Immunoglobulin motif DNA recognition and heterodimerization of the PEBP2/CBF Runt domain

Takashi Nagata[1,2], Vineet Gupta[1,2], Damian Sorce[2], Woo-Young Kim[3], Andrej Šali[2], Brian T. Chait[4], Katsuya Shigesada[3], Yoshiaki Ito[3] and Milton H. Werner[2]

[1]These authors contributed equally to this work. [2]Laboratories of Molecular Biophysics, The Rockefeller University, 1230 York Avenue, Box 42, New York, NY 10021, USA. [3]Institute for Virus Research, Kyoto University, Japan. [4]Laboratory of Mass Spectrometry and Gaseous Ion Chemistry, The Rockefeller University, 1230 York Avenue, Box 42, New York, NY 10021, USA.

The polyomavirus enhancer binding protein 2 (PEBP2) or core binding factor (CBF) is a heterodimeric enhancer binding protein that is associated with genetic regulation of hematopoiesis and osteogenesis. Aberrant forms of PEBP2/CBF are implicated in the cause of the acute human leukemias and in a disorder of bone development known as cleidocranial dysplasia. The common denominator in the natural and mutant forms of this protein is a highly conserved domain of PEBP2/CBFα, termed the Runt domain (RD), which is responsible for both DNA binding and heterodimerization with the β subunit of PEBP2/CBF. The three-dimensional structure of the RD bound to DNA has been determined to be an S-type immunoglobulin fold, establishing a structural relationship between the RD and the core DNA binding domains of NF-κB, NFAT1, p53 and the STAT proteins. NMR spectroscopy of a 43.6 kD RD–β–DNA ternary complex identified the surface of the RD in contact with the β subunit, suggesting a mechanism for the enhancement of RD DNA binding by β. Analysis of leukemogenic mutants within the RD provides molecular insights into the role of this factor in leukemogenesis and cleidocranial dysplasia.

The Runt domain is composed of ten anti-parallel strands of β-sheet: β_O (Val 63–Arg 64), β_A (Phe 70–Ser 73), β_B (Phe 89–Ala 93), β_C (Gly 100–Asn 109), β_D (Ser 114–Arg 118), $\beta_{C'}$ (Thr 121–Lys 125), β_E (Val 128–Phe 131), $\beta_{E'}$ (Arg 135–Arg 139), β_F (Phe 146–Val 152), β_G (Gln 158–Arg 164) (Figs 1*a*, 2*a*). Two parallel strands, β_{P1} (His 78–Arg 80) and β_{P2} (Lys 167–Thr 169), form at the loop between strands β_A and β_B (the β_A-β_B loop, residues Val 74–Ile 87) and the C-terminus (Figs 1*a*, 2*a*). An all-β structure is consistent with earlier conclusions derived from circular dichroism measurements[1].

The Ig motif was first observed among transcription factors for the tumor suppressor p53 (ref. 2). Ig motif DNA recognition has since been observed for NF-κB[3-4], the nuclear factor of activated T-cells NFAT1 (refs 5,6), STAT1 (ref. 7) and STAT3β[8]. The RD is unusual among Ig motif transcription factors in that it is at least 40 amino acids shorter than the other proteins that utilize this fold for DNA recognition. A second difference can be seen at the N-terminus (Fig. 2*a*). In Runt, the N-terminus forms an additional short anti-parallel strand of β-sheet (strand β_O) with strand β_A. This was evidenced by strong NOEs between backbone amide protons for residues Val 63 and Ser 72 and several NOEs between side chains for residues Leu 62, Val 63, Arg 64 and residues Leu 71, Cys 72 and Ser 73. In contrast, the N-terminus of NF-κB[3-4] and NFAT1 (refs 5, 6) loops around the backside of the protein domain while that of the STATs[7-8] is a helical segment leading from the coiled-coil domain of the protein.

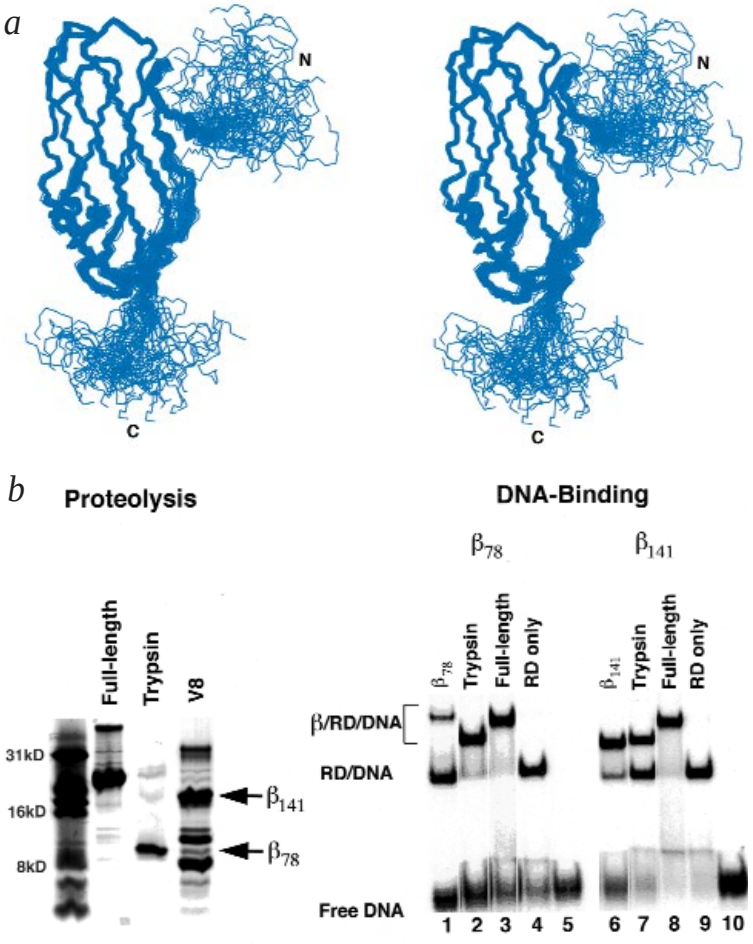

Fig.1 Three-dimensional structure of the Runt domain and its interaction with the β subunit. ***a***, Stereosuperposition of backbone heavy atoms of the 47 Runt domain three-dimensional structures, residues 51–178. The conformation of the N-terminal residues 51–62 and C-terminal residues 169–178 were poorly defined by the NOE data. Residues 179–188 are unstructured, display negative heteronuclear NOEs and are not shown. The C-terminal segment (residues 169–178) displayed heteronuclear NOEs ≤0.6, suggestive of some dynamic motion in this portion of the molecule. Heteronuclear NOEs were near zero for the N-terminal residues 51–61. ***b***, The core structured domain of PEBP2/CBFβ was determined by proteolysis/mass spectrometry mapping of the full-length protein which identified a 78 amino acid (β_{78}) and a ~140 amino acid (β_{141}) fragment that was resistant to proteolysis. EMSA of the Runt domain construct employed in this study demonstrates that the protein–DNA complex is a well-formed species in solution which is capable of binding both full-length β and the β_{141} fragment efficiently (compare lanes 6, 8 and 9), but binds β_{78} weakly (compares lane 1 with lane 6). The mobility of the supershifted complex formed with β_{141} was consistent with that observed from trypsin proteolysis of full-length β (compare lanes 6 and 7).

DNA-binding

Analysis of through-bond correlation spectra, NMR

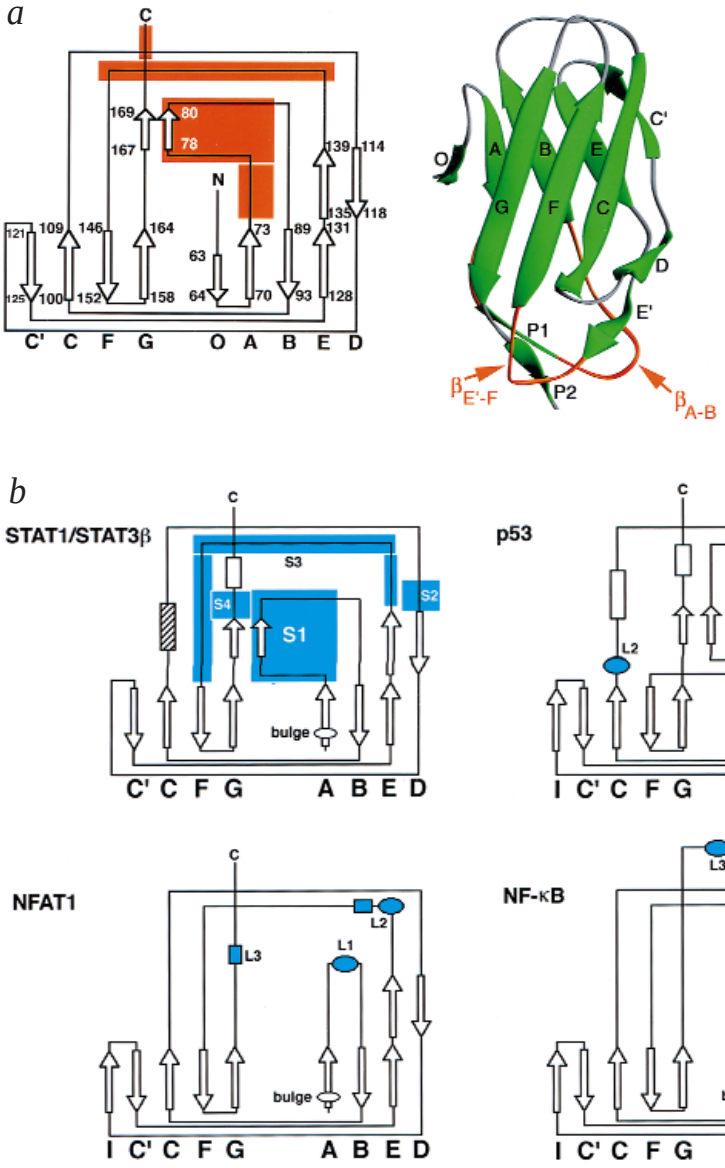

Fig. 2 Folding topology of Ig domain DNA binding proteins. ***a***, Folding topology of the Runt domain defined in this study. Sheets and helices are indicated by arrows and rectangles, respectively. Lettering at the bottom of each figure follows the established nomenclature of the immunoglobulin fold[33]. Numbering corresponds to the sequence of human AML1 (ref. 12). The hypothesized DNA binding elements for the Runt domain are shown in orange. RIBBONS[34] representation of the Runt domain (residues 62–170) indicating the elements of secondary structure (β-sheets green, loops gray and orange). The putative DNA binding loops are labeled in orange. ***b***, The topology five Ig motif transcription factors defined in their co-crystal structures with DNA: p53 (ref. 2), NF-κB[3–4], NFAT1 (refs 5, 6) STAT1 (ref. 7) and STAT3β[8]. The striped helix for STAT1/STAT3β was seen only in STAT3β[8]. 'L' or 'S' refers to the numbering of loops or segments, respectively, as defined for the indicated portions of each protein in their co-crystal structures with DNA. The structural elements in each case are not drawn to scale and serve only to allow easy comparison of the overall topology. Loops and segments in blue represent the regions identified to interact with the DNA in each case.

linewidths, the absence of any peak-doubling throughout and electrophoretic mobility shift assays (EMSA) suggested a single, high affinity protein–DNA complex in solution that was in slow exchange on the NMR time scale (Figs 1*b*, 3). Nonetheless, the orientation of the RD on the DNA could not be unambiguously defined from the NMR data. The principal deficiency of the data was the paucity of intermolecular NOEs that could be observed between the protein and the DNA. In order to define the interface, a large number of NOEs must be visible in a ^{13}C-edited/^{13}C-filtered NOE experiment which define protein–DNA contacts across the recognition sequence of the DNA[9]. These NOEs form the basis for docking the protein domain on the DNA, enabling the construction of a preliminary model of the protein–DNA complex that can be utilized for further NOE analysis. The RD–DNA complex in this study displayed only 18 intermolecular NOEs between a single amino acid, the methyl groups of Val 74, and the base (H_5 and H_6) and sugar ($H_{1',2',3',4'}$) protons of two cytidine residues in the core recognition sequence (ATGCGGTTA·TAACCGCAT). Two additional NOEs were observed between the H_β and H_δ protons of Arg 80 and $H_{3'}$ of a third cytidine in the core sequence, ATGCGGTTA·TAACCGCAT, although not to other protons of this nucleotide. The core recognition sequence was defined by methylation and ethylation interference experiments[10]. Such a low number of interfacial contacts was deemed insufficient to define the orientation of the protein domain on the DNA. An additional 10–20 NOEs might be inferred from conventional ^{13}C-edited NOESY data, but their interpretation necessitates a preexisting model for the interaction that would be derived from a ^{13}C-edited/^{13}C-filtered NOE experiment. A further complication was the difficulty in defining the conformation of the C-terminus of the RD that had been shown to be essential for DNA binding. The deletion of one or more of these three C-terminal arginines (174, 177 and 178) abolishes DNA binding[11,12]. The segment of the protein containing these arginines, residues Tyr 169–Arg 178, did not display many NOEs to the rest of the protein domain, leaving the conformation of this segment poorly defined (Fig. 1*a*). In addition, this segment did not display any spectroscopic characteristics

that suggested the presence of regular secondary structure (Fig. 1*a*). Thus, a segment of the protein deemed essential for DNA-recognition based on biochemical experiments could not be defined structurally in this study. Comparison to the free RD in solution was not possible due to the low solubility of the protein (<100 μM) between pH 4–8 and/or in the presence of 0–500 mM NaCl. Resonance assignment of the DNA was complete for the base protons and for sugar protons with the exception of $H_{5'}/H_{5''}$. The DNA displayed no spectroscopic characteristics that would suggest a conformation that significantly differed from B-form helix (data not shown).

Despite these limitations, the interfacial NOEs observed, mutagenesis and comparison of the RD-fold to the Ig-domains of other factors (Fig. 2) suggests three distinct regions of the RD which should be involved in DNA binding. The first segment is formed by the loop between β_A and β_B (the $\beta_{A\text{-}B}$ loop, residues Val 74–Ile 87) which probably plays a principal role in core sequence recognition. Precisely which residues of this loop, in addition to Val 74 and Arg 80, are involved is unknown. The analogous segments in five other factors play an essential role in DNA recognition in the major groove (Fig. 2*b*), even though the details of the interaction are distinct in (nearly) each case[2-8]. The RD mutants R80C (ref. 12), C81D (ref. 13) and K83N (ref. 12) identify positions that are located at the bottom of this loop and fail to bind DNA *in vitro*. R80L and K83N were identified in patients with acute leukemia[12], reinforcing the notion that this segment is essential to the protein's DNA binding/gene activation function. The positions equivalent to Arg 80 in the DNA binding loops of NF-κB and NFAT1 have been shown to form direct contacts with the DNA[3-6].

A second loop of the Ig-motif is frequently involved in binding the DNA backbone and/or bases at the minor groove 3' to the major groove recognition sequence. This element is represented by, for example, segment S3 in the STATs[7-8] and by loop L2 in NFAT1 (refs 5, 6). The analogous loop in the RD is comprised of residues Arg 139-Ser 145 (the $\beta_{E'\text{-}F}$ loop) (Fig. 2*a*). Mutagenesis of Ser 140 to Gly (ref. 14) or Asn (ref. 15) and Leu 148 to Asp (ref. 16) substantially weaken the DNA-binding affinity of the RD *in vitro*; these mutants traverse the length of the $\beta_{E'\text{-}F}$ loop. *In vivo*, a patient with cleidocranial dysplasia was found to carry a Ser to Asn mutation at position 140 (ref.15), an observation consistent with the hypothesis that the $\beta_{E'\text{-}F}$ loop is important to the function of the RD. Chemical footprinting also suggests a role for minor groove contacts by the RD as adenine methylation interference downstream of GCGG disrupts DNA binding[10].

The third segment of the RD that has been implicated in DNA binding resides in the C-terminus of the domain. Deletion of one or more of three arginine residues at the C-terminus (positions 174, 177 or 178) results in a complete loss of DNA binding *in vitro*[11-12]. None of these residues displayed unambiguous NOEs to the DNA in a ^{13}C-edited/^{13}C-filtered NOE experiment. No conclusion can be drawn from this observation due to the general difficulty in defining the conformation of the C-terminus (Fig. 1*a*). Elements of the C-terminal segments of other Ig-domains have been shown to participate in major groove recognition with the $\beta_{A\text{-}B}$ loop in a variety of ways[2-8].

Heterodimerization

Functional characterization of the RD demonstrated that this domain serves as both a DNA binding domain and heterodimerization surface for the β-subunit[11]. The binding surface for β on the RD was mapped by direct analysis of a ternary β–RD–DNA complex (43.6 kDa) in which only the RD was ^{13}C and/or ^{15}N-enriched, following identification of the minimal heterodimerizing fragment of β (β_{141}) (Fig. 1*b*, Methods). A plot of the change in total ^{1}H and ^{15}N chemical shifts of the RD upon addition of β_{141} reveals the likely heterodimerization surface (Figs 3, 4). The region with the greatest changes in chemical shift, Met 106–Ser 114, is consistent with mutagenesis data in which the point mutants A107T and G108R were no longer able to heterodimerize with the β subunit *in vitro*, but retained the ability to bind DNA (Fig. 1*b*). The proximity of the RD–β interaction surface to the putative $\beta_{E'\text{-}F}$ DNA binding loop suggests a mechanism for DNA-binding enhancement by β (Fig. 4). Direct contact of β with a DNA binding loop of the RD could stabilize the binding loop's conformation, increasing the binding affinity of the RD for DNA. It may also be possible that β forms contacts with the phosphodiester backbone directly[17].

Both heterodimerization and cooperative interactions with neighboring transcription factors indicate that the DNA binding affinity of the RD can be substantially increased in a heteromeric nucleoprotein complex as compared to that observed for the RD alone[18–20]. These observations suggest that the DNA binding surface of the RD is pliable, with a singular conformation for the DNA contacting loops formed only in the functional context of

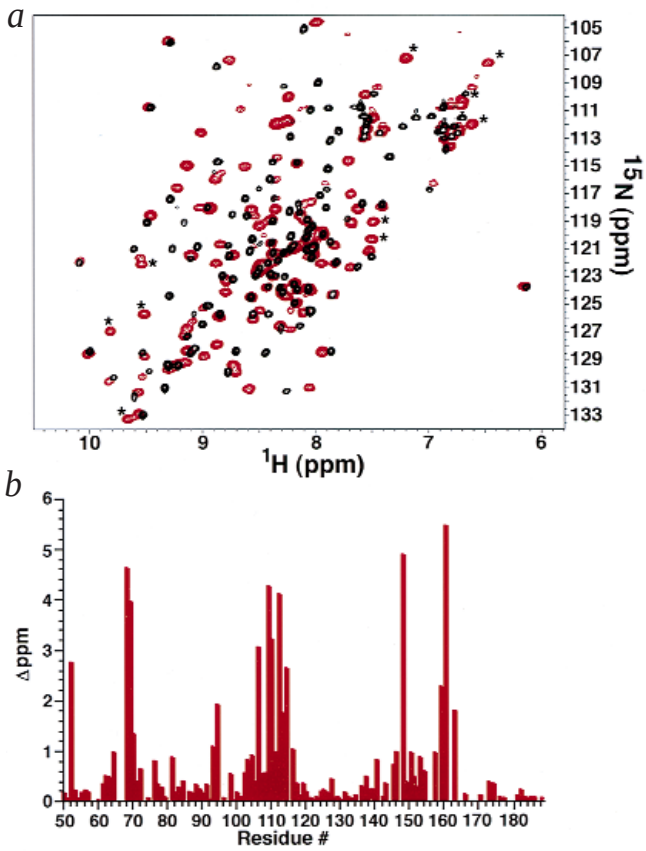

Fig. 3 Identification of the heterodimerization surface for PEBP2/CBFβ. ***a***, Shown in black is the ^{15}N-^{1}H HSQC spectrum of the RD in the 27 kDa RD–DNA complex and in red is RD in the 43.6 kDa β_{141}–RD–DNA ternary complex. Asterisks highlight highly shifted crosspeaks in the ternary complex spectrum. ***b***, A plot of the total change in ^{1}H and ^{15}N chemical shifts per residue indicates which residues of the RD are likely to be in contact with β_{141}. A clustering of chemical shift changes ≥1.0 p.p.m. occurs for residues Asp 66–Asn 71, Val 92–Gly 95, Val 105–Glu 116, Thr 147–Leu 148 and Gln158–His163. These residues are located in the $\beta_{O\text{-}A}$ loop, the $\beta_{B\text{-}C}$ loop, β_C-β_D, β_F and β_G, respectively.

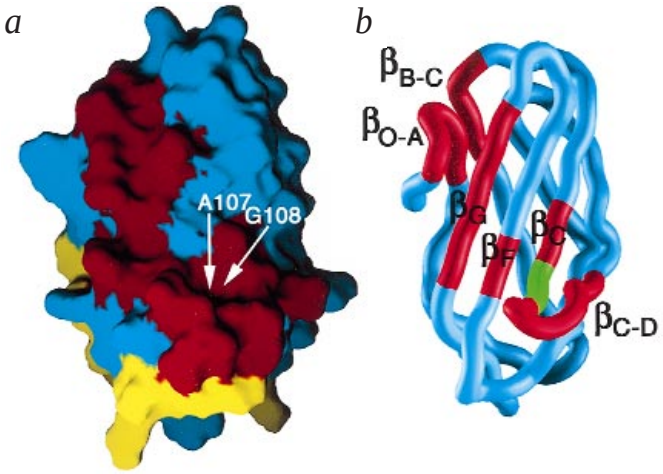

Fig. 4 Mapping of the β heterodimerization surface of the RD onto the three-dimensional structure. ***a***, Residues identified by chemical shift perturbation mapping are highlighted in red on the molecular surface[35] of the Runt domain (residues 62–170). Indicated residues are mutants that bind DNA, but fail to heterodimerize with β *in vitro* (see text). Yellow surfaces indicate the presumptive DNA-binding loops: the $\beta_{E'-F}$ loop (foreground) and the β_{A-B} loop (background). ***b***, A worm representation[35] of residues 62–170 illustrates the contiguous surface of interaction proposed (red): the β_{O-A} loop (Asp 66–Asn 71); a portion of the β_{B-C} loop (Val 92–Gly 95); portions of β_C, the β_{C-D} loop and portions of β_D (Val 105–Glu 116, Thr 147–Leu 148 and Gln 158–His 163, respectively). The heterodimerization mutants A107T and G108R are shown in green[14]. The orientation of the worm and surface is identical to that of Fig. 3*a*.

an oligomeric complex. The structurally related factor NFAT1 displayed similar conformational adaptability at the protein–DNA interface in solution, forming a short helical segment in one of the DNA recognition loops upon DNA binding[6]. In the ternary complex of NFAT1 with AP-1 (ref. 5), the Ig domain both rotates and tilts to form intimate contacts with the binding partner resulting in significant changes at the Ig domain–DNA interface. Moreover, the NFAT1 Ig domain binds DNA poorly in the absence of a binding partner, necessitating the introduction of a point mutant to enhance DNA binding[6]. All of these properties of NFAT1 are reminiscent of those described herein, suggesting that conformational adaptability is an important component of NFAT1 and Runt domain DNA binding and perhaps of an Ig DNA binding domain in general.

Implications for the role of the Runt domain in leukemogenesis

The RD is causally associated with the onset of the acute human leukemias[21]. Chromosomal translocations result in fusions of the RD gene to the ETO or TEL transcription factor genes, resulting in acute myeloblastic or acute pro-B cell lymphoblastic leukemia, respectively[21]. Thus, understanding the oligomeric state factor genes and heterodimerization mechanism of PEBP2/CBF and how these contribute to the regulation of gene expression could reveal new insights into the mechanism of leukemogenesis involving these fusion products. One intriguing question in this regard is the oligomeric state of PEBP2/CBF in the cytoplasm and how the protein is signaled to be translocated into the nucleus. Immunofluorescence labeling of PEBP2/CBF suggests that the β subunit localizes in the cytoplasm while full-length α is located principally in the nucleus[22–23]. Truncation of α at the C-terminus, however, co-localizes β with α in the nucleus[22]. It is possible, therefore, that the full-length α may exist in a conformation that does not permit its association with β. That the roles of the two subunits of PEBP2/CBF in co-localization are mutual, rather than one-sided, has been suggested by the observation that a leukemogenic fusion product resulting from chromosome 16 inversion, PEBP2/CBFβ-SMMHC, sequesters the α subunit with the β chimera in the cytoplasm[24–26]. The possibility that a β chimera, containing the heterodimerization surface for α, can preferentially co-localize α in the cytoplasm suggests an alternate mechanism by which the fusion products of PEBP2/CBF may promote the onset of the acute human leukemias. These chimeras may simply never get into the nucleus and/or they may heterodimerize with the wild-type gene products, leaving normal PEBP2/CBF unavailable for transcriptional activation. In support of this hypothesis are the results of *in vivo* co-localization experiments which demonstrate that when a portion of the RD-interaction domain of β is deleted in PEBP2/CBFβ-SMMHC, α translocates to the nucleus normally[25]. This demonstrates that cytoplasmic sequestration of PEBP2/CBF subunits may play a causative role in leukemogenic transformation in some cases[24–25].

Table 1 Structural statistics[1]

	<SA>
R.m.s. deviations from experimental distance restraints (Å)[2]	
All (1522)	0.048 ± 0.007
Sequential (li - jl = 1) (423)	0.036 ± 0.004
Short range (1 < li - jl < 5) (185)	0.057 ± 0.006
Long range (li - jl) > 5) (572)	0.051 ± 0.006
Intraresidue (294)	0.030 ± 0.004
Hydrogen bond (48)	0.105 ± 0.011
R.m.s. deviations from experimental	
Dihedral restraints (°) (159)	0.87 ± 0.12
$^3J_{NH\alpha}$ coupling constants (Hz) (72)	1.11 ± 0.05
Deviations from idealized covalent geometry	
Bonds (Å)	0.0052 ± 0.0003
Angles (°)	0.64 ± 0.02
Impropers (°)	0.62 ± 0.04
Coordinate precision[3]	
Backbone (residues 62–170)	0.59 ± 0.07
All non-hydrogen atoms (residues 62–170)	1.21 ± 0.09
Quality factors[4]	
% residues in most favorable Ramachandran (5,123)	78.0%
Prosa II Z score	-5.57 ± 0.25

[1]R.m.s. deviations are calculated relative to the mean coordinates <SA> calculated for the family of 47 simulated annealing structures excluding residues 179–188 which were completely disordered and displayed negative heteronuclear NOEs.
[2]No restraints between protons separated by three bonds were utilized (approximately 800 NOEs).
[3]The precision of the coordinates is defined as the average atomic r.m.s. difference between the 47 individual simulated annealing structures and the mean coordinates <SA> for residues 62–170.
[4]PROCHECK_NMR[36] and ProsaII[37] were used to assess the overall quality of the structures for residues 62–170.

Methods

Proteins and DNA. A 138 amino acid RD construct derived from cDNA for human AML1 (ref. 12) with a single cysteine to serine mutation[13] (residues 51–188, Cys 81→Ser) was overexpressed in a medium suitable for isotopic-enrichment in ^{13}C and ^{15}N and purified as described[11]. Full-length and truncated β subunit were subcloned and overexpressed in pET3a (Novagen) and purified by a combination of Q-Sepharose and Mono-Q chromatography. A 16 base pair oligonucleotide duplex was synthesized and purified using standard phosphoramidite chemistry (5'-GAGGATGCGGTTACTG-3'). Equimolar binary or ternary complexes were formed at 0.1–0.3 mg ml^{-1} of the labeled component in 50 mM Tris, 1 mM DTT, pH 8.0 at 400 mM NaCl and gradually exchanged by dialysis into a final buffer containing 10 mM sodium phosphate and 1 mM DTT-d_{10}, pH 7.2. The protein solution was concentrated by ultrafiltration to 1 mM for the NMR experiments.

Proteolysis. Trypsin and V8 were used to cleave PEBP2/CBFβ at protein:enzyme ratios of 19:1 and 4.7:1 (w/w) respectively for 30 min in a 20 μl reaction containing 25 mM Tris, 100 mM NaCl, 1 mM DTT, 50% (v/v) glycerol, pH 8.5. The reactions were quenched with either 1% trifluoroacetic acid, 250 ng leupeptin (trypsin) or 2 μg TLCK (V8). Electrospray mass spectrometry and N-terminal sequencing identified major fragments in the proteolysis mixtures as follows: 1–78 (trypsin), 1–139 or 1–144 (V8). The 141 amino acid fragment chosen (Fig.1*b*) is consistent with deletion analyses which suggested a truncated form of β retains full heterodimerization activity[18].

Electrophoretic mobility shift assays (EMSA). Truncated β proteins were incubated with pre-formed RD–DNA complex (12.5 nM RD and 1 nM ^{32}P-DNA) in 20 μl containing 20 mM HEPES-KOH, 2 mM EDTA, 1 mM DTT, 100 mM NaCl, 75 ng of sonicated salmon sperm DNA, 0.2 mg ml^{-1} BSA, 0.1% NP-40, pH 8.0. The resulting mixture was analyzed on 10% 0.25X TBE gels and electrophoresed at 10 mA and 4 °C for 2.5 h.

NMR spectroscopy. Data was collected at 36 °C on either a Bruker DMX500 or DMX600 NMR spectrometer equipped with a triple-resonance z-shielded gradient probe. The sequential assignment of the ^1H, ^{13}C and ^{15}N chemical shifts was achieved by standard techniques[9,27]. Four residues could not be identified (His 163–Ala 165 and Gly 172). NOEs involving protons of the protein were obtained from 3D ^{15}N-separated, 3D ^{13}C-separated and 4D $^{13}C/^{13}C$-separated NOE spectra (mixing times of 150, 75 and 120 ms, 75 ms and 180 ms, respectively) and a 3D ^{15}N-separated ROE spectrum recorded with a 30 ms mixing time and 6kHz spin-locking field.

The mapping of the heterodimerization surface of the RD for β_{141} was accomplished by measurement of the chemical shift changes in ^1H and ^{15}N chemical shifts of the RD from a single ^1H-^{15}N HSQC spectrum of the binary RD–DNA or ternary β_{141}–RD–DNA complex. Each spectrum was collected in 17 min with 4 scans per increment. The assignment of the backbone resonances of the RD in the ternary complex was achieved by comparison of free and bound ^{15}N-edited NOESY and sequential assignment with an HNCA experiment. Approximately 85% of the backbone resonances of the RD in the ternary complex were assigned. The changes in ^1H and ^{15}N chemical shifts between the binary and ternary RD spectra were analyzed as described[17].

Structure calculation. NOEs within the protein were grouped into four distance ranges as described[28]. Distances involving methyl protons, aromatic ring protons and non-stereospecifically assigned methylene protons were represented as a $(\Sigma r^{-6})^{-1/6}$ sum[29]. Protein backbone hydrogen bonding restraints ($r_{NH\text{-}O}$ = 1.5–2.8 Å, $r_{N\text{-}O}$ = 2.4–3.5 Å) within areas of regular secondary structure were introduced during the final stages of refinement. The minimum ranges employed for φ, ψ, χ_1 and χ_2 torsion angle restraints were ± 30°, ± 30°, ± 20° and ± 30°, respectively. The structures were calculated with the program X-PLOR-3.843[30] adapted to incorporate pseudo-potentials for $^3J_{NH\alpha}$ coupling constants[31] and a conformational database potential[32]. There were no hydrogen-bonding, electrostatic or 6-12 Lennard-Jones empirical potential energy terms in the target function.

Coordinates. The coordinates of the 47 structures have been deposited in the Protein Data Bank (accession code 1CMO).

Acknowledgments

The authors wish to acknowledge many stimulating discussions with L. Glaser, J. Hill, A. Kim, A. Seth and M. Osato. This work was supported in part by generous grants from the Sidney Kimmel Cancer Foundation, the Alexander and Alexandrine Sinsheimer Foundation and the New York Community Trust to M.H.W., a postdoctural fellowship to V.G. form the Leukemia Research Foundation and by grants from the Ministry of Education, Science, Culture and Sports of Japan to Y. I and K.S.

Correspondence should be addressed to M.H.W. *email: mwerner@portugal.rockefeller.edu*

Received 25 March, 1999; accepted 14 May, 1999.

1. Crute, B. E., Lewis, A. F., Wu, Z., Bushweller, J. H. & Speck, N. A. *J. Biol. Chem.* **271**, 26251–26260 (1996).
2. Cho, Y., Gorina, S., Jeffrey, P. D. & Pavletich, N. P. *Science* **265**, 346–355 (1994).
3. Ghosh, G., van Duyne, G., Ghosh, S. & Sigler, P. B. *Nature* **373**, 303–310 (1995).
4. Müller, C. W., Rey, F. A., Sodeoka, M., Verdine, G. L. & Harrison, S. C. *Nature* **373**, 311–317 (1995).
5. Chen, L., Glover, J. N. M., Hogan, P. G., Rao, A. & Harrison, S. C. *Nature* **392**, 42–48 (1998).
6. Zhou, P., Sun, L.-J. Dötsch, V., Wagner, G. & Verdine, G. L. *Cell* **92**, 687–696 (1998).
7. Chen, X. *et al. Cell* **93**, 827–839 (1998).
8. Becker, S., Groner, B. & Müller C. W. *Nature* **394**, 145–151 (1998).
9. Clore, G. M. & Gronenborn, A. M. *Protein Sci.* **3**, 372–390 (1994).
10. Kamachi, Y. *et al. J. Virol.* **64**, 4808–4819 (1990).
11. Kagoshima, H., Akamatsu, Y., Ito, Y. & Shigesada, K. *J. Biol. Chem.* **271**, 33074–33082 (1996).
12. Osato, M. *et al. Blood* **93**, 1817–1824 (1999).
13. Akamatsu, Y. *et al. J. Biol. Chem.* **272**, 14497–14500 (1997).
14. Akamatsu, Y., Tsukumo, S., Kagoshima, H., Tsurushita, N. & Shigesada, K. *Gene* **185**, 111–117 (1997).
15. Lee, B. *et al. Nature Genet.* **16**, 307–310 (1997).
16. Lenny, N., Meyers, S. & Hiebert, S. W. *Oncogene* **11**, 1761–1769 (1995).
17. Goger, M. *et al. Nature Struct. Biol.* **6**, 620–623 (1999).
18. Ogawa, E. *et al. Virol.* **194**, 314–331 (1993)
19. Wang, S. *et al. Mol. Cell. Biol.* **13**, 3324–3339. (1993)
20. Kim, W-Y. *et al. EMBO J.* **18**, 1609–1620 (1999).
21. Look, A.T. *Science* **278**, 1059–1064 (1997).
22. Lu, J. *et al. Mol. Cell. Biol.* **15**, 1651–1661 (1995).
23. Chiba, N. *et al. Oncogene* **14**, 2543–2552 (1997).
24. Kanno, Y., Kanno, T., Sakakura, C., Bae, S-C., & Ito Y. *Mol. Cell. Biol.* **18**, 4252–4261 (1998).
25. Adya, N., Stacy, T., Speck, N. A. & Liu, P. P. *Mol. Cell. Biol.* **18,** 7432–7443 (1998).
26. Liu, P. *et al. Cold Spring Harbor Symp. Quant. Biol.* **59**, 547–553 (1994).
27. Bax, A. & Grzesiek, S. *Acc. Chem. Res.* **26**, 131–138 (1993).
28. Omichinski, J. G., Pedone, P. V., Felsenfeld, G., Gronenborn, A. M. & Clore, G. M. *Nature Struct. Biol.* **4**, 122–132 (1997).
29. Nilges, M. *Prot. Struct. Funct. Genet.* ***17***, 295–309 (1993).
30. Brünger, A. T. *X-PLOR Manual, Version 3.1* (Yale University Press, New Haven, Connecticut; 1992).
31. Garrett, D. S. *et al. J. Magn Reson. (B)* **104**, 99–103 (1994).
32. Kuszewski, J., Gronenborn, A. M. & Clore, G. M. *J. Magn. Reson.* **125**, 171–177 (1997).
33. Bork, P., Holm, L. & Sander, C. *J. Mol. Biol.* **242**, 309–320 (1994).
34. Carson, M. *J. Mol. Graphics* **5**, 103–106 (1987).
35. Nicholls, A., Sharp, K. & Honig, B. *Proteins* **11,** 281–296. (1991)
36. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. & Thornton, J. M. *J. Biomol. NMR* **8**, 477–486. (1996)
37. Sippl, M. J. *Proteins* **17**, 355–362. (1993)