters, fitted to various motional models (S_f^2, τ_f in LS-2; S_f^2, τ_f, τ_c^{eff} in LS-3) (Skrynnikov et al. 2002). The inaccuracy of spectral density fitting was evaluated using 100 Monte Carlo steps and the adequacy of the Lipari-Szabo models using an *F*-test. A three parameter fit (LS-3) was not a statistically significant for any of the residues. The R scripts are available upon request.

Molecular dynamics

All molecular dynamics simulations were performed in GROMACS 3.3 (Lindahl et al. 2001; Van der Spoel et al. 2005) in a rhombic dodecahedron box, with a box size equal to the diameter of the system (largest distance between two atoms in the protein) plus 24 Å. This led to a minimal shell of water of 12 Å from the surface of the protein to the edge of the box made of over 3,800 water molecules. The system was neutralised by adding Cl^- ions. The energy of the system was minimised by the steepest descent algorithm and a 10 ps simulation was performed during which positions of the protein atoms were restrained. The system prepared in this way was used for

the MD simulation at 298 K in AMBER99ϕ (SPC water model) and OPLS/AA (TIP4P water model). To maintain the temperature and pressure (at 1 atm), a Berendsen et al. (1984) thermostat was used with the coupling constants of 0.1 ps and 1 ps, respectively. A cut-off of 10 Å was used for the van der Waals interactions and the particle mesh Ewald algorithm (Essmann et al. 1995) was used for Coulomb interactions with a switching distance of 10 Å. Neighbour lists were updated every tenth integration step. Bond lengths were constrained using the LINCS algorithm (Hess et al. 1997) for the protein and SETTLE (Miyamoto and Kollman 1992) for the water. Every 1,000th protein structure (every 2 ps) was saved for further analysis. The analysis of the trajectory was performed using the GROMACS 3.3 tools and custom scripts in Perl.

The separation of the tumbling from the internal motion during the simulation was achieved by superimposing backbone atoms of residues 10–55 to the first structure and calculating the H^N–N order parameter (S^2_{MD}) using 1,000 ps correlation functions as (Lipari and Szabo 1982a, b):

$$S^2_{MD} = \lim_{t \to 1{,}000ps} C_I(t) = \lim_{t \to 1{,}000ps} \langle P_2(\hat{\mu}(\tau) \cdot \hat{\mu}(\tau + t)) \rangle$$

where: $C_I(t)$ is the internal correlation function, approximated as a single exponential decay from unity to a limiting value S^2_{MD}, with a time constant τ_e. It is equal to the second order Legendre polynominal (P_2) of the change in the orientation of the normalised intermolecular bond vector with time ($\hat{\mu}(0)\hat{\mu}(t)$) The order parameters were calculated using the above equation for each ns of the simulation leading to 9 correlation functions/residue. Occurrence of rare transitions (motions, which are not sampled fully within the ns of the simulation) was evaluated by comparing the correlation functions for each ns and their convergence level. If such motions occurred during the simulation, they should change the correlation function and the reported value at $C(t_{max})$, leading to discrepancies between the consecutive correlation functions. Therefore the standard deviation of the order parameter was used to describe the imprecision of the evaluation due to the events which are under-sampled.

Comparison to crystal structures

Fifteen different crystal structures were used for comparison of free, DNA-bound or single point mutants of EnHD (1DU0A, 1DU0B, 1ENH, 1P7IA, 1P7IB, 1P7IC, 1P7ID, 1P7JA, 1P7JB, 1P7JC, 1P7JD, 2HDDA, 2HDDB, 3HDDA, 3HDDB). The resolutions (in Å) and R-free were, respectively: 1DU0: 2.0, 0.27; 1ENH: 2.1, 0.20, 1P7I: 2.1, 0.24; 1P7J: 2.1, 0.24; 2HDD 1.9, 0.25; 3HDD: 2.2, 0.23. The structures were superimposed on backbone atoms of residues 10–55.

Results and discussion

Backbone dihedral angles (TALOS and $^3J(H^NH^\alpha)$ couplings)

Chemical shift assignment is a first step towards the structure determination. The structural information present in chemical shifts can be translated into the backbone dihedral angles using the program TALOS (Cornilescu et al. 1999). This software searches a database of assigned proteins with high resolution crystal structures for similar backbone chemical shifts occurring in triplets of adjacent residues of the same residue type. Predicted ϕ/ψ dihedral angles are marked as "Good", if at least 9 out of 10 best matching triplets exhibit similar dihedral angles. Approximately 3 % of TALOS predictions are incorrect.

Comparison of "Good" TALOS predictions for EnHD with the backbone dihedral angles determined from 15 crystal structures is shown on Fig. 2a, b. For most of the residues, the TALOS-predicted dihedral angles were very close to the average angle determined from multiple crystal structures, suggesting that in most of the cases this predicted angle is no worse than the angle determined from crystal structure. This suggests that setting tight bounds on TALOS-predicted dihedral angles during the NMR structure determination might lead to an improved agreement with crystal structures. Further, for all of the accepted predictions in EnHD the ±2 SD boundaries included a large fraction of dihedral angles found in the crystal structures. Interestingly, the TALOS boundaries were also informative of the spread of the dihedral angle values found in the crystal structures (correlation coefficient of 0.86 for ϕ angle and 0.50 for the ψ angle), with the deviations smaller in the helical regions and larger in the loops, where some of the predictions were still marked as "Good". For example, T27 ψ angle has a TALOS SD ($\Delta\psi$) of 3.8° and we observed an SD of 4.4°; in comparison, the $\Delta\phi$ in TALOS predictions for this residue is 16° with the SD in this dihedral angle observed in the crystal structures of 15°.

Backbone dihedral angles can also be deduced from the $^3J(H^NH^\alpha)$ couplings, the magnitude of which is dependent on the ϕ angle (Pardi et al. 1983) via a well known Karplus (1959) relationship. Here the $^3J(H^NH^\alpha)$ couplings were measured using a quantitative J-coupling method (Vuister and Bax 1993) and the values were predicted from the crystal structures with the recent parameterisation of this formulae (Wang and Bax 1996).

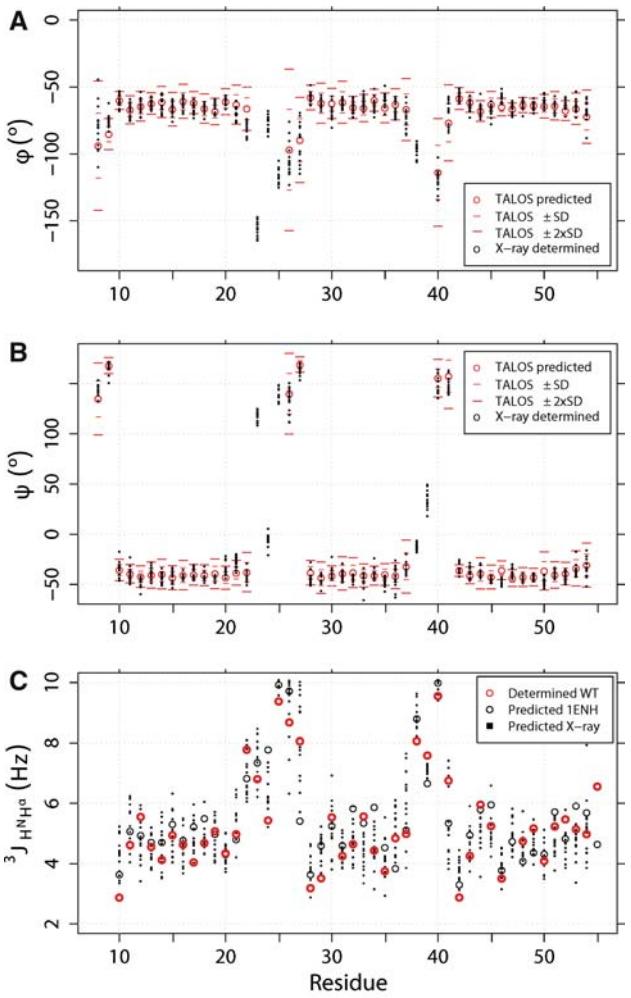

Fig. 2 (**a**) Comparison of ϕ dihedral angle prediction from TALOS with values determined from crystal structures. (**b**) Comparison of ψ dihedral angle prediction from TALOS with values determined from crystal structures. (**c**) $^3J(H^NH^\alpha)$ coupling measured in solution and compared to values calculated from crystal structures

$^3J(H^NH^\alpha)$ in EnHD WT were in good agreement with the values predicted from the crystal structure of the free, WT protein (1ENH) (see Fig. 2c). The couplings are <6 Hz for residues 10–21, 28–37 and 42–54, consistent with the presence of helical structure between these residues (Pardi et al. 1983; Smith et al. 1996). The largest deviations of the couplings from the 1ENH crystal structure occurred for residues R24, T27 and K55. The K55 J-coupling is likely to be influenced by helix fraying at the C-terminus. Comparison between the EnHD crystal structures shows that the residues in the loop region have large differences in the ϕ angle (Fig. 2a) and therefore the scalar coupling. For R24 the couplings in the crystal structures range from 5.2 to 7.8 Hz with the measured value equal to 5.4 Hz. For T27 the predicted couplings are in the 5.4–10.0 Hz range, while the measured value is 8.1 Hz. The root mean square deviation (RMSD) between the 1ENH predicted and measured couplings is 0.91 Hz for residues 10–55 and 0.70 Hz for the couplings between residues 10–55, excluding residues with deviations >1.5 Hz (R24, T27 and K55). This is as good as the RMSD obtained during the parameterisation of the Karplus equation based on the SNase crystal structure (Vuister and Bax 1993). Since there is a variation in the ϕ angle values between different crystal structures, which leads to a variation in predictions of the J-coupling (average SD between J predictions from crystal structure for each residue of 0.6 Hz), obtaining a lower RMSD from parameterisation to a single crystal structure might be difficult.

Distance restraints (NOEs)

Solving NMR structures of proteins critically depends on nuclear Overhauser enhancement (NOE) experiments, which provide information on internuclear distances (Neuhaus and Williamson 2000). Here, assigned NOEs from a NOESY spectrum acquired with 150 ms mixing time were compared with multiple crystal structures. None of the NOEs have been found to be inconsistent with all of the determined crystal structures, but no crystal structure satisfied all of the NOEs, as shown in Table 1. Two major reasons have been found for this disagreement: (a) side-chains on the surface of the protein, which displayed multiple conformations in the collection of crystal structures, (b) residues in the core of the protein, which had multiple, but discreet conformations. Differences in the positioning of the surface residues are usually ascribed to crystal contact artifacts and the disagreement with r^{-6} weighted NOEs might not be surprising. Some of the conformational differences between hydrophobic residues in the core of the protein were described by Stollar et al. (2003), suggesting that different conformational states were frozen out in the crystal lattice. For example, we observed that the L13 $H^{\delta1}$–W48$^{\varepsilon1}$ NOE is stronger than predicted from the 1ENH crystal structure. This stronger NOE could be explained either by persistent structural change occurring in solution (one state) or presence of different rotameric states of L13, as observed in the different variants of EnHD crystal structures (see Fig. 3). However, distinguishing between these two scenarios (i.e. one-state or two-state) is not possible with the current density of experimental restraints (Bonvin and Brunger 1996). The intensity of the NOEs between the two methyl groups of L26 to other atoms seemed to be inverted with respect to the distances expected from the 1ENH crystal structure. Together with the observation of the Y25 H^N–L26 H^γ NOE this suggests orientation of the L26 side-chain different to that observed in the WT crystal structure and is more consistent with the orientation seen in 5 out of 8 K52A/E mutant crystal structures as well

Table 1 Summary of conformational restraints and statistics for the 25 accepted structures of EnHD

A. Structural constraints	
NOEs	
Strong (1.8–2.7 Å)	59
Medium (1.8–3.6 Å)	148
Weak (1.8–5.5 Å)	468
Intra-residue ($\|i - j\| = 0$)	0
Sequential ($\|i - j\| = 1$)	260
Medium range ($1 < \|i - j\| < 5$)	209
Long range ($\|i - j\| \geq 5$)	206
Total	675
Hydrogen bonds (each contributes 2 restraints)	30
Dihedral angles from TALOS	72 (36 phi and 36 psi angles)
$^3J(H^NH^\alpha)$ couplings	44
B. Statistics for the accepted structures	(kcal mol^{-1})
E (overall)	57 ± 2
E (bond)	1.1 ± 1
E (angle)	32.8 ± 0.4
E (improper)	1.8 ± 0.2
E (van der Waals)	11 ± 1
E (noe)	1.6 ± 0.3
E (J-coupling)	8.0 ± 0.4
E (dihedral)	0.0 ± 0.1
Average backbone RMS from the mean structure	(Å)
Residues all	3.2 ± 1.1
Residues 10–55	0.39 ± 0.11
Average backbone pairwise RMS	(Å)
Residues all	3.3 ± 1.0 (min: 1.4; max: 6.2)
Residues 10–55	0.43 ± 0.10 (min: 0.22; max: 0.85)
Ramachandran analysis (all residues—25 models)	(%)
Residues in most favoured regions	86.3
Residues in additional allowed regions	12.1
Residues in generously allowed regions	1.1
Residues in disallowed regions	0.4
Ramachandran analysis (residues 10–55—25 models)	(%)
Residues in most favoured regions	96.4
Residues in additional allowed regions	3.6
Residues in generously allowed regions	0.0
Residues in disallowed regions	0.0

as the DNA-bound EnHD structure. This observation, however, does not exclude the possibility of alternative leucine rotameric states being present, since the difference in the NOE intensities for H^δs is small and the observation of, for example, an L26 H^γ–R30 H^N NOE, which would allow establishing an alternative conformation, is precluded by spectral overlap. Similar discrepancies in rotameric states were also found for other homeodomain structures (Table 2).

Solution structure

The solution structure of EnHD (pdbID: 2JWT) closely resembles the crystal structures of the protein (Fig. 1), with the backbone pairwise RMSD between the lowest energy structure and the 1ENH crystal structure of 0.59 Å for residues 10–55. The structural statistics for the accepted structures is summarised in Table 1. The structure is well defined with backbone RMSD of 0.39 Å for residues in 10–55 and ~14 restraints found/residue. There are only small deviations from the covalent geometry and experimental restraints. The torsion angles are mostly in the favourable region of ϕ, ψ space with no residues from 10 to 55 in the disallowed region. The maximum pairwise backbone RMSD between the members of the ensemble is 0.85 Å, which is more than the RMSD to the crystal structure. The accuracy of the structure, as judged by backbone RMSD, allowed by the collected NMR restraints should be regarded as lower than this value (Spronk et al. 2003).

While the overall fold of the protein is the same as the determined crystal structure, there are small differences between both of them. Differences in the location of the side-chains on the surface of the protein and location of the loop regions are likely to be caused by a lack of restraints, intrinsic mobility or crystal packing. The differences in the core of the protein are less likely to be artifacts of the structure determination, but then the differences should usually be reflected in the lack of agreement of the restraints with the crystal structure of the protein, as has been described above.

H^N–N RDCs were measured to evaluate how accurate the crystal and the solution structures were for residues 10–55. The average Cornilescu et al. (1998) Q factor for crystal structures was 0.17 ± 0.02. No increase in Q factor (i.e. relative change in backbone H^N–N orientations) was observed for structures bound to DNA, suggesting that DNA binding does not change relative orientation of the helices. Our determined NMR structure had $\langle Q \rangle = 0.22 \pm 0.05$, averaged over all members of the ensemble, suggesting that the H^N–N bond vector orientation is less accurate than observed in the crystal structure.

Backbone and side-chain dynamics

It has been suggested that intrinsic differences in the ensemble members of NMR (and ever more so) crystal

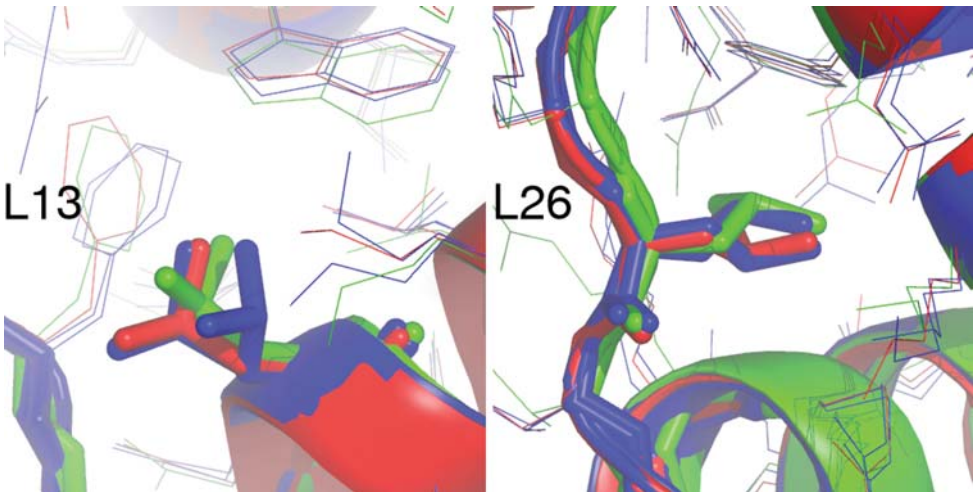

Fig. 3 Comparison of side-chain orientations between the solution and crystal structures. Lowest energy WT solution structure (green); WT crystal structure (1ENH) (red); K52A/E mutants crystal structures (1P7ID/1P7JA) (blue)

Table 2 NOE is regarded as violated if the distance in the crystal structure is >0.5 Å from that predicted by NOE, including correction for the proton multiplicity of the group (Fletcher et al. 1996)

Structure	1DU0B	1ENH	1P7IA	1P7IB
Violated NOEs	11	7	7	7
Structure	1P7ID	1P7JA	1P7JB	1P7JC
Violated NOEs	6	12	8	8
Structure	2HDDA	2HDDB	3HDDA	3HDDB
Violated NOEs	10	12	8	13

structures sample the underlying dynamics of the molecule under native conditions (Best et al. 2006). To probe this assumption, the H^N–N backbone dipolar relaxation values as well as ^2H quadrupolar side-chain relaxation parameters were measured and order parameters calculated. They were then compared to the order parameters predicted from the MD simulations (S^2_{MD}) (which usually reproduce well the backbone dynamics (Philippopoulos and Lim 1995; Philippopoulos et al. 1997)) as well as to the order parameters predicted from the crystal structures ($S^2_{timeless}$).

The fitted backbone S^2 values were high for secondary structure elements ($<S^2> = 0.84 \pm 0.05$) with values tailing off at the N- and C-termini of the protein (Fig. 4a). Mobility of the H^N–N bond vector could be described for most of the residues in the helices with the simplest model ("model 1"), which requires only S^2 as an adjustable parameter and assumes that the motions on the fast time are too fast ($t_e < 20$ ps) to be characterised. For residues in the loop region and the beginning of H2 the internal motion was slower ($t_e > 20$ ps) and these residues were fitted with model 2. Residues N21, R24 and Y25 had chemical exchange contribution to R_2, suggesting motion on a ms timescale (model 3).

The measurement of the side-chain dynamics was performed using the method described by (Millet et al. 2002; Muhandiram et al. 1995). All five independent deuterium

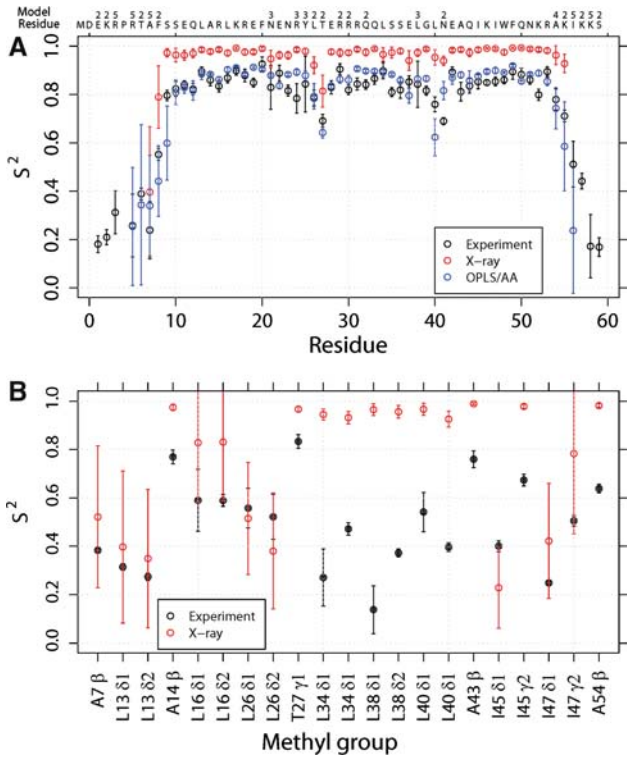

Fig. 4 Backbone (**a**) and side-chain (**b**) dynamics measured using NMR and predicted from CSE or MD simulations. The model used to fit the backbone relaxation data is shown on the top of the plot with no number reported meaning model 1

relaxation rates (D_z, D_+, $3D_Z^2 - 2$, $D_+D_z - D_zD_+$, D_+^2) were recorded at 11.7 T (500 MHz ^1H) and inequality and consistency relationships (Millet et al. 2002) were used to confirm that the data was correctly acquired. Within the measurement error the data satisfied both relationships. For the consistency relationships, the correlation coefficient between the single measured relaxation rates (left hand side of the equations in (Millet et al. 2002)) and the rates calculated from two measurements was at least 0.94. In

general, it was observed that $H^{\delta 1}$ of Leu side-chains were of lower intensity in our spectra, possibly due to some impurities during protein expression, leading to an increased fitting error for these methyl groups, apparent in the aforementioned relationships. Therefore, any subsequent per-residue analysis involved only $H^{\delta 2}$ Leu methyl group, for which data of better quality was obtained.

The spectral density values obtained from solving the relaxation equation were used to obtain side-chain order parameters (S^2_{axis}). Residues at the N-terminus of the protein displayed characteristic low S^2_{axis}, consistent with high mobility of this region as expected from backbone dynamics. L13, only partially buried, present in the first turn of helix 1 (H1) had low order parameters for H^{δ}. This residue is present in multiple conformations in the crystal structures and might undergo rotameric transitions. The short chain of the solvent-exposed A14 had order parameter just 0.07 lower than the corresponding value for the backbone amide group. Fully buried L16 had the highest order parameter (0.6) for all leucines. Interestingly, the partially buried L26, which is present in multiple conformations in the EnHD crystal structures and implicated in the network of conformational adjustments (Stollar et al. 2003) had, for a leucine, a relatively high order parameter of 0.5. The H^N–N bond vector of T27 was mobile ($S^2_{axis} = 0.7$), yet the methyl group of this residue remained relatively rigid ($S^2_{axis} = 0.83$). The OH group of this residue provides a hydrogen bond to N-cap helix 2 (H2) and no rotameric transition around C^{α}–C^{β} bond is possible with the oxygen bound to R30 H^N. The mobility of the Leu H^{δ} from H2 and the turn region similarly affected the deuterium relaxation, so that the calculated order parameters are $\sim$0.4. The mobility of the shorter-chain A43 H^{β} and fully buried I45 H^{γ} was significantly lower ($S^2_{axis} > 0.7$). The mobility of the fully buried I45 was lower than the solvent-exposed I47 or I56, which is present in the frayed region of H3.

The distribution of the S^2_{axis} order parameters for methyl side-chains follows the trend observed for other proteins (Best et al. 2004), with residues at the N- and C-termini displaying low order parameters and methyl groups on the longer side-chains having lower order parameters than ones on the shorter ones.

Molecular dynamics simulations

Our simulations were done at 298°K in all-atom optimised potentials for liquid simulations (OPLS/AA) (Kaminski et al. 2001) and AMBER99ϕ (Sorin and Pande 2005) force-fields using explicit water model in GROMACS 3.3 simulation package (Lindahl et al. 2001; Van der Spoel et al. 2005). The protein was stable under both force-fields after 10 ns of the MD simulation with RMSD to the starting structure for the residues 10–55 of 0.6 Å. and 0.5 Å for OPLS/AA and AMBER99ϕ, respectively. The pattern of backbone order parameters was similar in both force-fields with low values at the N- and C-termini of the protein and lower values in the loop regions (see Fig. 4a). The order parameters were on average 0.03 higher for the AMBER99ϕ, than for OPLS/AA and displayed a similar pattern and magnitude to the simulation of (Zhao et al. 2006) in the structured region of the protein. Their simulations, performed under similar conditions (Gromacs/AMBER94/SPC) displayed lower mobility of the N- and C-termini of the protein, with $S^2 > 0.8$ for residues 5–10. The order parameters predicted by the OPLS/AA force-field were closer to the experimentally determined values, with excellent agreement for H1. The force-field also correctly reproduced the decrease in order parameters for residues L26 and T27 and the decrease in the order parameters for the turn region. While the pattern of the OPLS-simulated values matched less closely the experiment in the helix-turn-helix motif, the force-field correctly predicted the most rigid residue in H3.

Since the OPLS/AA force-field more closely predicted the experimentally derived order parameters, we used it to compare to the measured values and models fitted by the Tensor 2 program in a greater detail, as shown on Fig. 5. We focused on residues L26–E28 for the loop region 1, which displayed variability in the S^2 values and yet they were correctly predicted by the force-field. Multiple correlation functions, calculated for each ns of the simulation, suggested that the L26 H^N–N bond vector might undergo motion on ns time-scale, due to significant differences in the correlation function for each ns. Similarly, for T27, the correlation function reports different values for each ns of the simulation, suggesting that rare transitions occur for this residue. On contrary, such transitions were not observed for the first residue in H2—E28—for which the correlation function looked similar for each ns of the simulation.

The agreement of the predicted order parameters for the methyl groups was significantly worse in both force-fields, with comparison to the experimental data (data not shown). Most of the side-chains undergo rotameric transitions in the solution and did so during the simulation (Best et al. 2004, 2005). During the 10 ns simulation the correlation functions for A14 and A43 H^{β}, L26 and L34 H^{δ}, T27 $H^{\gamma 2}$ converged (i.e. no rotameric transition were observed), leading to high ($S^2_{axis} > 0.85$) order parameters in both force-fields. Other methyl groups had large error in the precision of the determined parameters due to the rotameric transitions, which were not sampled fully in a 10 ns run. Longer simulations will be needed to obtain the convergence of the correlation function and the equilibrium distribution of the rotamers.

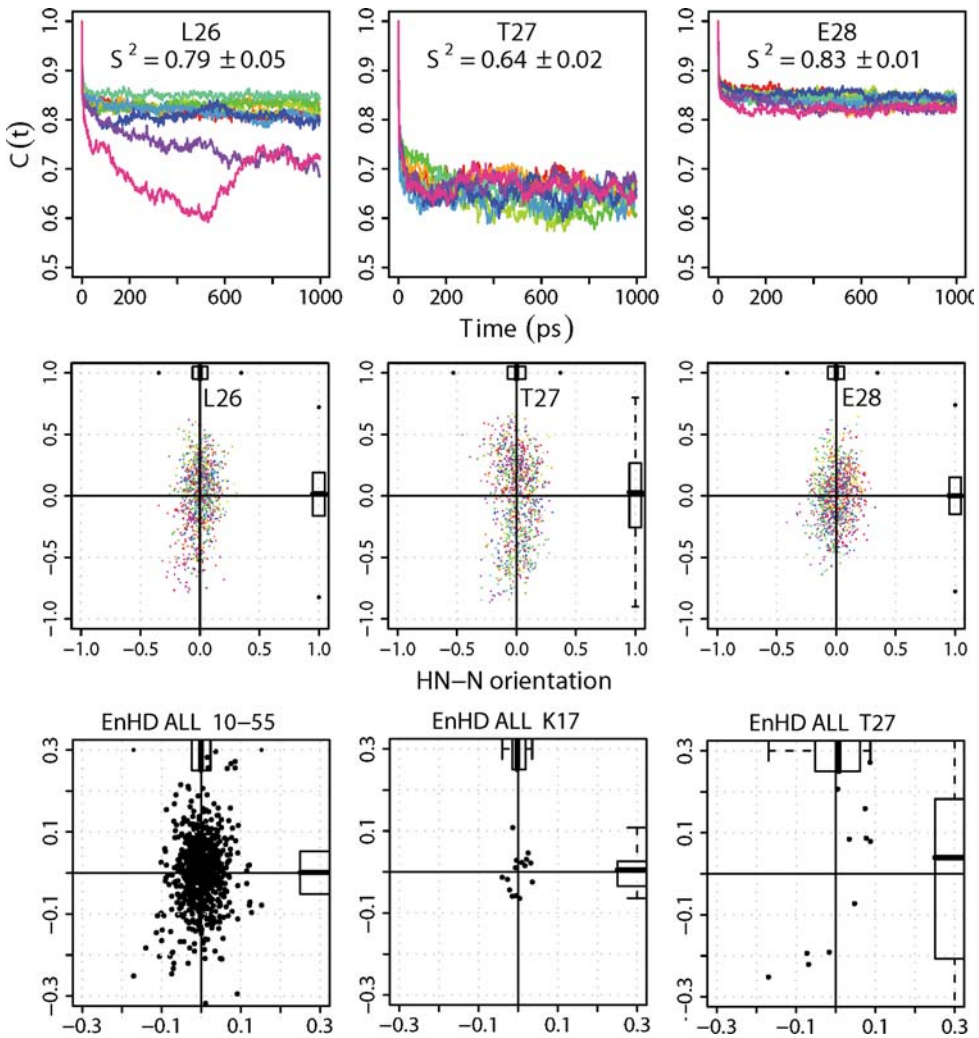

Fig. 5 Orientation of H^N–N bond vector in OPLS/AA simulation and in CSE. *Top panel:* time correlation function, calculated for each ns of the simulation. *Middle panel:* Projection of H^N–N bond vector during the simulation. The amide N atom is at the origin with the mean H^N–N bond vector pointing directly out of the page and the XY projection of the mean N–C^α vector pointing to the right. The box-and-whisker plots of the x- and y- coordinate, showing median and the inter-quartile range are drawn on the side. Rare transitions were observed for L26, since the correlation functions do not converge for each ns. T27 was fitted with model 2 experimentally, and here it undergoes motions with longer correlation times. Correlation function for E28 quickly plateaued at 0.83 and is similar for each ns of the simulation suggesting no rare transitions, in agreement with the experiments. *Bottom panel:* Projection of H^N–N bond vector in the CSE for all superimposed residues (left), for the residue with the highest S^2_{timeless} (K17) and the lowest (T27). The scale in the bottom panel plots is over three times lower than in the middle panel one, to conform to the lower spread of the vectors observed in the crystal structures

Comparison of dynamics with crystal structures

The suggestion that the disorder shown in the crystal structure ensembles (CSE)—a collection of crystal structures of the given protein or its single point mutants—provide a representative sample of the structural fluctuations of a protein under native conditions is based predominantly on a comparison of backbone and side-chain mobility with the disorder present between multiple crystal structures (Best et al. 2006).

The RMSD between 15 fitted structures for residues 10–55 is 0.36 Å—slightly lower than that obtained during the MD simulations. The backbone and side-chain order parameters were obtained using this CSE. In such an ensemble there is no "time" concept, which is important for the S^2 parameter determination. We, therefore, evaluated order parameters by calculating the second order Legendre polynomial between all the available pairs of H^N–N bond vectors for the given residue. This is equivalent to calculating the spectral density function at random time points:

$$S^2_{timeless} = \langle P_2(\hat{\mu}_n \cdot \hat{\mu}_m) \rangle = \sum_{n=1}^{allstruct.} \sum_{m \neq n}^{allstruct.} (3(\hat{\mu}_n \cdot \hat{\mu}_m)^2 - 1)/(2n(n-1))$$

To evaluate how well this "time-less" definition can predict order parameters, we draw at random 15 structures from the OPLS MD simulation and calculated S^2 values for each of the residues. The Pearson's correlation coefficient between the S^2 values calculated using traditional method (from Figs. 4, 5a) and this "time-less" approximation was 0.96, with RMSD between two values of 0.07 and no systematic error on the predicted values. This suggests that while using only 15 crystal structures introduces a small random error, no systematic one should be present if the observed distribution of H^N–N bond vectors in CSE is similar to that in the solution. The same structure (m = n) case was not used in $S^2_{timeless}$ comparison to decrease the bias of $S^2 = 1$ resulting from this case.

The order parameters predicted in such a way from the crystal structure ensemble were significantly higher than obtained from the experiment or MD simulations (Fig. 4a). In the helices the predicted values were almost uniformly high (average of 0.98), suggesting highly similar orientation of H^N–N bond vector for a given residue between the crystal structures, as observed on Fig. 5. The pattern of S^2 values did not follow well the one experimentally obtained. Such an uniform orientation of the H^N–N bond vector between the crystal structures is consistent with excellent agreement of the measured RDCs with the RDCs predicted from high resolution crystal structures (Bax 2003). The predicted S^2 values in the un-aligned N-terminus (A7 and F8) of the protein were significantly lower, consistent with the experiment (as well as with the lack of superposition). Residues R24 and Y25, present in the loop region, had high order parameters, yet they were mobile according to the NMR relaxation data. On the other hand, the predicted order parameters for residues L26 and T27 were significantly lower than the average. These residues displayed lower order parameters from NMR data. It is worth pointing out that the backbone for these residues is solvent exposed and interacts with other molecules in the crystal lattice.

Due to the uniformly high $S^2_{timeless}$ value calculated from CSE, the RMSD between the predicted and the determined datasets for EnHD was 0.14—similar to the value obtained for other proteins (Best et al. 2006), suggesting strong systematic deviation in $S^2_{timeless}$. The above calculations show that the disorder present in the crystal structure ensemble is significantly lower than mobility probed by the NMR relaxation measurements or observed in MD simulations at room temperature (i.e. native temperature for this protein) making it unlikely that they capture a *representative* subset of structures of the true native state ensemble. This, however, does not rule out that the disorder in the crystal structures samples a subset of the native mobility, which is severely restricted due to crystal packing. Testing this claim using disorder in H^N–N bond vectors might prove to be difficult due to the intrinsically small differences observed, which might also be influenced by structure determination protocols.

A larger structural change observed in the CSE ensemble is related to different rotameric states of the side-chains. Unless the accuracy of the crystal structure is low, the reported rotameric state is unlikely to be an artefact of the calculation protocol (as it is more likely to be the case in the H^N–N bond vector distribution, described above). The best method to determine the equilibrium distribution of the rotameric state of the side-chain would be the analysis of the $^3J_{CC}$ and $^3J_{CN}$ couplings, combined with side-chain dipolar couplings (Chou et al. 2003). Alternatively, it was shown that the methyl group side-chain order parameters (S^2_{axis}), measured using deuterium relaxation, are sensitive to the relative distribution of the rotamers (Best et al. 2005; Chou et al. 2003).

EnHD displayed variation in the rotameric distribution of the side-chains with methyl groups between different crystal structures. L26, in particular, was implicated in the network of conformational adjustments, where it occupied one of two side-chain rotameric conformations (*-gauche* and *trans*) and different positions in the hydrophobic core in the structures of K52 mutants (Stollar et al. 2003). Furthermore, it was also observed that the side-chains of the partially buried L13 and fully buried I45 occupied different conformations in the crystal structures. Prediction of the "time-less" order parameters ($S^2_{timeless}$) from the CSE yielded low order parameters for these residues, consistent with the experimental observations (see Fig. 4b). Prediction of $S^2_{timeless}$ for most of the other residues yielded values close to 1, as it was observed previously for the backbone predictions. This is inconsistent with the determined order parameters for L34, L38 and L40, which were $\sim$0.4, showing that the "time-less" order parameter determined from CSE ensemble does not predict the dynamics undergoing in the solution reliably. A similar conclusion was reached by Higman et al. (2004), who analysed the distribution of asparagine and glutamine side-chain conformations in solution and crystal for lysozyme using RDCs.

The RMSD between the predicted and determined values is 0.36 with correlation of 0.42, which is similar to the values obtained from other proteins (Best et al. 2006). It was noted that the agreement might be improved significantly by increasing the ensemble size, with over 330 crystal structures giving RMSD of 0.17 between the predicted and the determined values. When evaluating these numbers, it is worth noting that a uniform, randomly distributed variable in

a range of 0.0–1.0 would give, on average, RMSD of $1/(2\sqrt{3}) = 0.29$ to a fixed value of 0.5 (Riley et al. 2000). If we assume that the order parameters are uniformly distributed in a range of 0.2–0.8, then the RMSD to a mean value of 0.5 would be $(0.8 - 0.2)/2\sqrt{3} = 0.17$. Obviously, the distribution of order parameters resembles more Gaussian or Poisson-like curves with slight differences occurring for different amino acids (Best et al. 2004), for which the RMSD to an average value would be even lower than the one quoted here for a uniform distribution.

Conclusions

In this study we compared directly acquired NMR parameters to multiple crystal structures of EnHD to account for possible inaccuracies in the data analysis, influence of crystal contacts and dynamics of the protein. We found that the TALOS-predicted backbone dihedral angles reproduce very well the dihedral angles found in the multiple crystal structures with the TALOS-standard deviation being similar to the standard deviation of the ϕ/ψ space found between the crystal structures of the same protein. Reproducibility of the $^{3}J(H^{N}H^{\alpha})$ couplings was similar to that found during parameterisation of the Karplus equations, but improving on the agreement (i.e. obtaining a lower RMSD) between the solution measured and crystal structure predicted couplings might require using multiple crystal structures. While no crystal structure satisfied all of the solution NOEs, only a small fraction of them was violated. The violations usually occurred for the surface residues or for the residues for which multiple conformations were observed in the crystal lattice. This suggests that the NMR and crystallography data are consistent with each other, but multiple structures might be needed to satisfy all datasets (Clore and Schwieters 2006; Schwieters and Clore 2007; Zhang et al. 2007).

Comparison of the NMR relaxation data to multiple crystal structures allowed us to investigate if the disorder seen in crystal structures can be accounted by intrinsic dynamics of EnHD. The backbone order parameters estimated from the CSE of EnHD are too high to represent the dynamics of the molecule under native conditions. The RMSD between the predicted and determined side-chain order parameters (which are, to some extent, sensitive to the rotameric state) in EnHD is close to a value expected from a random distribution. Overall, the analysis of the backbone and side-chain dynamics data does not support the hypothesis that the CSE capture the *representative* subset of structures of the true native state ensemble for EnHD. This, however, does not rule out the possibility that the disorder in the crystal structures sub-samples, to some extent, the mobility, which is severely restricted due to crystal packing with the crystal forces likely to have more influence on the surface residues.

Acknowledgments T.L.R. was supported by the Medical Research Council and Trinity College. Alan R. Fersht and Stefan M. V. Freund are thanked for their support and encouragement as well as their reading the manuscript and helpful comments.

References

Bax A (2003) Weak alignment offers new NMR opportunities to study protein structure and dynamics. Protein Sci 12:1–16

Bax A, Ikura M, Kay LE, Barbato G, Spera S (1991) Multidimensional triple resonance NMR spectroscopy of isotopically uniformly enriched proteins: a powerful new strategy for structure determination. Ciba Found Symp 161, 108–119; discussion 119–135

Berendsen HJC, Postma JPM, Vangunsteren WF, Dinola A, Haak JR (1984) Molecular-dynamics with coupling to an external bath. J Chem Phys 81:3684–3690

Best RB, Clarke J, Karplus M (2004) The origin of protein sidechain order parameter distributions. J Am Chem Soc 126:7734–7735

Best RB, Clarke J, Karplus M (2005) What contributions to protein side-chain dynamics are probed by NMR experiments? A molecular dynamics simulation analysis. J Mol Biol 349:185–203

Best RB, Lindorff-Larsen K, DePristo MA, Vendruscolo M (2006) Relation between native ensembles and experimental structures of proteins. Proc Natl Acad Sci USA 103:10901–10906

Billeter M (1992) Comparison of protein structures determined by NMR in solution and by X-ray diffraction in single crystals. Q Rev Biophys 25:325–377

Bonvin AM, Brunger AT (1996) Do NOE distances contain enough information to assess the relative populations of multi-conformer structures? J Biomol NMR 7:72–76

Brunger AT (1997) X-ray crystallography and NMR reveal complementary views of structure and dynamics. Nat Struct Biol 4 (Suppl):862–865

Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS et al (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 54:905–921

Chou JJ, Gaemers S, Howder B, Louis JM, Bax A (2001) A simple apparatus for generating stretched polyacrylamide gels, yielding uniform alignment of proteins and detergent micelles. J Biomol NMR 21:377–382

Chou JJ, Case DA, Bax A (2003) Insights into the mobility of methyl-bearing side chains in proteins from (3)J(CC) and (3)J(CN) couplings. J Am Chem Soc 125:8959–8966

Clarke ND, Kissinger CR, Desjarlais J, Gilliland GL, Pabo CO (1994) Structural studies of the engrailed homeodomain. Protein Sci 3:1779–1787

Clore GM, Schwieters CD (2006) Concordance of residual dipolar couplings, backbone order paramters and crystallographic B-factors for a small alpha/beta protein: a unified picture of high probability, fast atomic motions in proteins. J Mol Biol 355:879–886

Cordier F, Grzesiek S (1999) Direct observation of hydrogen bonds in proteins by interresidue (3h)J(NC′) scalar couplings. J Am Chem Soc 121:1601–1602

Cordier F, Dingley AJ, Grzesiek S (1999) A doublet-separated sensitivity-enhanced HSQC for the determination of scalar and dipolar one-bond J-couplings. J Biomol NMR 13:175–180

Cornilescu G, Marquardt JL, Ottiger M, Bax A (1998) Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. J Am Chem Soc 120:6836–6837

Cornilescu G, Delaglio F, Bax A (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. J Biomol NMR 13:289–302

DePristo MA, de Bakker PI, Blundell TL (2004) Heterogeneity and inaccuracy in protein structures solved by X-ray crystallography. Structure 12:831–838

Dosset P, Hus JC, Blackledge M, Marion D (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. J Biomol NMR 16:23–28

Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. J Chem Phys 103:8577–8593

Farrow NA, Zhang OW, Forman-Kay JD, Kay LE (1995) Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer. Biochemistry 34:868–878

Fletcher CM, Jones DNM, Diamond R, Neuhaus D (1996) Treatment of NOE constraints involving equivalent or nonstereoassigned protons in calculations of biomacromolecular structures. J Biomol NMR 8:292–310

Fraenkel E, Rould MA, Chambers KA, Pabo CO (1998) Engrailed homeodomain-DNA complex at 2.2 A resolution: a detailed view of the interface and comparison with other engrailed structures. J Mol Biol 284:351–361

Goddard TD, Kneller DG SPARKY 3. University of California, San Francisco

Grant RA, Rould MA, Klemm JD, Pabo CO (2000) Exploring the role of glutamine 50 in the homeodomain-DNA interface: crystal structure of engrailed (Gln50 → Ala) complex at 2.0 A. Biochemistry 39:8187–8192

Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997). LINCS: a linear constraint solver for molecular simulations. J Computat Chem 18:1463–1472

Higman VA, Boyd J, Smith LJ, Redfield C (2004) Asparagine and glutamine side-chain conformation in solution and crystal: a comparison for hen egg-white lysozyme using residual dipolar couplings. J Biomol NMR 30:327–346

Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. J Phys Chem B 105:6474–6487

Karplus M (1959) Contact electron-spin coupling of nuclear magnetic moments. J Chem Phys 30:11–15

Kissinger CR, Liu BS, Martin-Blanco E, Kornberg TB, Pabo CO (1990) Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: a framework for understanding homeodomain-DNA interactions. Cell 63:579–590

Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR 8:477–486

Lindahl E, Hess B, van der Spoel D (2001) GROMACS 3.0: a package for molecular simulation and trajectory analysis. J Mol Model 7:306–317

Lipari G, Szabo A (1982a) Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. 1. Theory and range of validity. J Am Chem Soc 104:4546–4559

Lipari G, Szabo A (1982b) Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. 2. Analysis of experimental results. J Am Chem Soc 104:4559–4570

Liu YJ, Zhao DQ, Altman R, Jardetzky O (1992) A systematic comparison of 3 structure determination methods from NMR data—dependence upon quality and quantity of data. J Biomol NMR 2:373–388

Mandel AM, Akke M, Palmer AG (1995) Backbone dynamics of *Escherichia coli* ribonuclease HI—correlations with structure and function in an active enzyme. J Mol Biol 246:144–163

Mayor U, Johnson CM, Daggett V, Fersht AR (2000) Protein folding and unfolding in microseconds to nanoseconds by experiment and simulation. Proc Natl Acad Sci USA 97:13518–13522

Mayor U, Guydosh NR, Johnson CM, Grossmann JG, Sato S, Jas GS, Freund SM, Alonso DO, Daggett V, Fersht AR (2003) The complete folding pathway of a protein from nanoseconds to microseconds. Nature 421:863–867

Millet O, Muhandiram DR, Skrynnikov NR, Kay LE (2002) Deuterium spin probes of side-chain dynamics in proteins. 1. Measurement of five relaxation rates per deuteron in (13)C-labeled and fractionally (2)H-enriched proteins in solution. J Am Chem Soc 124:6439–6448

Miyamoto S, Kollman PA (1992) Settle—an analytical version of the shake and rattle algorithm for rigid water models. J Comput Chem 13:952–962

Muhandiram DR, Yamazaki T, Sykes BD, Kay LE (1995) Measurement of H-2 T-1 and T-1p relaxation-times in uniformly C-13-labeled and fractionally H-2-labeled proteins in solution. J Am Chem Soc 117:11536–11544

Neidhard FC, Bloch PL, Smith DF (1974) Culture medium for enterobacteria. J Bacteriol 119:736–747

Neri D, Szyperski T, Otting G, Senn H, Wuthrich K (1989) Stereospecific nuclear magnetic-resonance assignments of the methyl-groups of valine and leucine in the DNA-binding domain of the 434-repressor by biosynthetically directed fractional C-13 labeling. Biochemistry 28:7510–7516

Neuhaus D, Williamson MP (2000) The nuclear overhauser effect in structural and conformational analysis Wiley-VCH

Pardi A, Wagner G, Wuthrich K (1983) Protein Conformation and proton NMR chemical-shifts. Eur J Biochem 137:445–454

Philippopoulos M, Lim C (1995) Molecular dynamics simulation of *E. coli* ribonuclease H1 in solution: correlation with NMR and X-ray data and insights into biological function. J Mol Biol 254:771–792

Philippopoulos M, Mandel AM, Palmer AG 3rd, Lim C (1997) Accuracy and precision of NMR relaxation experiments and MD simulations for characterizing protein dynamics. Proteins 28:481–493

Piotto M, Saudek V, Sklenar V (1992) Gradient-tailored excitation for single-quantum NMR-spectroscopy of aqueous-solutions. J Biomol NMR 2:661–665

Religa TL, Markson JS, Mayor U, Freund SM, Fersht AR (2005) Solution structure of a protein denatured state and folding intermediate. Nature 437:1053–1056

Renner C, Schleicher M, Moroder L, Holak TA (2002) Practical aspects of the 2D N-15–{H-1}-NOE experiment. J Biomol NMR 23:23–33

Riley KF, Hobson MP, Bence SJ (2000) Mathematical methods for physics and engineering. A comprehensive guide. Cambridge University Press, Cambridge

Santiveri CM, Perez-Canadillas JM, Vadivelu MK, Allen MD, Rutherford TJ, Watkins NA, Bycroft M (2004) NMR structure of the alpha-hemoglobin stabilizing protein: insights into conformational heterogeneity and binding. J Biol Chem 279:34963–34970

Schwieters CD, Clore GM (2007) A physical picture of atomic motions within the Dickerson DNA dodecamer in solution derived from joint ensemble refinement against NMR and large-angle X-ray scattering data. Biochemistry 46:1152–1166

Skrynnikov NR, Millet O, Kay LE (2002) Deuterium spin probes of side-chain dynamics in proteins. 2. Spectral density mapping and identification of nanosecond time-scale side-chain motions. J Am Chem Soc 124:6449–6460

Smith JL, Hendrickson WA, Honzatko RB, Sheriff S (1984) Discrete disorder in protein crystals. Acta Crystallogr A 40:C51–C51

Smith JL, Hendrickson WA, Honzatko RB, Sheriff S (1986) Structural heterogeneity in protein crystals. Biochemistry 25:5018–5027

Smith LJ, Bolin KA, Schwalbe H, MacArthur MW, Thornton JM, Dobson CM (1996) Analysis of main chain torsion angles in proteins: prediction of NMR coupling constants for native and random coil conformations. J Mol Biol 255:494–506

Snyder DA, Bhattacharya A, Huang YPJ, Montelione GT (2005) Assessing precision and accuracy of protein structures derived from NMR data. Proteins 59:655–661

Sorin EJ, Pande VS (2005) Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. Biophys J 88: 2472–2493

Spronk CA, Nabuurs SB, Bonvin AM, Krieger E, Vuister GW, Vriend G (2003) The precision of NMR structure ensembles revisited. J Biomol NMR 25:225–234

Stollar EJ, Mayor U, Lovell SC, Federici L, Freund SM, Fersht AR, Luisi BF (2003) Crystal structures of engrailed homeodomain mutants: implications for stability and dynamics. J Biol Chem 278:43699–43708

Tjandra N, Bax A (1997) Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. Science 278:1111–1114

Tucker-Kellogg L, Rould MA, Chambers KA, Ades SE, Sauer RT, Pabo CO (1997) Engrailed (Gln50 → Lys) homeodomain-DNA complex at 1.9 A resolution: structural basis for enhanced affinity and altered specificity. Structure 5:1047–1054

Van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) Gromacs: fast, flexible, and free. J Comput Chem 26:1701–1718

Vuister GW, Bax A (1993) Quantitative J correlation—a new approach for measuring homonuclear 3-bond J(H(N)H(Alpha) coupling-constants in N-15-enriched proteins. J Am Chem Soc 115:7772–7777

Wang AC, Bax A (1996) Determination of the backbone dihedral angles phi in human ubiquitin from reparametrized empirical Karplus equations. J Am Chem Soc 118:2483–2494

Wagner G, Hyberts SG, Havel TF (1992) NMR structure determination in solution: a critique and comparison with X-ray crystallography. Annu Rev Biophys Biomol Struct 21:167–198

Wuthrich K (1986) NMR of proteins and nucleic acids. Wiley

Zhang XJ, Wozniak JA, Matthews BW (1995) Protein flexibility and adaptability seen in 25 crystal forms of T4 lysozyme. J Mol Biol 250:527–552

Zhang Q, Stelzer AC, Fisher CK, Al-Hashimi HM (2007) Visualizing spatially correlated dynamics that directs RNA conformational transitions. Nature 450:1263–1267

Zhao X, Huang XR, Sun CC (2006) Molecular dynamics analysis of the engrailed homeodomain-DNA recognition. J Struct Biol 155:426–437

Zweckstetter M, Bax A (2000) Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. J Am Chem Soc 122: 3791–3792