J. Mol. Biol. (1997) **273**, 19–25

JMB

COMMUNICATION

Solution Structure of the Iγ Subdomain of the Mu End DNA-binding Domain of Phage Mu Transposase

Robert T. Clubb[1], Silke Schumacher[1], Kiyoshi Mizuuchi[2] Angela M. Gronenborn[1]* and G. Marius Clore[1]*

[1]*Laboratories of Chemical Physics and* [2]*Molecular Biology, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases National Institutes of Health Bethesda, MD 20892-0520 USA*

**Corresponding authors*

The MuA transposase of phase Mu is a large modular protein that plays a central role in transposition. We show that the Mu end DNA-binding domain, Iβγ, which is responsible for binding the DNA attachment sites at each end of the Mu genome, comprises two subdomains, Iβ and Iγ, that are structurally autonomous and do not interact with each other in the absence of DNA. The solution structure of the Iγ subdomain has been determined by multidimensional NMR spectroscopy. The structure of Iγ comprises a four helix bundle and, despite the absence of any significant sequence identity, the topology of the first three helices is very similar to that of the homeodomain family of helix-turn-helix DNA-binding proteins. The helix-turn-helix motif of Iγ, however, differs from that of the homeodomains in so far as the loop is longer and the second helix is shorter, reminiscent of that in the POU-specific domain.

Keywords: phage Mu; transposase; Iγ subdomain; solution structure; helix-turn-helix

The MuA protein of phage Mu plays a key role in transposition, a recombination reaction that moves a genetic element, the transposon, from one location on the chromosome to a new and often distantly located site. Specifically, MuA pairs the ends of the phage DNA, cleaves the donor and target DNA and promotes strand transfer (Symonds *et al.*, 1987; Mizuuchi, 1992; Lavoie & Chaconas, 1995). MuA has a total molecular mass of 75 kDa and consists of five functional domains (Figure 1(a)). Domain Iα enhances stable synaptic complex formation by binding to the internal activation sequence (IAS); the Mu end DNA binding domain Iβγ is required for transposition and binds to the L and R attachment (*att*) sites, contacting two consecutive major grooves located on one face of an asymmetric 22 base-pair site and inducing a bend in the DNA of 60 to 90° (Kuo *et al.*, 1991; Zou *et al.*, 1991); the catalytic domain II participates in the catalysis of both cleavage and strand transfer reactions; domain IIIα is required for the assembly of an active MuA-DNA complex and has endonuclease activity; and finally domain IIIβ mediates the interaction between MuA and MuB protein, an ATP-dependent target DNA activator (Lavoie & Chaconas, 1995). The three-dimensional structures of the IAS-binding domain Iα and the catalytic domain II have been solved by NMR (Clubb *et al.*, 1994) and crystallography (Rice & Mizuuchi, 1995). Domain Iα is a member of the winged helix-turn-helix family of DNA-binding proteins, albeit with a permutation of the secondary structure elements. The catalytic domain II is structurally similar to the catalytic cores of HIV and ASV integrase, ribonuclease H and the Holliday-junction resolving enzyme RuvC. Here, we present the solution structure of the Iγ subdomain by NMR spectroscopy.

Limited trypsin proteolysis of the Iβγ domain (>95% pure) results in the appearance of a single band at ~9.4 kDa on SDS-PAGE. N-terminal sequencing of this band reveals the presence of two co-migrating peptides: an N-terminal fragment (Iβ) beginning at Ile77, and a C-terminal fragment

Present address: R. T. Clubb, Department of Chemistry and Biochemistry, University of California, Los Angeles at Los Angeles, CA 90095-1570, USA

Abbreviations used: IAS, internal activation sequence; ASV, ; HSQC, heteronuclear single quantum coherence; HTH, helix-turn-helix; ppm, parts per million; ROE, rotating frame Overhauser enhancement; NOE, nuclear Overhauser enhancement.

0022–2836/97/410019–07 $25.00/0/mb971312

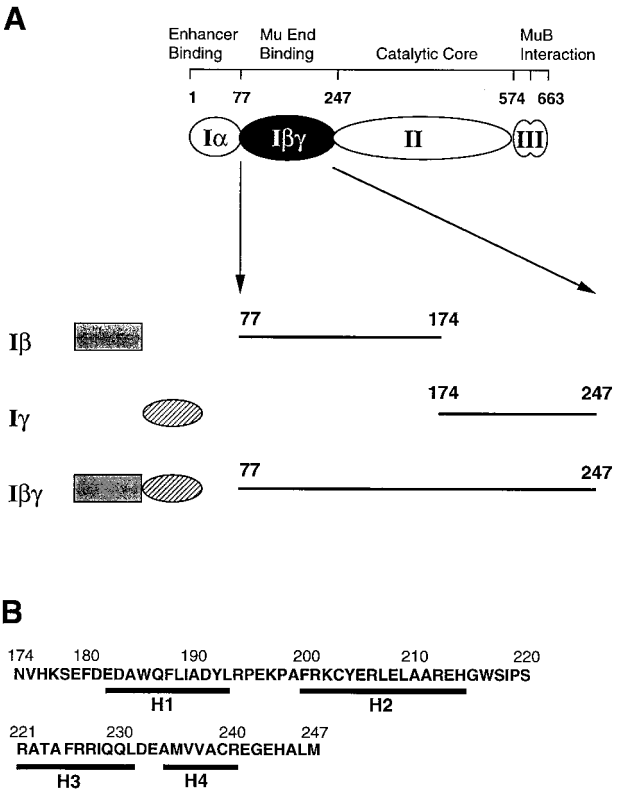

Figure 1. a, A drawing of the domain organization of MuA transposase, and of the various fragments studied by NMR. b, Sequence of Iγ (residues 174 to 247) with the location of the four helices determined by NMR indicated. MuA deletion variants were constructed using PCR methodology. Plasmid pMK609 (Mizuuchi & Mizuuchi, 1989), which contains DNA encoding the MuA transposase, was targeted for deletion. Appropriate restriction sites were engineered into the PCR primers allowing selectively amplified DNAs to be ligated into the phage T7 RNA polymerase expression plasmid pET3c (Novagen). The expression plasmids were then transferred to *E. coli* strain BL21(DE3) for protein expression. Three deletion constructs of the MuA transposase were constructed and comprise residues 77 to 247 (Iβγ), residues 77 to 174 (Iβ), and residues 174 to 247 (Iγ) of MuA transposase.

(Iγ) starting at Asn174 (data not shown). The Iβ and Iγ domains share no significant sequence identity and differ in length (98 residues for Iβ *versus* 74 for Iγ; Kim & Harshey, 1995). Iβγ, Iβ and Iγ were overexpressed in *Escherichia coli*, uniformly labeled with ^{15}N and purified to homogeneity for NMR analysis. ^{1}H-^{15}N HSQC spectra of Iβ, Iγ and Iβγ exhibit excellent spectral dispersion indicative of folded proteins (Figure 2a, b and d). Interestingly, in the absence of DNA, the ^{1}H-^{15}N HSQC spectrum of the intact Iβγ domain provides no evidence for any interaction between the Iβ and Iγ subdomains, since the sum of the ^{1}H-^{15}N HSQC spectra of the individual subdomains (Figure 2c) is virtually identical with the spectrum of the full-length domain construct (Figure 2d).

The three-dimensional structure of the Iγ subdomain (residues 174 to 247) was determined using multidimensional (3D and 4D) double and triple resonance NMR spectroscopy, making use of uniformly (>95%) ^{15}N and $^{15}N/^{13}C$ labeled samples (for details of the experiments and the original references, see the reviews by Clore & Gronenborn, 1991; Bax & Grzesiek, 1993; Bax *et al.*, 1994). The structure was solved on the basis of 1293 experimental restraints. A superposition of the final ensemble of 30 simulated annealing structures is shown in Figure 3 and a summary of the structural statistics is provided in Table 1. Residues 173 to 179 at the N terminus and residues 241 to 247 at the C terminus of the Iγ subdomain are disordered. The remainder of the Iγ subdomain (residues 180 to 240) is well defined with a precision of 0.49(±0.09) Å for the backbone atoms and 1.18(±0.11) Å for all atoms.

The Iγ subdomain comprises a four helix bundle tightly packed around a hydrophobic core consisting of aliphatic and aromatic amino acid residues (Figure 3). Helices 1 (residues 182 to 193) and 2 (residues 200 to 214) are oriented antiparallel to each other at an angle of ~150°. Helix 3 (residues 221-231) crosses helices 1 and 2 at angles of ~60° and ~120°, respectively. In the view shown in Figure 3, the top of the hydrophobic core is closed off by helix 4 (residues 234 to 240), which is oriented approximately orthogonal to the other three helices. The secondary ^{13}C shifts are indicative of the presence of an N-terminal capping box for helix 3 with the C^{β} and C^{α} of Ser220 shifted 1.84 ppm downfield and 0.16 ppm upfield, respectively (Gronenborn & Clore, 1994). This is confirmed in the structure, which shows that the hydroxyl group of Ser220 is hydrogen bonded to the backbone amide group of Thr223.

Excluding the C-terminal helix 4, the fold of the Iγ subdomain is remarkably similar to that of the homeodomain family of helix-turn-helix (HTH) DNA-binding proteins, although their amino acid sequences are completely unrelated (sequence identities <10%). Thus, for example, helices 1 to 3 of Iγ can be superimposed onto helices 1 to 3 of the Oct-1 POU homeodomain (Klemm *et al.*, 1994) with a C^{α} atomic rms difference of 2 Å for 41 residues (Figure 4). In the homeodomain family of proteins, helices 2 and 3, separated by a four residue turn, constitute the HTH motif (Kissinger *et al.*, 1990; Klemm *et al.*, 1994). The HTH motif consists of two helices that cross each other at an angle of ~120°; major groove recognition is achieved by residues in the second helix of the HTH motif and in the loop connecting the two helices. By analogy, polar side-chains in the recognition helix of the HTH motif of Iγ that could potentially interact with the DNA major groove include Arg221, Arg226, Arg227, Gln229 and Gln230. Interestingly, Arg226, Gln229 and Gln230 are located at analogous positions to residues that form hydrogen

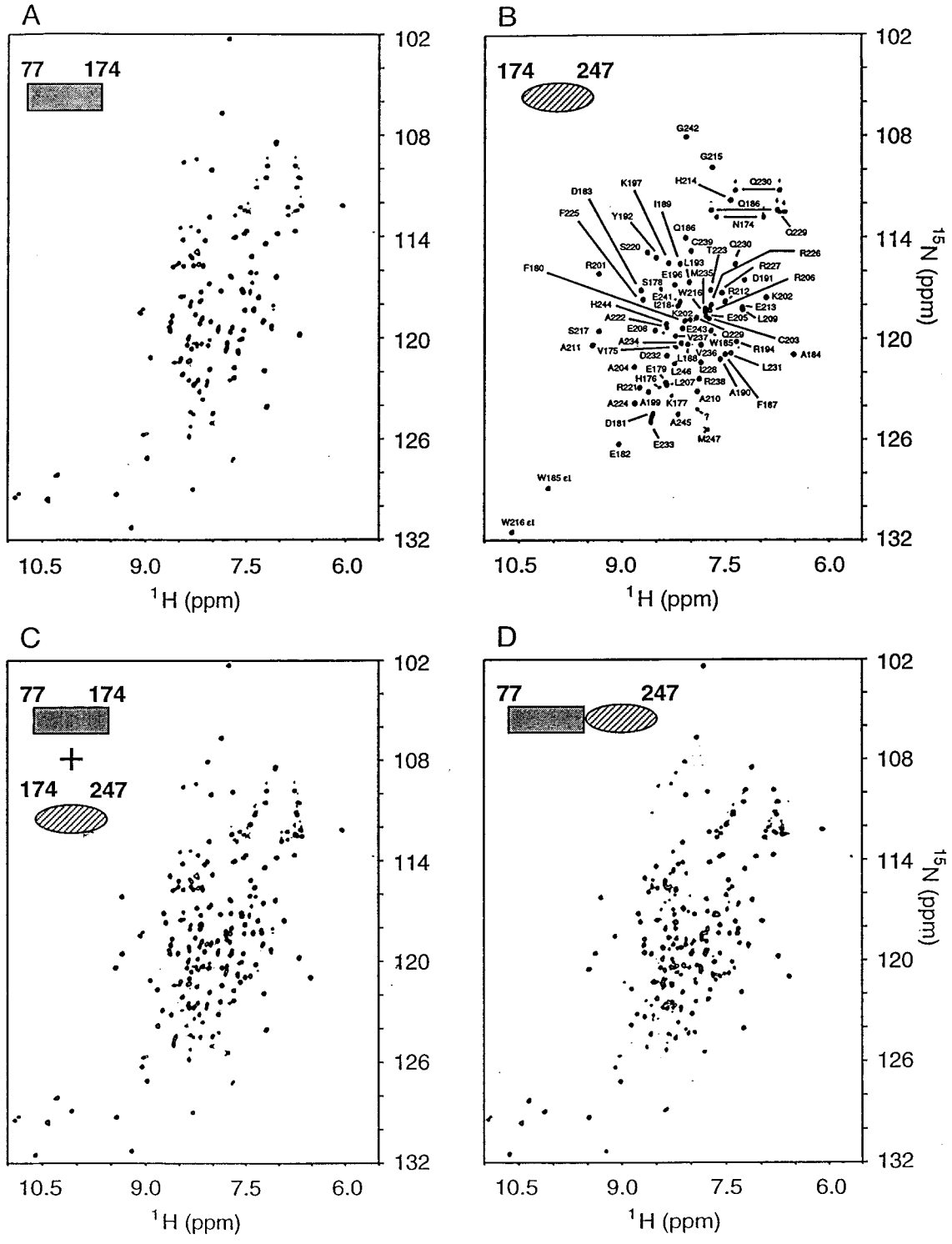

Figure 2. ^{1}H-^{15}N HSQC spectra of various fragments of the Mu end DNA binding domain Iβγ. a, Iβ (residues 77 to 174), b, Iγ (residues 174 to 247), c, the sum of the individual Iβ and Iγ spectra shown in a, and b; and d, Iβγ (residues 77 to 247). The proteins were uniformly labeled with either ^{15}N or ^{15}N/^{13}C by growing the bacteria on minimal medium containing $^{15}NH_4Cl$ and/or $^{13}C_6$-glucose as the sole nitrogen and carbon sources, respectively. Purification was similar to that described for Iα (Clubb *et al.*, 1994). Each NMR sample contained 1 mM protein, 50 mM phosphate buffer (pH 6.3), 100 mM NaCl and 2.4 mM deuterated dithiothreitol. All NMR spectra were recorded at 30°C on Bruker AMX500 or AMX600 spectrometers equipped with a *z*-shielded gradient triple resonance probe.

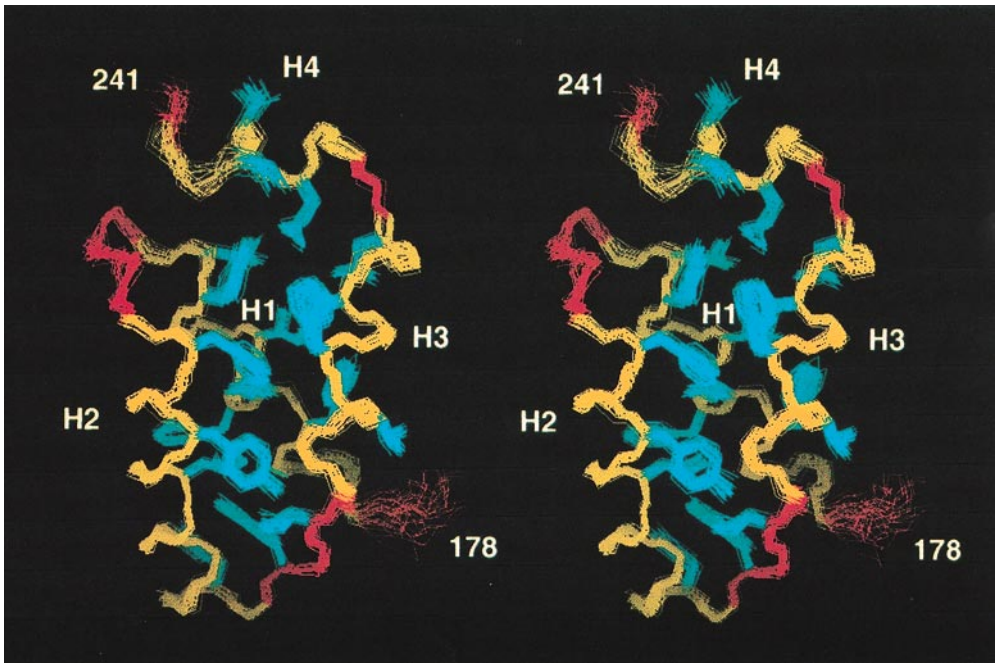

Figure 3. Stereoview showing a superposition of the final 30 simulated annealing structures of the Iγ subdomain. The backbone is displayed in yellow for the helices and red for the remaining residues, and side-chains are shown in blue. Complete ^{1}H, ^{15}N and ^{13}C assignments were obtained using double and triple resonance 3D NMR spectroscopy (Clore & Gronenborn, 1991; Bax & Grzesiek, 1993). These include 3D CBCANH, CBCA(CO)NH, HBHA(CO)NH, C(CO)NH, H(CCO)NH, HCCH-COSY, HCCH-TOCSY, HNHA and ^{15}N-separated HOHAHA experiments. Approximate interproton distance restraints, classified into four distance ranges, 1.8 to 2.7 (1.8 to 2.8 for distance pairs involving amide groups), 1.8 to 3.3 (1.8 to 3.5 for distance pairs involving amide groups), 1.8 to 5.0 and 1.8 to 6.0 Å, corresponding to strong, medium, weak and very weak nuclear Overhauser enhancements (NOEs) were obtained from 3D ^{15}N-separated NOE (70 and 100 ms mixing time) and ROE (45 ms mixing time), 3D ^{13}C-separated NOE (120 ms mixing time) and ROE (25 ms mixing time), 4D $^{15}N/^{13}C$ (100 ms mixing time) and $^{13}C/^{13}C$ separated (100 ms mixing time) NOE spectra. Three-bond coupling constants were obtained by quantitative J correlation spectroscopy (Bax *et al*., 1994). NMR spectra were processed using NMRPipe (Delaglio *et al*., 1995) and analysed using the programs PIPP, CAPP and STAPP (Garrett *et al*., 1991). Stereospecific assignments and ϕ and χ_1 torsion angle restraints were obtained using the conformational grid search program STEREOSEARCH (Nilges *et al*., 1990) on the basis of $^{3}J_{HN\alpha}$ and $^{3}J_{\alpha\beta}$ coupling constants, and intraresidue and sequential interresidue ROEs involving NH, $C^{\alpha}H$ and $C^{\beta}H$ protons. Information from $^{3}J_{NH\beta}$, $^{3}J_{COH\beta}$ coupling constants was also employed for identifying the appropriate χ_1 rotamer and for detecting rotamer averaging. In addition, in the case of threonine and valine residues $^{3}J_{C\gamma N}$ and $^{3}J_{C\gamma CO}$ coupling constants were used for determining the side-chain conformation. χ_2 torsion angle restraints for leucine residues were obtained from $^{3}J_{CC}$ coupling constants and the pattern of intraresidue NOEs (Powers *et al*., 1993).

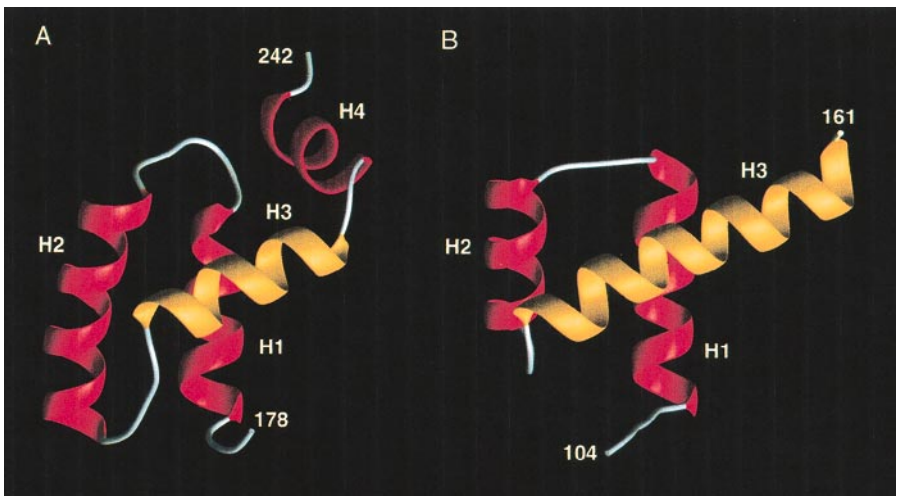

Figure 4. Comparison of the Iγ domain (a) with the Oct-1 POU homeodomain (b). The two structures can be superimposed with a C^{α} atomic rms difference of 2 Å for 41 residues (residues 179 to 194, 199 to 212 and 223 to 233 of Iγ and residues 107 to 122, 126 to 139 and 147 to 157 of the Oct-1 POU homeodomain) comprising principally helices 1, 2 and 3. Helices 1, 2 and 3 of Iγ comprise residues 182 to 192, 200 to 214 and 221 to 231, respectively, while in the recognition helix of the POU homedomain these three helices extend from residues 110 to 122, 128 to 137 and 142 to 160, respectively. The HTH motif is shown in yellow, and the other helices in red. The coordinates of the Oct-1 POU homeodomain are taken from Klemm *et al*. (1994). The Figure was generated with the program MOLMOL (Konradi *et al*., 1996).

bonds to DNA bases in the homeodomains, while Arg221 and Arg226 correspond to residues that form hydrogen bonds to the DNA bases in a large number of prokaryotic HTH proteins (Pabo & Sauer, 1992).

The HTH motif of Iγ is a variant of the classical HTH motif in so far that the two helices are connected by a six residue loop rather than the four residue turn seen in both the homeodomains and most prokaryotic HTH proteins (Pabo & Sauer,

Table 1. Structural statistics

	⟨SA⟩	($\overline{SA}$)r
Structural statistics		
RMS deviations from exptl distance restraints (Å)[a]		
All (1009)	0.027 ± 0.003	0.028
Interresidue sequential ($\|i-j\|=1$) (264)[b]	0.013 ± 0.007	0.008
Interresidue short range ($1 < \|i-j\| \leqslant 5$) (282)	0.042 ± 0.04	0.047
Interresidue long range ($\|i-j\| > 5$) (178)	0.030 ± 0.006	0.025
Intraresidue (245)	0.009 ± 0.006	0.003
H-bonds (40)	0.003 ± 0.004	0
RMS deviations from exptl		
Dihedral restraints (deg) (108)[a]	0.274 ± 0.099	0.161
$^3J_{HN\alpha}$ coupling constants (Hz) (47)[a]	0.78 ± 0.08	0.79
RMS deviations from exptl ^{13}C shifts		
$^{13}C^{\alpha}$ (ppm) (66)	1.20 ± 0.06	1.11
$^{13}C^{\beta}$ (ppm) (63)	0.90 ± 0.05	0.92
Deviations from idealized covalent geometry		
Bonds (Å) (1240)	0.004 ± 0.0005	0.005
Angles (deg) (2228)	0.518 ± 0.027	0.764
Impropers (deg) (676)	0.597 ± 0.067	1.108
Measures of structure quality		
$E_{L\text{-}J}$ (kcal.mol^{-1})[c]	-321 ± 10	−282
PROCHECK[d]		
Residues in most favorable region of Ramachandran plot (%)	93.0 ± 1.9	90.9
Number of bad contacts/100 residues	6.7 ± 2.9	3.3
H-bond energy	0.73 ± 0.07	0.80
Coordinate precision[e]		
Backbone (Å)	0.49 ± 0.09	
All atoms (Å)	1.18 ± 0.11	

The structures were calculated by simulated annealing (Nilges *et al.*, 1988) using the program XPLOR-3.1 (Brünger, 1993) modified to incorporate pseudo-potentials for coupling constant (Garrett *et al.*, 1994) and secondary ^{13}C chemical shift (Kuszewski *et al.*, 1995) restraints, and a conformational database potential (Kuszewski *et al.*, 1996, 1997). The latter biases sampling during simulated annealing refinement to conformations that are likely to be energetically possible by effectively limiting the choices of dihedral angles to those that are known to be physically realizable. The notation of the NMR structures is as follows: ⟨SA⟩ are the final 30 simulated annealing structures; $\overline{SA}$ is the mean structure obtained by averaging the coordinates of the individual SA structures (residues 180 to 240) best fit to each other; ($\overline{SA}$)r is the restrained minimized mean structure obtained by restrained regularization of the mean structure $\overline{SA}$. The number of terms for the various restraints is given in parentheses. The final force constants employed for the various terms in the target function used for simulated annealing are as follows: 1000 kcal mol^{-1} Å^{-2} for bond lengths, 500 kcal mol^{-1} rad^{-2} for angles and improper torsions (which serve to maintain planarity and chirality), 4 kcal mol^{-1} Å^{-4} for the quartic van der Waals repulsion term (with the hard sphere effective van der Waals radii set to 0.8 times their value used in the CHARMM PARAM19/20 parameters), 30 kcal mol^{-1} Å^{-2} for the experimental distance restraints (interproton distances and hydrogen bonds), 200 kcal mol^{-1} rad^{-2} for the torsion angle restraints, 1 kcal mol^{-1} Hz^{-2} for the coupling constant restraints, 0.5 kcal mol^{-1} ppm^{-2} for the carbon chemical shift restraints, and 1.0 for the conformational database potential. The terms for covalent geometry, coupling constants and ^{13}C shifts are described by harmonic potentials, while the interproton distance and torsion angle restraints are represented by square-well potentials.

[a] None of the structures exhibited distance violations greater than 0.5 Å, dihedral angle violations greater than 5°, or $^3J_{HN\alpha}$ coupling constant violations greater than 2 Hz. The torsion angles restraints comprise 68 φ, 27 χ_1 and 13 χ_2 angles. (The latter include the χ_2 of the four Phe and two Tyr residues that were restrained to 90(±30)°, in addition to the χ_2 angles of all three Ile and four out of the five Leu.). The hydrogen bonding restraints, which consist of two distances per backbone-backbone hydrogen bond, were included only in the final stages of refinement using standard criteria based on amide exchange, $^3J_{HN\alpha}$ couplings and secondary ^{13}C shifts.

[b] Only structurally useful intraresidue NOEs are included in the intraresidue interproton distance restraints. Thus, intraresidue NOEs between protons separated by two bonds or between non-stereospecifically assigned protons separated by three bonds are not incorporated in the restraints.

[c] $E_{L\text{-}J}$ is the Lennard-Jones van der Waals energy calculated with the CHARMM PARAM19/20 protein parameters (Brooks *et al.*, 1993) and is not included in the target function for simulated annealing or restrained minimization.

[d] The programs PROCHECK (Laskowski *et al.*, 1993) was used to assess the overall quality of the structures. More than 85% of residues in the most favourable region of the Ramachandran plot, less than ten bad contacts per 100 residues, and a hydrogen bond energy of 0.6 to 1.0 are expected for a good quality structure. The dihedral angle *G*-factors (which should be greater than −0.5 for a good quality structure) for the φ/ψ, χ_1/χ_2, χ_1 and χ_3/χ_4 distributions are 0.56 ± 0.07, 0.19 ± 0.10, 0.06 ± 0.30 and -0.17 ± 0.25, respectively. The PROCHECK statistics apply to the ordered region of Iγ comprising residues 180 to 240.

[e] The precision of the atomic coordinates is defined as the average rms difference between the 30 final simulated annealing structures and the mean coordinates, $\overline{SA}$. The values given relate to residues 180 to 240, since the N (residues 173 to 179) and C (residues 241 to 247) termini are disordered. The values given for the backbone atoms relate to the N, C^{α}, C and O atoms; those given for all atoms refer only to non-hydrogen atoms.

1992). Thus the longer turn in the HTH motif of Iγ is reminiscent of that seen in the POU-specific domain (Klemm *et al*., 1994; Wright, 1994; Dekker *et al*., 1993).

Acknowledgments

We thank J. Boyes, J. Huth and H. Savilahti for useful discussions; G. Poy and R. Tschudin for technical support; F. Delaglio and D. S. Garrett for software support. This work was supported by a Leukemia Society of America post-doctoral fellowship (to R. T. C.), a Deutscher Akademischer Austauschdienst predoctoral fellowship (to S. S), and the AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (to G. M. C., A. M. G. and K. M.). The coordinates of the 30 simulated annealing structures and the restrained regularized mean structure of Iγ, together with the experimental restraints and ^{1}H, ^{15}N and ^{13}C resonance assignments have been deposited in the Brookhaven Protein Data Bank (accession codes 2EZH, 2EZI and 2EZHMR).

References

Bax, A. & Grzesiek, S. (1993). Methodological advances in protein NMR. *Acc. Chem. Res.* **26**, 131–138.

Bax, A., Vuister, G. W., Grzesiek, S., Delaglio, F., Wang, A. C., Tschudin, R. & Zhu, G. (1994). Measurement of homo- and heteronuclear J couplings from quantitative J correlation. *Methods Enzymol.* **239**, 79–105.

Brennan, R. G. (1993). The winged-helix DNA-binding motif: another helix-turn-helix takeoff. *Cell,* **74**, 773–776.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1993). CHARMM: a program for macromolecuar energy minimizaion and dynamics calculations. *J. Comput. Chem.* **4**, 187–217.

Brünger, A. T. (1993). *XPLOR 3.1: A System for X-ray Crystallography and NMR,* Yale University Press, New Haven, CT.

Clore, G. M. & Gronenborn, A. M. (1991). Structures of larger proteins in solution: three- and four-dimensional heteronuclear NMR spectroscopy. *Science,* **252**, 1390–1399.

Clubb, R. T., Omichinski, J. G., Savilahti, H., Mizuuchi, K., Gronenborn, A. M. & Clore, G. M. (1994). A novel class of winged helix-turn-helix protein: the DNA binding domain of Mu transposase. *Structure,* **2**, 1041–1048.

Dekker, N., Cox, M., Boelens, R., Verrijzer, C. P., van der Vliet, P. C. & Kaptein, R. (1993). Solution structure of the POU-specific DNA binding domain of Oct-1. *Nature,* **362**, 852–855.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR,* **6**, 277–293.

Garrett, D. S., Powers, R., Gronenborn, A. M. & Clore, G. M. (1991). A common sense approach to peak picking in two-, three- and four-dimensional spectra using automatic computer analysis of contour diagrams. *J. Magn. Reson.* **95**, 214–220.

Garrett, D. S., Kuszewski, J., Hancock, T. J., Lodi, P. J., Vuister, G. W., Gronenborn, A. M. & Clore, G. M. (1994). The impact of direct refinement against three-bond NH-CαH coupling constants on protein structure determination by NMR. *J. Magn. Reson. ser B,* **104**, 99–103.

Gronenborn, A. M. & Clore, G. M. (1994). Identification of N-terminal helix capping boxes by means of ^{13}C chemical shifts. *J. Biomol. NMR,* **4**, 455–458.

Kim, K. & Harshey, R. M. (1995). Mutational analysis of the att DNA-binding domain of phage Mu transposase. *Nucl. Acids Res.* **23**, 3937–3943.

Kissinger, C. R., Liu, B., Martin-Blanco, E., Kornberg, T. B. & Pabo, C. O. (1990). Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions. *Cell,* **63,** pp. 579–590.

Klemm, J. D., Rould, M. A., Aurora, R., Herr, W. & Pabo, C. O. (1994). Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell,* **77**, 21–32.

Konradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 52–55.

Kuo, C. -F., Zou, A., Jayaram, M., Getzoff, E. & Harshey, R. M. (1991). DNA-protein complexes during attachment-site synapsis in Mu DNA transposition. *EMBO J.* **10**, 1585–1591.

Kuszewski, J., Qin, J., Gronenborn, A. M. & Clore, G. M. (1995). The impact of direct refinement against $^{13}C\alpha$ and $^{13}C\beta$ chemical shifts on protein structure determination by NMR. *J. Magn. Reson. ser. B,* **106**, 92–96.

Kuszewski, J., Gronenborn, A. M. & Clore, G. M. (1996). Improving the quality of NMR and crystallographic protein structures by means of a conformational database potential derived from structure databases. *Protein Sci.* **5**, 1067–1080.

Kuszewski, J., Gronenborn, A. M. & Clore, G. M. (1997). Improvements and extensions in the conformational database potential for the refinement of NMR and X-ray structures of proteins and nucleic acids. *J. Magn. Reson.* **125**, 171–177.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.

Lavoie, B. D. & Chaconas, G. (1995). Transposition of phage Mu DNA. *Curr. Top. Microbiol. Immunol.* **204**, 83–102.

Mizuuchi, K. (1992). Transpositional recombination: mechanistic insights from studies of Mu and other elements. *Annu. Rev. Biochem.* **61**, 1011–1051.

Mizuuchi, M. & Mizuuchi, K. (1989). Efficient Mu transposition requires interaction of transposase with a DNA sequence at the Mu operator: implication for regulation. *Cell,* **58**, 399–408.

Nilges, M., Clore, G. M. & Gronenborn, A. M. (1988). Determination of three-dimensional structures of proteins from interproton distance data by hybrid distance geometry-dynamical simulated annealing calculations. *FEBS Letters,* **229**, 317–324.

Nilges, M., Clore, G. M. & Gronenborn, A. M. (1990). ^{1}H-NMR stereospecific assignments by conformational database searches. *Biopolymers,* **29**, 813–822.

Pabo, C. O. & Sauer, R. T. (1992). Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Biochem.* **61**, 1053–1095.

Powers, R., Garrett, D. S., March, C. J., Frieden, E. A., Gronenborn, A. M. & Clore, G. M. (1993). The high resolution three-dimensional solution structure of human interleukin-4 determined by multidimensional heteronuclear magnetic resonance spectroscopy. *Biochemistry,* **32**, 6744–6762.

Rice, P. A. & Mizuuchi, K. (1995). Structure of the bacteriophage Mu transposase core: a common structural motif for DNA transposition and retroviral integration. *Cell,* **82**, 209–220.

Symonds, N., Toussaint, A., van de Putte, P. & Howe, M. M. (1987). *Phage Mu,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Wright, P. E. (1994). POU domains and homeodomains. *Curr. Op. Struct. Biol.* **4**, 22–27.

Zou, A., Leung, P. C. & Harshey, R. M. (1991). Transposase contacts with Mu DNA ends. *J. Biol. Chem.* **266**, 20476–20482.

Edited by P. E. Wright

(Received 30 May 1997; received in revised form 24 July 1997; accepted 29 July 1997)