Structure and Backbone Dynamics of Apo-CBFβ in Solution[†,‡]

Magnus Wolf-Watz,[§] Thomas Grundström,[||] and Torleif Härd*[,§]

Department of Biotechnology, Center for Structural Biochemistry, Royal Institute of Technology (KTH), Novum, S-141 57 Huddinge, Sweden, and Department of Cell and Molecular Biology, Umeå University, S-901 87 Umeå, Sweden

Received April 6, 2001; Revised Manuscript Received June 27, 2001

ABSTRACT: Runx proteins constitute a family of mammalian transcription factors that interact with DNA through their evolutionarily conserved Runt domain. CBFβ, alternatively denoted PEBP2β, is the non-DNA-binding heterodimer partner and acts to enhance the DNA binding affinity of Runx proteins. Runx proteins and CBFβ are associated with a variety of biological functions and human diseases; they are, for example, together the most frequent targets for chromosomal rearrangements in acute human leukemias. We have determined the solution structure and characterized the backbone dynamics of C-terminally truncated fragments containing residues 1−141 of CBFβ. The present apo-CBFβ structure is very similar to that seen in a Runt−CBFβ complex. An evaluation of backbone ^{15}N NMR relaxation parameters shows that CBFβ is a rigid molecule with high order parameters throughout the backbone; the only regions displaying significant dynamics are a long loop and the C-terminal α-helix. A few residues display relaxation behavior indicating conformational exchange on microsecond to millisecond time scales, but only one of these is located at the Runt binding surface. Our structure and dynamics analysis of CBFβ therefore suggests that the protein binds to Runt without large conformational changes or induced folding ("lock-and-key" interaction). The apo-CBFβ structure presented here exhibits several significant differences with two other published NMR ensembles of very similar protein fragments. The differences are located in four regions outside of the central β-barrel, whereas the β-barrel itself is almost identical in the three NMR structures. The comparison illustrates that independently determined NMR structures may display rather large differences in backbone conformation in regions that appear to be well-defined in each of the calculated NMR ensembles.

CBFβ, also called PEBP2β, is the non-DNA-binding heterodimer partner for the Runx family of mammalian transcription factors. All three known Runx proteins share the evolutionarily conserved Runt homology domain (RD)[1] named after the *Drosophila melanogaster* gene *Runt* (*1*). The RD is sufficient for both DNA binding and heterodimerization with CBFβ. The ternary CBFβ−RD−DNA complex is formed at a 1:1:1 ratio (*2*). *Runx1,* also denoted *AML1, PEBP2αB,* or *CBFα2,* is essential for blood cell formation, since mice with homozygous disruption of this gene totally lack definitive hematopoietic stem cells and consequently lack all blood cells (*3*, *4*). Hematopoietic stem cells are also important in promoting angiogenesis (*5*). *CBFβ* deficient mice have the same defects as *Runx1* deficient mice, showing that CBFβ is required for the function of the Runx1 protein also in vivo (*6*). The dominantly inherited human disease familial platelet disorder with propensity to acute myeloid leukemia (FPD/AML) is caused by defects in one of the two copies of the *Runx1* gene (*7*). Point mutations found in FPD/AML and in several percent of acute leukemias cluster within the RD. Together, the Runx1/AML1 and CBFβ genes are the most frequent targets for chromosomal rearrangements in acute human leukemias (*6*, *8*−*10*).

Runx2 is essential for differentiation of osteoblasts and formation of bone. The dominantly inherited human disease cleidocranial dysplasia, characterized by multiple skeletal abnormalities, is caused by defects in one of the two copies of the *Runx2* gene (*9*, *11*).

To date, only one mammalian *CBFβ* gene has been identified (*12*, *13*), whereas two *Drosophila CBFβ* genes, *Brother* and *Big-brother,* have been found (*14*). CBFβ is ubiquitously expressed, localized to the cytoplasm, and transported to the nucleus after heterodimerization with Runx proteins (*15*). The 135 N-terminal amino acid residues of CBFβ are sufficient for heterodimerization with the RD (*16*). CBFβ binding to Runt proteins induces a several-fold increase in DNA binding affinity by decreasing the rates of dissociation from DNA (*12*). The mechanism is allosteric and involves the stabilization of one minor and one major groove DNA binding loop in the RD (*17*). As anticipated from biochemical data (*12*), X-ray diffraction studies have

[†] This work was supported by grants to T.H. from the Swedish Foundation for Strategic Research (SSF), the Swedish Research Council, and the Knut and Alice Wallenberg Foundation, and by a grant to T.G. from the Swedish Cancer Society.

[‡] The atomic coordinates of the CBFβ ensemble are deposited at the RCSB Protein Data Bank as entry 1ILF.

* To whom correspondence should be addressed. Fax: +46-8-608 92 90. E-mail: Torleif.Hard@biochem.kth.se.

[§] Royal Institute of Technology.

[||] Umeå University.

[1] Abbreviations: CBFβ_{141}, amino acid residues 1−141 of core binding factor β; FPD/AML, familial platelet disorder with propensity to acute myeloid leukemia; RD, Runt homology domain; rmsd, root-mean-square deviation; NMR, nuclear magnetic resonance; HSQC, heteronuclear single-quantum coherence; NOE, nuclear Overhauser effect; TOCSY, total correlation spectroscopy; SA, simulated annealing.

10.1021/bi010713+ CCC: $20.00
Published on Web 08/30/2001

shown that CBFβ does not make any DNA contacts (*17*, *18*).

Here we describe the solution structure and a characterization of backbone ^{15}N dynamics of apo-CBFβ. The work is part of our continuing studies of the functional and structural biology of the AML1/Runx-1 transcription factor (*19*−*21*). In particular, the studies presented here were carried out to investigate how and to what extent CBFβ undergoes conformational changes when binding to the RD. CBFβ consists of a partially open six-stranded up-and-down β-barrel capped by α-helices. The question of conformational change has come up since two previously determined NMR structures of CBFβ (*22*, *23*) display very different backbone conformations in several regions outside the central β-barrel. The conformations of these regions are also in both cases different from those seen in the RD-complexed CBFβ (*24*). A reexamination of the structure of apo-CBFβ is therefore needed to resolve the issue of conformational change and understand how CBFβ functions at a molecular level. Complementary studies of backbone dynamics were carried out to investigate if inherent flexibility in apo-CBFβ might be implicated in the RD−CBFβ recognition mechanism. The structure that we determine here differs significantly from the two previous NMR structures of CBFβ, but it is on the other hand very similar to that observed in the RD−CBFβ complex. We also find that the backbone of apo-CBFβ is quite rigid with high order parameters for sub-nanosecond dynamics and only limited contributions from chemical exchange, with the exception of a long surface loop and the C-terminal α-helix. Our structural and dynamic data therefore suggest that RD binding proceeds without backbone rearrangements or stabilization of flexible regions in CBFβ.

EXPERIMENTAL PROCEDURES

Cloning, Purification, and Sample Preparation. Two different plasmid constructs were used for the overproduction of amino acid residues 1−141 of CBFβ (CBFβ_{141}). For the structure determination, we subcloned DNA corresponding to CBFβ_{141} into the pET-15b plasmid (Novagen; this plasmid encodes an N-terminal 6*-polyhistidine tag plus 14 other non-native N-terminal amino acid residues). Expression and purification were performed as described previously (*21*). For backbone dynamics measurements, we instead used the plasmid pET-11c (Novagen) that produces native proteins. CBFβ_{141} produced with pET-11c was purified by using DEAE-Sepharose CL-6B, Phenyl Sepharose CL-4B, and Superdex 75 gel filtration columns (all columns were purchased from Pharmacia Biotech). The purified proteins were finally resuspended in the appropriate NMR buffers, desalted on PD-10 columns (Pharmacia Biotech), and concentrated in Centricon tubes (Amicon). Unlabeled, ^{15}N-labeled, and ^{15}N- and ^{13}C-labeled CBFβ_{141} samples were produced. NMR samples were concentrated to 0.8−1.6 mM in 20−150 mM NaCl, 0 or 0.1 mM EDTA, 25 mM NaP_i (pH 6.5), 5 mM DTT, and 0.1% NaN_3.

NMR Spectroscopy. NMR data were acquired at 600 and 800 MHz on Varian Inova spectrometers and at 600 MHz on a Bruker AVANCE spectrometer. The spectrometers were equipped with three rf channels, pulsed field gradients, and 5 mm triple-resonance (^{1}H, ^{15}N, and ^{13}C) probes. NMR spectra used for the structure determination were recorded at 24 °C, and descriptions of these experiments can be found in ref *25*. Backbone resonance assignments of CBFβ_{141} were obtained on the basis of complementary HNCA/HN(CO)-CA and HNCACB/CBCA(CO)NH experiments. Aromatic side chain assignments were obtained from a DQF-COSY spectrum recorded in D_2O. The remaining side chain and all carbonyl carbon assignments were obtained from ref *26*. Side chain assignments were verified with a three-dimensional (3D) HCCH-TOCSY experiment. NOE-derived distance restraints were obtained from a 120 ms ^{15}N-edited 3D NOESY-HSQC spectrum acquired in H_2O, a 100 ms ^{13}C-edited 3D NOESY-HSQC spectrum acquired in D_2O, and a 90 ms two-dimensional (2D) NOESY spectrum acquired in D_2O. A 3D HNHA spectrum was acquired to obtain $^3J_{HN\alpha}$ scalar coupling constants. NMR data were processed using the Azara suit of programs (W. Boucher, University of Cambridge, Cambridge, U.K.), XWINNMR (Bruker), VNMR 5.3 (Varian), and NMRPipe (*27*) and subsequently analyzed using Ansig for Windows (*28*).

Measurements of ^{15}N R_1 and R_2 relaxation rates and steady-state heteronuclear ^{1}H−^{15}N NOEs were taken at 24 °C using the experiments described in ref *29*. The R_1 experiments were acquired with relaxation delays of 10, 100, 200, 300, 400, 500, 600, and 700 ms with a duplicate data set acquired at 200 ms to estimate experimental uncertainties. The CPMG delay in the R_2 measurement was 900 μs, and data were acquired with relaxation delays of 15.6, 31.2, 46.8, 62.5, 78.1, 93.7, 109.3, and 156.2 ms, with a duplicate data set acquired at 78.1 ms. All relaxation data points were obtained by integration of cross-peaks in two-dimensional spectra using procedures available within XWINNMR. Experimental errors in R_1, R_2, and steady-state ^{1}H−^{15}N NOEs were estimated as described in ref *30*. R_1 and R_2 relaxation rates were calculated by nonlinear fitting of relaxation decays to the equation $A \times \exp(-R_i t)$ using Mathematica (Wolfram Research).

NOE and Hydrogen Bond Distance Restraints. NOE cross-peaks were assigned and integrated within Ansig for Windows. Calibration of cross-peak intensities was carried out against known distances in antiparallel β-sheet structures (*31*). The NOE distance restraints were divided into three classes: strong, medium, and weak corresponding to 1.8−2.7, 1.8−3.6, and 1.8−6.0 Å, respectively. To account for the increased apparent intensities of methyl resonances, 0.5 Å was added to the upper distance limits for all cross-peaks involving methyl protons. Redundant intraresidue NOE distances were removed with Ansig for Windows and AQUA (*32*). Non-stereoassigned diastereotopic groups were treated as suggested in ref *33*. Multiplicity corrections were achieved by using r^{-6} summation in the structure calculations.

Slowly exchanging amide protons were identified by acquiring a ^{15}N HSQC spectrum on a lyophilized CBFβ_{141} sample dissolved in 100% D_2O. Backbone amide resonances visible after 20 min were identified as hydrogen bond donors. Main chain hydrogen bond acceptors were identified from stretches of regular secondary structure. Two distance restraints were incorporated for each hydrogen bond: the donor hydrogen to acceptor oxygen distance was set to 1.70−2.20 Å and the donor nitrogen to acceptor oxygen distance to 2.70−3.20 Å.

Dihedral Angles and Stereospecific Assignments. Restraints for backbone ϕ dihedral angles were derived from the 3D HNHA experiment. Calculated $^3J_{HN\alpha}$ values were

multiplied by a factor of 1.15 to account for the different relaxation properties of the diagonal and cross-peak signals (*34*). Dihedral ϕ angle restraints were divided into three groups depending on the magnitude of the $^3J_{HNH\alpha}$ values. For $^3J_{HNH\alpha}$ scalar couplings larger than 9.0 Hz, larger than 8.0 Hz, or in the range of 3.0−5.0 Hz, ϕ values were restrained to $-120 \pm 20°$, $-120 \pm 30°$, or $-60 \pm 30°$, respectively. Coupling constants in the range of 5.0−8.0 Hz were assumed to be averaged, and no ϕ values were assigned to the corresponding residues. Additional ϕ and ψ torsion angle restraints were obtained with the program TALOS (*35*). These angles were restrained to within 20° of the predicted value for those residues that were assigned good predictions by the program. Stereospecific assignments of β-methylene resonances and χ^1 torsion angles were determined from the relative intensities of intraresidue H^N-H^β NOEs in a 50 ms ^{15}N-edited 3D NOESY-HSQC spectrum in conjunction with the relative size of $^3J_{\alpha\beta}$ coupling constants estimated from a 30 ms ^{15}N-edited 3D TOCSY-HSQC spectrum (*36*). Stereospecific assignments of valine and leucine methyl groups were obtained by acquiring a constant-time (27 ms) ^{13}C HSQC spectrum on a 10% ^{13}C-labeled CBFβ_{141} sample dissolved in 100% D_2O (*37*). All methyl groups except those of Val106 could be stereospecifically assigned in this way.

Structure Calculations. CBFβ_{141} structures were calculated with the program CNS 1.0 (*38*) using torsion angle dynamics simulated annealing (SA). Force constants were set to their default values given in the original anneal.inp protocol, except for the force constant for dihedral angle restraints which was set to 1 kcal mol^{-1} rad^{-2}. Other SA parameters were set to default anneal.inp values, but the molecules were initially heated to 70 000 K (instead of 50 000 K) with a longer first cooling period of 3700 steps (yielding 55 ps of molecular dynamics during the first cooling period). A second Cartesian cooling period was also included. The 20 non-native N-terminal amino acid residues, including the polyhistidine tag, were omitted in the structure calculations. A total of 58 SA structures were calculated, and the 25 structures with total CNS energies of <390 kcal/mol were selected to represent the solution structure ensemble of CBFβ_{141}. These structures were analyzed using the programs PROCHECK-NMR (*32*) and MOLMOL (*39*).

Evaluation of ^{15}N Relaxation Data. Relaxation data were first used to determine the extent of overall rotational diffusion and rotational diffusion anisotropy and then quantitated in terms of local dynamics parameters. To examine overall rotation, we selected R_2/R_1 ratios for residues which do not display rapid local dynamics (^{1}H−^{15}N NOE > 0.65) and which also have limited contributions to R_2 from conformational exchange (*40*). The energy-minimized average CBFβ_{141} structure was used to fit an optimal diffusion model to selected experimental R_2/R_1 ratios using the programs pdbinertia 1.1 and quadric_diffusion 1.11 (A. G. Palmer, Columbia University, New York, NY). An axial symmetric diffusion model was selected on the basis of *F* tests (*41*) (Results). Values for local rotational correlation times, τ_M, were then calculated from optimized overall diffusion coefficients and local orientations of the NH vectors with respect to the symmetry axis for rotational diffusion (*30, 42*).

Local backbone dynamics were quantitated using parameters in spectral density functions of the Lipari−Szabo type (*43*−*45*). Model selection was carried out using the program Modelfree 4.01 (A. G. Palmer, Columbia University). With three measured relaxation parameters, it is possible to fit up to three parameters in five different ways. The five models include fitting S^2 (model 1), S^2 and τ_e (model 2), S^2 and R_{ex} (model 3), S_s^2, τ_e, and R_{ex} (model 4), or S_f^2, S^2, and τ_e (model 5), where S^2 ($=S_s^2S_f^2$) is the square of the generalized order parameter reflecting the amplitude of internal motions, S_s^2 and S_f^2 are the squares of the order parameters for dynamics on slow and fast time scales, respectively, τ_e is the effective correlation time for internal motions, and R_{ex} is a conformational exchange contribution to the R_2 relaxation rate. For each residue, we first tested if model 1 was sufficient for a good fit of experimental data. The fit was considered good if Γ_i was less than $\chi^2_{0.10,2}$. Goodness of fit, Γ_i, and critical values, $\chi^2_{\alpha,\nu}$, were defined and used as described in ref *30*. Next, we examined if a fit of model 3 to the remaining residues yielded an improved fit, and model 3 was selected if Γ_i was less than $\chi^2_{0.10,1}$. Residues with relaxation behavior that did not fit model 1 or 3 were analyzed using model 2. With these, we also tested model 4 if the following three criteria were fulfilled: (i) Γ_i for model 2 was greater than $15\chi^2_{0.10,1}$, (ii) $\Gamma_i = 0$ for model 4, and (iii) the uncertainty in the R_{ex} value (with model 4) was smaller than the R_{ex} value. Model 5 was tested in a similar way. Following model selection, we classified the fits as very good if Γ_i was less than $\chi^2_{0.10,\nu}$, good if Γ_i was less than $4\chi^2_{0.10,\nu}$, acceptable if Γ_i was less than $9\chi^2_{0.10,\nu}$, and poor if Γ_i was greater than $9\chi^2_{0.10,\nu}$. Residues with poor fits were excluded from further analysis. Uncertainties in optimized dynamics parameters were obtained from Monte Carlo simulations (*46*).

RESULTS

Backbone Conformation of Apo-CBFβ. The SA ensemble of CBFβ_{141} structures is shown in Figure 1, and structural statistics are summarized in Table 1. The overall topology is that of a partially open up-and-down six-stranded β-barrel surrounded by α-helices of different lengths (Figure 2), and it is very similar to the topology observed in other studies of CBFβ (*17*, *18*, *22*−*24*). The structural precision of the ensemble is good with the exception of a long loop (Ala71−Glu84) and α-helix 4. These regions are less well determined due to a lack of long-range distance restraints. The long loop is followed by a short stretch of extended peptide (Tyr85−Asp87). The conformation of this peptide fragment is well-defined, and it is very likely that it forms an extension of the β-barrel by interacting with β-strand 4 (we have identified one putative hydrogen bond and several NOEs supporting such an extended secondary structure). We were able to identify 11 of the 19 observable main chain amide moieties in the polyhistidine tag. According to the chemical shift index based on the α-carbon chemical shifts (*47*), the polyhistidine tag adopts a random coil conformation. Only intraresidue NOEs could be observed for the resonances in the polyhistidine tag.

Assignment and Conformation of the Tyr29 Hydroxyl. The side chain of Tyr29 occupied two different conformations during the early stages of the iterative structure refinement process. One conformation was solvent-exposed, and the other was buried in a hydrophobic environment that consisted of side chains from residues in α-helix 3 and some surrounding regions. Further examination of the 2D NOESY

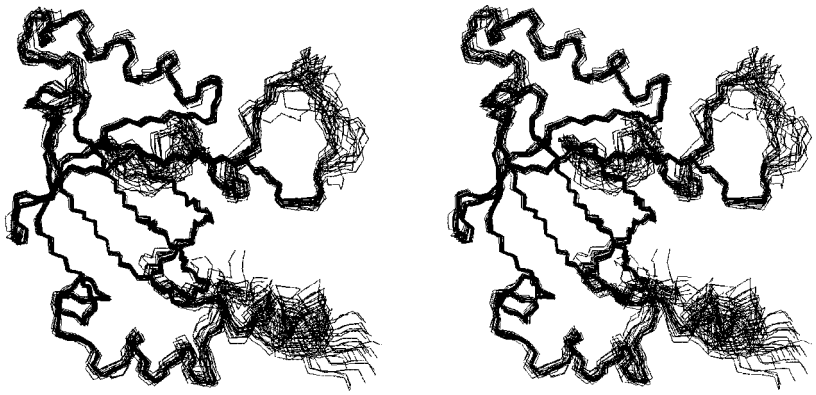

FIGURE 1: Stereoview of the 25 SA models representing the solution structure ensemble of CBFβ_{141}. The backbone atoms of the β-strands are superimposed. All molecular illustrations in this article were prepared with MOLMOL (*39*).

Table 1: Structural Statistics for the Ensemble of SA Models of CBFβ_{141}

no. of NOE-derived distances	
intraresidue	596
sequential ($\|i - j\| = 1$)	516
medium-range ($1 < \|i - j\| < 4$)	363
long-range ($\|i - j\| > 4$)	664
total	2139
no. of hydrogen bonds	43
no. of dihedral angle restraints	
ϕ obtained from HNHA	44
ϕ obtained from TALOS	52
ψ obtained from TALOS	83
χ^1	36
no. of stereospecific assignments	
β-methylene	36
Val and Leu methyl groups	32
Ramachandran plot analysis	
residues in most favored regions	82.5%
residues in additional allowed regions	15.2%
residues in generously allowed regions	2.2%
residues in disallowed regions	0.1%
coordinate precision	
residues in β-strands (24−29, 52−57, 62−67, 95−103, 106−114, and 120−127)	
rmsd from mean (backbone atoms)	0.27 Å
rmsd from mean (all heavy atoms)	0.71 Å
residues 1−141	
rmsd from mean (backbone atoms)	0.78 Å
rmsd from mean (all heavy atoms)	1.39 Å
deviations from idealized covalent geometry	
bonds	0.0028 ± 0.00008 Å
angles	$0.47 \pm 0.01°$
impropers	$0.36 \pm 0.02°$
no. of violations	
distance restraints (>0.5 Å)	0
dihedral angle restraints (>5°)	0.24

spectrum led to assignment of the H$^\eta$ proton of Tyr29 resonating at 6.79 ppm. Within the Tyr29 side chain, the H$^\eta$ proton has a strong NOE to the H$^\delta$ protons and a weak NOE to the H$^\epsilon$ protons. These NOEs lack corresponding DQF-COSY cross-peaks, as expected. Several long-range NOEs could subsequently be assigned between the H$^\eta$ proton and side chain protons of, for example, Phe44 in α-helix 3. When these distance restraints were included in the structure calculations, the side chain of Tyr29 relaxed into the buried conformation (Figure 2). A buried and slowly exchanging tyrosyl H$^\eta$ proton is expected to participate in hydrogen bonding. The structural precision of the Tyr29 side chain is well-defined in relation to neighboring groups, and it appears

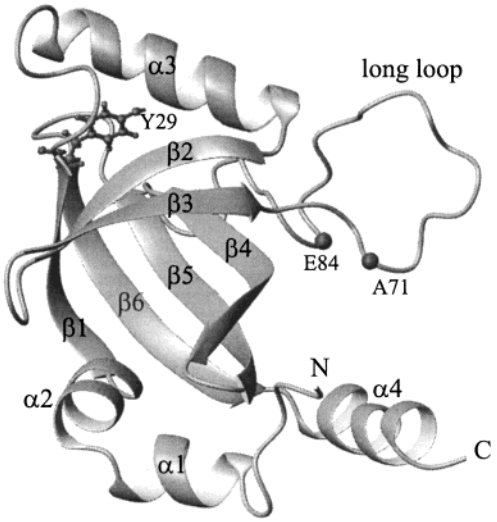

FIGURE 2: Ribbon drawing of CBFβ_{141} with notation of β-strands, α-helices, and the long loop. The first CBFβ_{141} structure of the SA ensemble is shown here. The side chain of Y29 is also shown, and the C$^\alpha$ atoms of A71 and E84 are highlighted to mark the location of the long loop.

to be fixed by hydrogen bonds between the hydroxyl oxygen and the backbone amide of Glu54, and between the hydroxyl proton and the backbone carbonyl oxygen of Arg40.

Backbone Dynamics of CBFβ. The backbone ^{15}N relaxation parameters of CBFβ_{141} at a ^{1}H frequency of 600 MHz are shown in Figure 3. To quantify these parameters with the model free approach, we first determined which overall rotational diffusion model would best fit the relaxation data. The R_2/R_1 ratios from 98 selected NH vectors were found to fit an axially symmetric diffusion model with an average rotational correlation time, $\tau_{M,iso}$, of 10.00 ± 0.11 ns and a ratio of rotational diffusion coefficients, D_{par}/D_{per}, of 1.21 ± 0.021 (Table 2). Our estimate of $\tau_{M,iso}$ is somewhat lower than the value of 11.4 ns that was estimated in ref *48*. An axial diffusion tensor with these parameters corresponds to local correlation times between 9.58 and 10.65 ns depending on the orientation of the NH vector with respect to the diffusion symmetry axis. Local τ_M constants were fixed to calculated values before model selection and optimization of other local dynamics parameters. Of the total of 111 residues with determined relaxation parameters, 9 were excluded due to poor fits to all models. Of the remaining

Table 2: Rotational Diffusion Models for CBFβ_{141}[a]

model	$\tau_{M,iso}$ (ns)[b]	D_{par}/D_{per}	D_x/D_y	χ^2_{red}	F[c]	p[d]
isotropic	10.02 ± 0.11	–	–	6.970	–	–
axial symmetric	10.00 ± 0.11	1.21 ± 0.021	–	5.921	6.66	0.0004
anisotropic	10.01 ± 0.11	1.21 ± 0.022[e]	1.08 ± 0.025	5.932	0.90	0.41

[a] The evaluation is based on 98 R_2/R_1 values and the averaged and energy-minimized CBFβ_{141} structure, as described in the text. [b] Effective correlation time [$\tau_{M,iso} = 1/6D_{iso}$, where $D_{iso} = (2D_{per} + D_{par})/3$]. [c] F ratio (*41*) for assessing the validity of a reduction in χ^2_{red} when additional parameters are added to the model. [d] Probability that the reduction in χ^2_{red} is achieved by chance. [e] $D_{par}/D_{per} = 2D_z/(D_x + D_y)$.

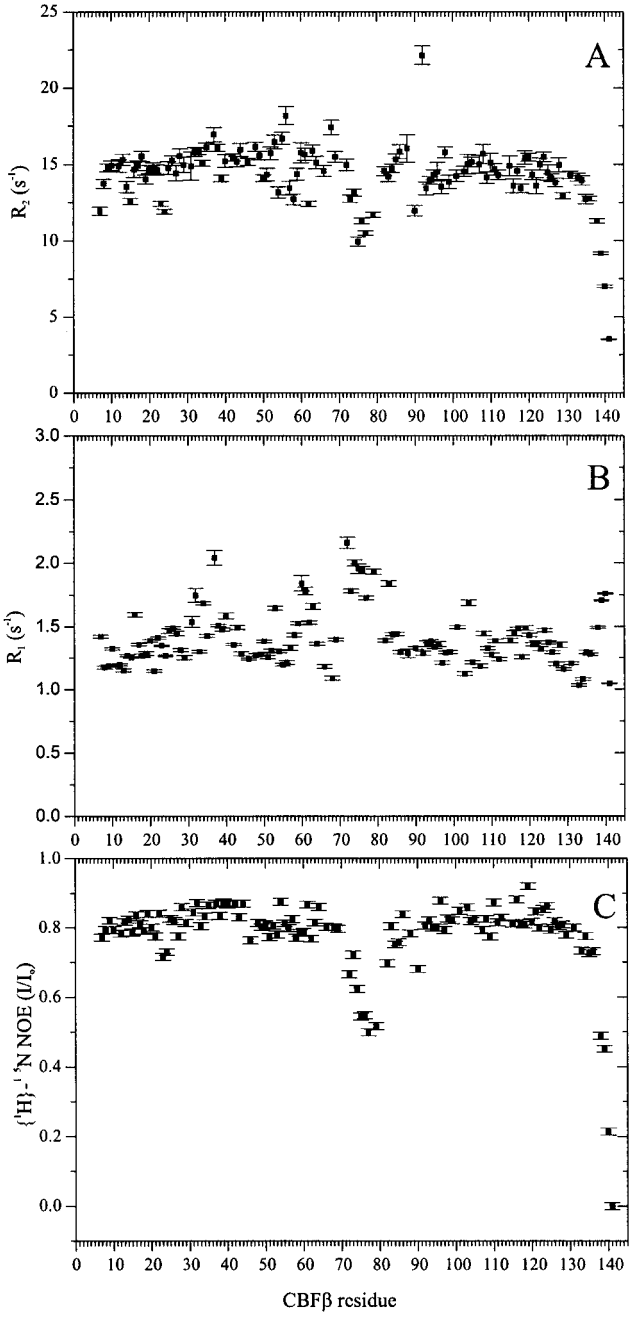

FIGURE 3: Backbone ^{15}N relaxation parameters of CBFβ_{141} measured at a ^{1}H resonance frequency of 600 MHz. R_2 (A), R_1 (B), and $\{^1H\}-^{15}N$ NOE values (C) could be determined for 111 of 135 total non-proline amino acid residues.

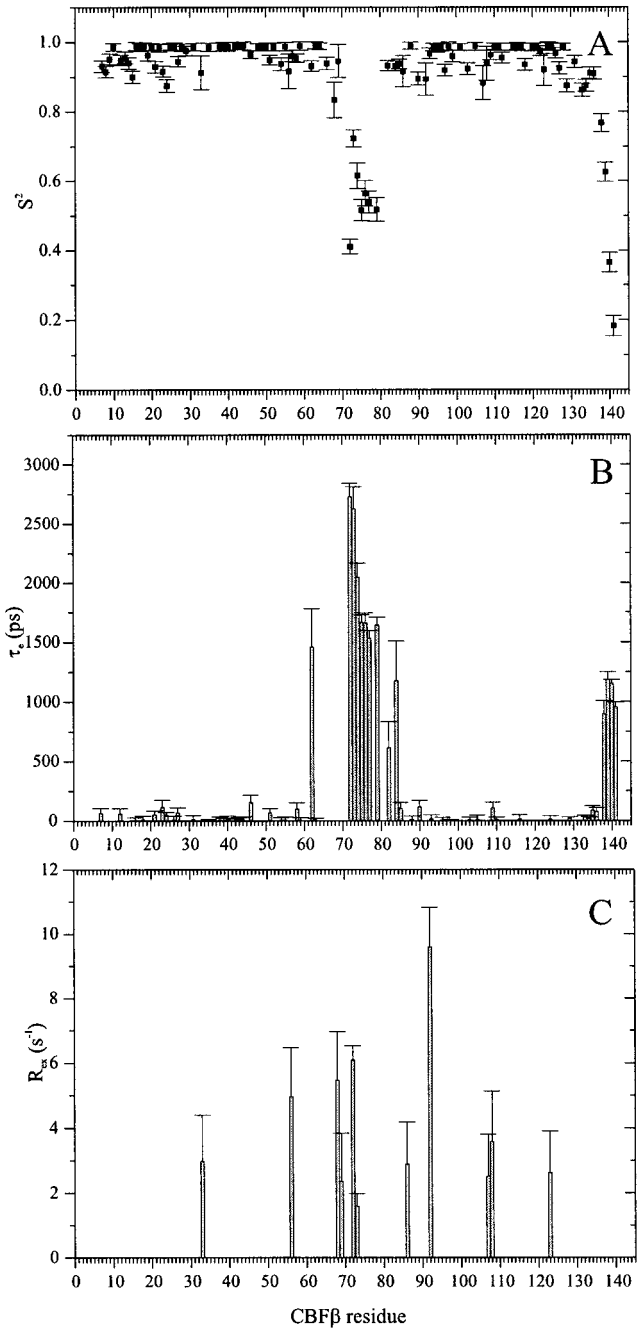

FIGURE 4: Model free backbone dynamics parameters and R_{ex} values from analysis of ^{15}N NMR relaxation in CBFβ_{141}. In all, 102 of the 111 amino acid residues with determined relaxation parameters could be fitted with models 1−5 (see the text). (A) Order parameters (S^2), (B) effective correlation times for internal motions (τ_e), and (C) conformational exchange contributions (R_{ex}) to transverse relaxation rates.

102 residues, we could fit 38, 52, 9, 2, and 1 relaxation parameter set to models 1−5, respectively. Of these, 50, 42, and 10 fits were judged to be very good, good, and acceptable, respectively. Local dynamics parameters and R_{ex} terms for CBFβ are shown in Figure 4.

Backbone order parameters of CBFβ_{141} are high, indicating a well-ordered and rigid molecule. This is expected since a

Table 3: Precision of the Three CBFβ NMR Ensembles and Comparison with CBFβ−RD Complex

	1ILF (this study)	2JHB	1CL3
precision of apo-CBFβ			
secondary structure C^{α} [a]	0.46 Å	0.50 Å	0.47 Å
heavy atoms in all residues[b]	1.37 Å	1.24 Å	1.32 Å
comparison to the CBFβ−RD complex[c]			
secondary structure C^{α} [d]	1.03 ± 0.04 Å	1.92 ± 0.08 Å	2.65 ± 0.05 Å

[a] The rmsd (pairwise against mean) is calculated for C^{α} atoms in secondary structure elements determined in this study, i.e., residues 8−14, 16−22, 24−29, 37−50, 52−57, 62−67, 95−103, 106−114, 120−127, and 129−139. [b] Residues 4−141, which are present in all three CBFβ ensembles. [c] Average rmsd values and standard deviations in pairwise comparisons of individual NMR models in the three ensembles to the B-chain CBFβ monomer in the CBFβ−RD complex (*24*). [d] As defined by MOLMOL (*39*) for the B-chain CBFβ monomer in the X-ray structure of the CBFβ−RD complex (residues 8−14, 16−22, 25−29, 37−49, 52−57, 62−67, 86, 87, 94−103, 106−115, 120−127, and 129−133).

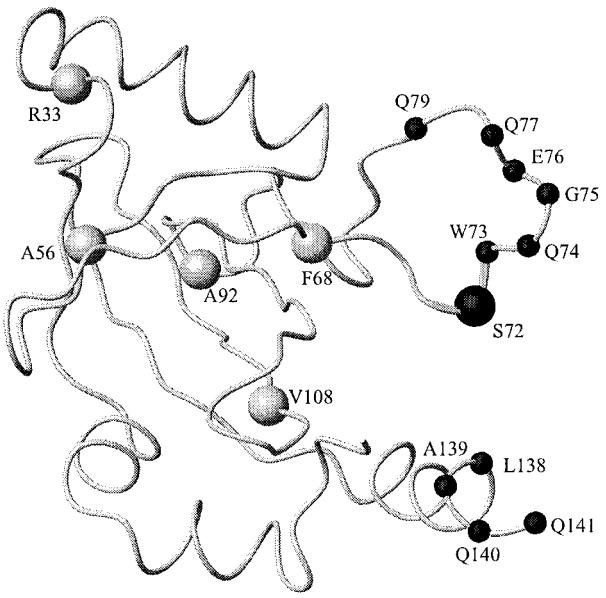

FIGURE 5: Location of significant internal motions or conformational exchange in the CBFβ_{141} backbone. Small black spheres indicate residues with an S^2 of <0.8. Large gray spheres indicate residues with an R_{ex} of >3.0 s^{-1}. The large black sphere indicates an S^2 of <0.8 and an R_{ex} of >3.0 s^{-1}.

β-barrel is a rigid fold with numerous hydrogen bonds connecting the antiparallel β-strands and leaving only limited degrees of freedom for backbone nuclei within the barrel. There are only two regions with lower order parameters: the long loop and the C-terminal α-helix. Both these regions have low structural precision in our ensemble (Figure 1). The reason for the lack of long-range NOEs in these regions therefore appears to be the large amplitude of the internal motions. Only 11 residues were fitted with conformational exchange contributions to the observed R_2 relaxation rates, with the largest being 9.6 s^{-1} for A92 and the smallest 1.59 s^{-1} for W73. Amino acid residues with significant internal motions or conformational exchange contributions are mapped on the CBFβ_{141} backbone structure in Figure 5.

Comparison of Apo-CBFβ Structures Determined by NMR. Two independently determined NMR structures of CBFβ have been published [PDB entries 2JHB (*23*) and 1CL3 (*22*)]. We have compared these to each other and to our apo-CBFβ structure. The structural precision of backbone and side chain atoms is very similar in the three NMR ensembles (Table 3). The β-barrel parts also overlap very well (Figure 6A); the rmsd value following a superposition of the β-strands of all 70 models in the three ensembles is only 0.78 Å. If instead residues 4−141 are superimposed, the rmsd increases to 2.6 Å, indicating that there are structural differences in regions outside of the β-barrel. Inspection of the structures reveals that the differences are localized to four regions (Figure 6). These regions have, in all three ensembles, somewhat lower precisions than the β-barrel core. Regions 1, 2, and 4 are still relatively well defined, whereas region 3, containing the long loop, is less defined in all structures.

Region 1 (Gln8−Ser22 and Glu129−Ala139) contains α-helices 1, 2, and 4, which have different relative orientations in the three ensembles. The conformation we observe here, as well as the conformations of regions 2−4, is supported by several restraints derived from NOEs, *J* couplings, and chemical shifts, of which we mention a few critical ones. For instance, we have identified long-range NOEs between Ser22 H^{α} in α-helix 2 and Glu126 H^{N} in β-strand 6, between Val4 H^{α} in α-helix 1 and Gly105 H^{N} in the turn between β-strands 4 and 5, and between the H^{δ} protons of Phe12 in α-helix 1 and the H^{γ} protons of Glu129 in the beginning of α-helix 4. These NOEs contribute to the relative orientations of helices 1, 2, and 4 that we observe. Furthermore, the $^3J_{HNH\alpha}$ coupling constant for Asp128 was measured to be 8.8 Hz. This value was interpreted as a ϕ angle restraint of $-120 \pm 30°$. We also included a ψ angle restraint for Asp128 obtained from TALOS of $122 \pm 20°$. This ϕ and ψ restraint pair directs Asp128 into an allowed region of the Ramachandran diagram. Differences in the main chain torsion angles at Asp128 might inflict the different orientations of C-terminal α-helix 4 in the three ensembles.

Region 2 (Tyr29−Asp50) contains α-helix 3 and eight preceding amino acid residues. We assigned long-range NOEs in this region between the side chains of Tyr29 and His37 (for instance, between Tyr29 H^{α} and H37 $H^{\delta 2}$ protons), for which resonances were resolved in the 2D NOESY spectrum acquired in D_2O. The $^3J_{HNH\alpha}$ coupling constant for Glu38 (the second residue of α-helix 3) was measured to be 3.7 Hz, and this value corresponds to a ϕ angle restraint of $-60 \pm 30°$. The ψ angle suggested by TALSO for Glu38 was $-44 \pm 20°$, which is consistent with a peptide in a helical conformation.

Region 3 (Pro70−Ser82) contains a large fragment of the long loop. The loop is less well defined, but its conformation in the three NMR ensembles appears to be significantly different also when the lower resolution is considered (Figure 6A). In region 3, we assigned long-range NOEs between the H^{β} methyl protons of Ala71 and the H^{δ} protons of Arg131.

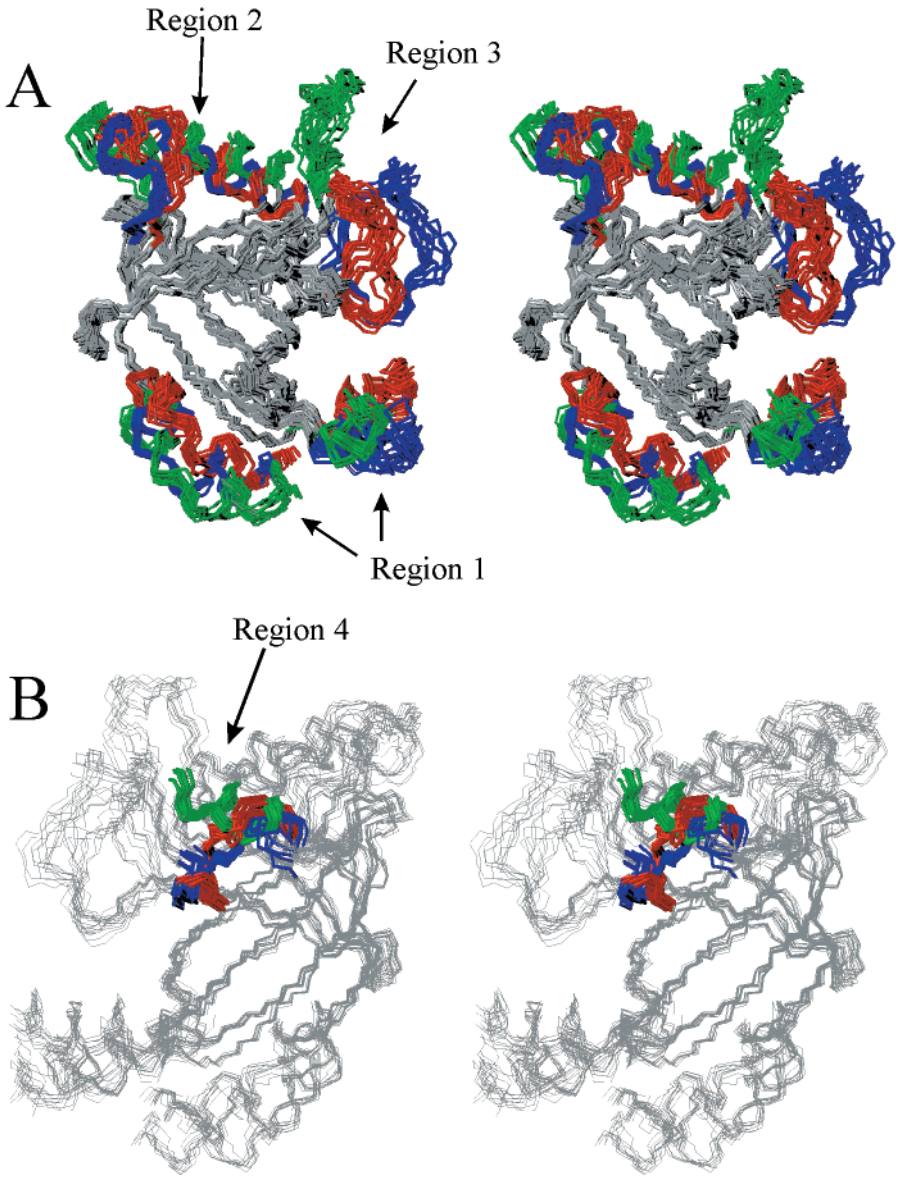

FIGURE 6: (A) Comparison of three CBFβ NMR structures determined in different laboratories. The models were superimposed on C^{α} atoms in β-strands. For clarity, residues Met1−Asp7, Gln140, and Gln141 are not shown and only the first 10 models in each ensemble are presented. Three regions with different backbone conformations in the three ensembles are colored blue, red, and green for the 1ILF, 2JHB, and 1CL3 ensembles, respectively. Region 4 is partially hidden in this view and is shown in panel B. The core β-barrel of CBFβ is colored gray. (B) Conformations of region 4 in the three NMR ensembles. Residues Arg83−Arg90 are color-coded as the residues are in panel A. The view is rotated by approximately 180° compared to the view in panel A.

For Gln79, TALOS suggested ϕ and ψ angles of −105° and 140°, respectively. These torsion angles force Gln79 to adopt an extended conformation.

Region 4 (Arg83−Arg90) contains residues terminating and following the long loop. Here, the 1CL3 ensemble contains an eight-residue 3_{10}-helix that is absent in our ensemble and the 2JHB ensembles (Figure 6B). We find no NOE data supporting the presence of this helical segment. In contrast, we identified NOEs that are consistent with a putative backbone hydrogen bond between the amide proton of Lys98 (β-strand 4) and the backbone carbonyl oxygen of Tyr85. This observation is in good agreement with the possibility that this fragment indeed adopts a β-strand also in the apo state. The chemical shift index for CBFβ has been published previously (*26*), and it is also consistent with a β-strand conformation. The $^3J_{HNH\alpha}$ coupling constant for Glu89 was measured to be 9.9 Hz, a value which was interpreted as a ϕ angle of $-120 \pm 30°$. The ψ angle suggested by TALOS for Glu89 was $142 \pm 30°$. This is clear evidence that Glu89 adopts an extended conformation and is not part of a helical segment. Furthermore, long-range NOEs were assigned between the H^N resonances of Asp87 and Tyr96, and between H^{α} of Leu88 and H^N of Tyr96, and these are consistent with the calculated structure.

The origin of the differences between the three NMR structures may, at least in part, arise due to experimental differences. For instance, we note that the N-termini of the three studied molecules differ in length. The possible influence of the longer N-terminal fragment with a histidine tag used here and other differences in experimental conditions are discussed below.

Comparison of Apo- and RD-Bound CBFβ. We have made a comparison of the three NMR apo-CBFβ structures and CBFβ in the crystallographically determined structure of a

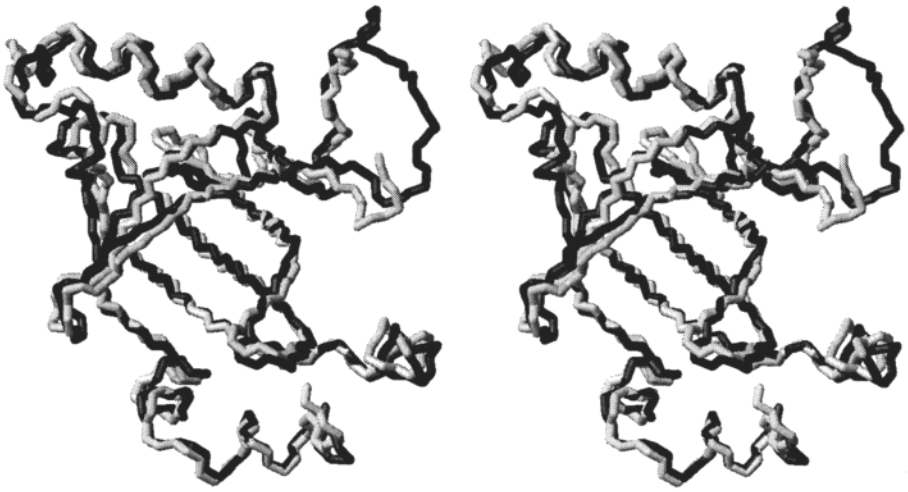

FIGURE 7: Backbone stereoview comparison of the B-chain CBFβ monomer in the CBFβ−RD X-ray structure (gray) (*24*) and our mean and energy-minimized NMR structure (black). Amino acids 5−135 are shown, with the exception of Gln74−Gln77 which are missing in the X-ray structure. The two molecules were superimposed on C^{α} atoms in secondary structure elements found in the X-ray structure (see Table 2).

CBFβ−RD complex (*24*) (PDB entry 1E50). The rmsd between the four different CBFβ chains (B, D, E, and F) in the X-ray structure is 0.06 Å for the C^{α} atoms in residues 2−70 and 80−135. The chains therefore have identical conformations, and we chose one of them, the B-chain monomer, for the comparison. The B-chain is also very similar to CBFβ in the quaternary RD−CBFβ−C/EBPβ−DNA complex (*17*) with an rmsd of 0.54 Å for a superimposition of C^{α} atoms in residues 2−70 and 82−135. The differences between the apo-CBFβ models and the X-ray structure of the CBFβ−RD complex were measured, and the smallest rmsd values found for the individual NMR models are summarized in Table 3. The backbone conformation of our mean and energy-minimized apo-CBFβ structure is very similar to that of the RD-bound structure (Figure 7) with an average rmsd of 1.0 Å, which can be compared to rmsds of 1.9 and 2.7 Å for the 2JHB and 1CL3 ensembles, respectively. The conformations of the long loop in the apo and RD-bound states cannot be compared, as some residues in the loop are absent in the X-ray structure.

DISCUSSION

We have determined the solution structure and quantified the backbone dynamics of CBFβ in solution. When comparing our apo-CBFβ structure to that observed in the heterodimeric CBFβ−RD complex (*24*), we find no evidence of substantial rearrangements of the backbone conformation upon RD binding. Not only the β-barrel but also the surrounding α-helices adopt very similar conformations in the two states. Analysis of the backbone ^{15}N relaxation parameters also reveals that most of apo-CBFβ is very rigid with high S^2 values for backbone NH vectors. The only parts of CBFβ that display significant dynamics on the nanosecond time scale are residues in the long loop and the C-terminal α-helix. The long loop probably maintains its flexible nature in the dimeric and ternary states since it could not be traced in either of two crystallographic studies of these two complexes (*17*, *24*). The role of the C-terminal helix is not clear, but it is distant from the RD-interacting surface. The dynamics on longer time scales for CBFβ are also restricted as there are only a few backbone amides with conformational exchange contributions to observed ^{15}N R_2 relaxation rates. These residues are distributed over the molecule with no obvious correlation to RD-binding regions (defined in ref *24*), with the exception of Arg33 that directly interacts with the RD (*24*). Arg33 has an R_{ex} term of 3.0 s^{-1}, and the possibility that conformational exchange might be important for RD binding in this region cannot be excluded. Overall, though, it appears that conformational changes do not occur upon binding to the RD and that local dynamics in apo-CBFβ are not implicated in the recognition process. The crystallographic analyses also show that there is no structural adaptation of CBFβ between the dimeric (CBFβ−RD) and trimeric (CBFβ−RD−DNA) states (*17*, *24*). On the basis of the data presented here and these structures, we therefore conclude that the role of CBFβ is to provide a rigid template for the RD to stabilize the optimal DNA-binding conformation of the RD.

The β-barrel part of our structure is very similar to that observed in the two independently determined NMR structures of apo-CBFβ. There are, however, also significant differences between the three ensembles, such as in the relative orientation of α-helices 1−4, the conformation of the long loop, and the secondary structure of residues Arg83−Arg90. Several experimental observations supporting the conformations reported here have been described above. We have also examined to what extent our restraints are compatible with the 2JHB and 1CL3 models. We find that a large number of NOEs (and coupling constants) are incompatible with these structures. Most of these NOEs were identified in the 2D NOESY spectrum acquired at 800 MHz using a sample dissolved in D_2O. It appears that high-resolution NOESY data for samples dissolved in D_2O were not collected in either of the two other structure determinations, and it is possible that the observed differences in part arise from this fact. However, we cannot draw any further conclusions in this matter, because we have not examined the experimental data used for the calculations of the 2JHB and 1CL3 structures that, to the best of our knowledge, have not been published. It is also important to note that the three

structures were not determined under identical experimental conditions. First, the three protein fragments differ in the way they were subcloned and purified. The 1CL3 ensemble was determined from a native protein fragment. The 2JHB ensemble contains two additional amino acid residues N-terminal to Met1, and our CBFβ fragment contains an N-terminal 6*-polyhistidine tag that in total adds 20 amino acid residues N-terminal to Met1. The chemical shift index for the polyhistidine tag and the fact that only intraresidue NOEs can be observed in the N-terminus suggest that the tag adopts a random coil conformation and probably is completely flexible. It is unlikely that a flexible purification tag would affect the structure of CBFβ and cause all the conformational differences that we observe, but the possibility cannot be ruled out. Second, the three structures have been determined at similar pH values, but there are some differences in ionic strength and temperature. Our structure was determined in 20−150 mM NaCl and at 24 °C, whereas the 1CL3 and 2JHB structures were determined without salt and at temperatures of 30 and 20 °C, respectively. Here too, it is unlikely that the small differences in experimental conditions cause rather substantial differences in the conformations of four regions in a protein, but salt and temperature effects cannot be completely ruled out. To summarize, it is not possible to draw conclusions about the origin of conformational differences in three published NMR models of apo-CBFβ. However, it is interesting to note that NMR studies of a protein by independent investigators using very similar techniques can yield structures that differ in regions that appear to be well-resolved in the individual ensembles.

ACKNOWLEDGMENT

We thank Pharmacia & UpJohn and the Swedish NMR Center for the use of their 800 MHz NMR spectrometers. Dr. Alan J. Warren is acknowledged for giving us access to the coordinates of the CBFβ−RD complex. We also thank Dr. Wayne Boucher and the Department of Biochemistry, University of Cambridge, for the kind distribution of the Azara suit of programs. Dr. Peter Allard is acknowledged for assistance with both NMR spectroscopy and analysis of relaxation data.

REFERENCES

1. Kania, M. A., Bonner, A. S., Duffy, J. B., and Gergen, J. P. (1990) *Genes Dev. 4*, 1701−1713.
2. Tang, Y. Y., Crute, B. E., Kelley, J. J., Huang, X., Yan, J., Shi, J., Hartman, K. L., Laue, T. M., Speck, N. A., and Bushweller, J. H. (2000) *FEBS Lett. 470*, 167−172.
3. Okuda, T., van Deursen, J., Hiebert, S. W., Grosveld, G., and Downing, J. R. (1996) *Cell 84*, 321−330.
4. Wang, Q., Stacy, T., Binder, M., Marín-Padilla, M., Sharpe, A. H., and Speck, N. A. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 3444−3449.
5. Takakura, N., Watanabe, T., Suenobu, S., Yamada, Y., Noda, T., Ito, Y., Satake, M., and Suda, T. (2000) *Cell 102*, 199−209.
6. Speck, N. A., Stacy, T., Wang, Q., North, T., Gu, T. L., Miller, J., Binder, M., and Marín-Padilla, M. (1999) *Cancer Res. 59* (Suppl.), 1789−1793.
7. Song, W. J., Sullivan, M. G., Legare, R. D., Hutchings, S., Tan, X., Kufrin, D., Ratajczak, J., Resende, I. C., Haworth, C., Hock, R., Loh, M., Felix, C., Roy, D. C., Busque, L., Kurnit, D., Willman, C., Gewirtz, A. M., Speck, N. A., Bushweller, J. H., Li, F. P., Gardiner, K., Poncz, M., Maris, J. M., and Gilliland, D. G. (1999) *Nat. Genet. 23*, 166−175.
8. Friedman, A. D. (1999) *Leukemia 13*, 1932−1942.
9. Westendorf, J. J., and Hiebert, S. W. (1999) *J. Cell. Biochem. 32/33* (Suppl.), 51−58.
10. Lutterbach, B., and Hiebert, S. W. (2000) *Gene 245*, 223−235.
11. Mundlos, S. (1999) *J. Med. Genet. 36*, 177−182.
12. Wang, S., Wang, Q., Crute, B. E., Melnikova, I. N., Keller, S. R., and Speck, N. A. (1993) *Mol. Cell. Biol. 13*, 3324−3339.
13. Ogawa, E., Inuzuka, M., Maruyama, M., Satake, M., Naito-Fujimoto, M., Ito, Y., and Shigesada, K. (1993) *Virology 194*, 314−331.
14. Golling, G., Li, L. H., Pepling, M., Stebbins, M., and Gergen, J. P. (1996) *Mol. Cell. Biol. 16*, 932−942.
15. Lu, J., Maruyama, M., Satake, M., Bae, S. C., Ogawa, E., Kagoshima, H., Shigesada, K., and Ito, Y. (1995) *Mol. Cell. Biol. 15*, 1651−1661.
16. Kagoshima, H., Akamatsu, Y., Ito, Y., and Shigesada, K. (1996) *J. Biol. Chem. 271*, 33074−33082.
17. Tahirov, T. H., Inoue-Bungo, T., Morii, H., Fujikawa, A., Sasaki, M., Kimura, K., Shiina, M., Sato, K., Kumasaka, T., Yamamoto, M., Ishii, S., and Ogata, K. (2001) *Cell 104*, 755−767.
18. Bravo, J., Li, Z., Speck, N. A., and Warren, A. J. (2001) *Nat. Struct. Biol. 8*, 371−378.
19. Bäckström, S., Huang, S. H., Wolf-Watz, M., Xie, X. Q., Härd, T., Grundström, T., and Sauer, U. H. (2001) *Acta Crystallogr. D57*, 269−271.
20. Wolf-Watz, M., Xie, X. Q., Holm, M., Grundström, T., and Härd, T. (1999) *Eur. J. Biochem. 261*, 251−260.
21. Wolf-Watz, M., Bäckstrom, S., Grundström, T., Sauer, U., and Härd, T. (2001) *FEBS Lett. 488*, 81−84.
22. Goger, M., Gupta, V., Kim, W. Y., Shigesada, K., Ito, Y., and Werner, M. H. (1999) *Nat. Struct. Biol. 6*, 620−623.
23. Huang, X., Peng, J. W., Speck, N. A., and Bushweller, J. H. (1999) *Nat. Struct. Biol. 6*, 624−627.
24. Warren, A. J., Bravo, J., Williams, R. L., and Rabbitts, T. H. (2000) *EMBO J. 19*, 3004−3015.
25. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., and Skelton, N. J. (1996) *Protein NMR Spectroscopy Principles and Practice*, Academic Press, San Diego.
26. Huang, X., Speck, N. A., and Bushweller, J. H. (1998) *J. Biomol. NMR 12*, 459−460.
27. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
28. Helgstrand, M., Kraulis, P., Allard, P., and Härd, T. (2000) *J. Biomol. NMR 18*, 329−336.
29. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) *Biochemistry 33*, 5984−6003.
30. Wikström, A., Berglund, H., Hambraeus, C., van den Berg, S., and Härd, T. (1999) *J. Mol. Biol. 289*, 963−979.
31. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
32. Laskowski, R. A., Rullmannn, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−486.
33. Fletcher, C. M., Jones, D. N. M., Diamond, R., and Neuhaus, D. (1996) *J. Biomol. NMR 8*, 292−310.
34. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
35. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR 13*, 289−302.
36. Wagner, G., Hyberts, S. G., and Havel, T. F. (1992) *Annu. Rev. Biophys. Biomol. Struct. 21*, 167−198.
37. Neri, D., Szyperski, T., Otting, G., Senn, H., and Wüthrich, K. (1989) *Biochemistry 28*, 427−433.
38. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski,

J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
39. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
40. Tjandra, N., Feller, S. E., Pastor, R. W., and Bax, A. (1995) *J. Am. Chem. Soc. 117*, 12562−12566.
41. Devore, J. L. (1995) *Probability and Statistics for Engineering and the Sciences*, 4th ed., Wadsworth, Inc./International Thomson Publishing Inc., New York.
42. Lee, L. K., Rance, M., Chazin, W. J., and Palmer, A. G. (1997) *J. Biomol. NMR 9*, 287−298.
43. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4546−4559.
44. Lipari, G., and Szabo, A. (1982) *J. Am. Chem. Soc. 104*, 4559−4570.
45. Clore, G. M., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C., and Gronenborn, A. M. (1990) *J. Am. Chem. Soc. 112*, 4989−4991.
46. Mandel, A. M., Akke, M., and Palmer, A. G. (1995) *J. Mol. Biol. 246*, 144−163.
47. Wishart, D. S., and Sykes, B. D. (1994) *J. Biomol. NMR 4*, 171−180.
48. Huang, X., Crute, B. E., Sun, C., Tang, Y. Y., Kelley, J. J., Lewis, A. F., Hartman, K. L., Laue, T. M., Speck, N. A., and Bushweller, J. H. B. (1998) *J. Biol. Chem. 4*, 2480−2487.

BI010713+