SUPPLEMENTARY MATERIAL

Structure and dynamics of KH domains from FBP bound to single stranded DNA

Demetrios T. Braddock[*,†] , John M. Louis[*], James L. Baber[*], David Levens[†] & G. Marius Clore[*]

[*]*Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0501, U.S.A.*

[†]*Laboratory of Pathology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, U.S.A.*

A. Structural Statistics

Table Structural statistics

	$<SA_{KH3}>$	$<SA_{KH4}>$
R.m.s. deviation from experimental restraints		
distances (Å) (1095/949)	0.05±0.00	0.04±0.00
torsion angles (°) (244/261)	0.40±0.05	0.46±0.05
$^{3}J_{HN\alpha}$ couplings (Hz) (33/36)	1.04±0.03	1.26±0.05
$^{13}C\alpha/\beta$ chemical shifts (Hz) (120/121)	1.14±0.03	1.06±0.05
Dipolar coupling R-factors* (%)		
$^{1}D_{NH}$ (61/61)	8.8±0.4	3.9±0.2
$^{1}D_{NC'}$ (47/39)	38.6±0.6	22.9±0.5
$^{2}D_{HNC'}$ (46/40)	35.6±0.7	23.1±1.1
Deviations from idealized covalent geometry		
bonds (Å)	0.003±0	0.003±0
angles (°)	0.57±0.01	0.59±0.01
Impropers (°)	0.59±0.07	0.54±0.09
% residues in most favorable region of Ramachandran map	96.0±0.9	90.0±2.4
Coordinate precision (Å)†		
protein backbone + DNA heavy atoms	0.30±0.05	0.38±0.08
protein heavy atoms + DNA heavy atoms	0.64±0.05	0.70±0.05

$<SA_{KH3}>$ and $<SA_{KH4}>$ are the final 80 simulated annealing structures for the KH3 and KH4 halves, respectively, of the FBP3/4M29 ssDNA complex. The corresponding number of terms for the various restraints is given in

parentheses. None of the structures exhibit interproton distance violations >0.5 Å or torsion angle violations >5°. There are 1029 and 877 structurally useful interproton distance restraints for the KH3 and KH4 halves of the complex, respectively, comprising (KH3/KH4): 938/756 interproton distances within the protein [157/169 intraresidue, and 274/216 sequential (|i-j|)=1), 247/155 medium ($1<|i-j| \leq 5$) and 260/216 long ($|i-j|>5$) range interresidue restraints]; 41/53 within the DNA; and 50/68 intermolecular between protein and DNA. 66/72 distance restraints for 33/36 backbone hydrogen bonds located in helices and sheets were added during the final stages of refinement using standard criteria[8].

*The dipolar coupling R-factor[30] which scales between 0 and 100% is defined as the ratio of the r.m.s. deviation between observed and calculated values to the expected r.m.s. deviation if the vectors were randomly oriented, given by $\{2D_a^2[4+3\eta^2]/5\}^{1/2}$ where D_a is the magnitude of the axial component of the alignment tensor and η the rhombicity . The values of D_a^{NH} and η, derived from the distribution of normalized dipolar couplings as described[15], are -7.2 Hz and 0.25, respectively, for the KH3 half of the complex, and -14.5 Hz and 0.16, respectively, for the KH4 half of the complex.

†The precision of the coordinates is defined as the average atomic r.m.s. difference between the individual 80 simulated annealing structures and the corresponding mean coordinates obtained by best-fitting to residues 2-74 of the protein and 16-21 of the ssDNA for the KH3 half of the complex, and to residues 104-174 of the protein and 4-11 of the ssDNA for the KH4 half of the complex. The values for the coordinate precision also refer to the same residues.

B. Explanation for the different magnitudes of the alignment tensor of the KH3 and KH4 halves of the complex in a dilute liquid crystalline medium of phage fd.

If the orientation of the KH4 and KH3 halves of the complex were fixed (i.e. no interdomain motion) the measured dipolar couplings for the KH4 and KH3 domains would be described by a single alignment tensor (i.e. the magnitude and orientation of the alignment tensors for the two domains would be identical). The fact that the magnitude of the alignment tensor for the KH3 domain is half that of the KH4 domain automatically indicates, in a completely unambiguous manner, that there must be significant interdomain motion. It does not, however, imply that reorientation of the KH3 half of the complex is of larger magnitude than that for the KH4 half. Alignment of macromolecules by a dilute liquid crystalline medium arises from two factors: (a) in a neutral liquid crystalline medium such as bicelles alignment is dominated by steric effects and can be predicted on the basis of molecular shape; (b) in charged liquid crystalline media such as phage fd, electrostatic factors also come into play (cf. Zweckstetter & Bax, *JACS 122*, 3791 (2000)). The shape and size of the KH3 and KH4 halves of the complex are similar so that discrimination in the magnitude of the alignment tensors for the KH3 and KH4 halves of the complex would be expected to be small in a neutral liquid crystalline medium. In a charged liquid crystalline medium (such as phage fd), however, differences in charge will result in differences in the magnitude of the alignment tensors for the two halves of the complex providing there is significant internal motion. Significant in this context means that the amplitude of the motion must be greater than about 20°. The more negative the domain, the less it will be aligned by the negatively charged rod-shaped phage fd particles. Dividing the complex in two, there are effectively 16 phosphates, corresponding to nucleotides 14-29, associated with the KH3 half of the complex, but only 13, corresponding to nucleotides 1-13, with the KH4 half. Hence, the smaller magnitude of the alignment tensor for the KH3 half of the complex relative to the KH4 half.

The take home message is that the difference in the magnitude of the alignment tensor for the dipolar couplings measured for the KH3 and KH4 halves of the complex is

diagnostic of sizeable interdomain motion. In the presence of significant interdomain motion, the alignment tensor for each domain will be affected by its size, shape and electrostatic properties. The dipolar couplings, however, do not provide any information regarding the time scale and magnitude of the internal motions (other than these have to span a cone of semi-angle greater than about 20°). The characterization of the internal motions in these terms is obtained from the ^{15}N relaxation data.

C. Determination of translational (D_{trans}) and rotational (D_w) diffusion coefficients and other parameters derived from the model free analysis of the ^{15}N relaxation data.

1. D_{trans}[19]

$$D_{trans} = k_B T/6\pi\eta R$$

where D_{trans} is the translational diffusion coefficient of domain i, k_B the Boltzmann constant, T the temperature (308K), η the viscosity of the solvent, and R the hydrodynamic radius of domain i. At 35°C, D_{trans} ~1.9×10^{-10} $m^2 s^{-1}$ for an effective hydrodynamic radius of ~16.5 Å for each KH domain. The value of R can readily be determined from the coordinates of each KH domain by generating an accessible surface and calculating the volume of the molecule using the program GRASP.

2. D_w[19]

$$D_w = \{\cos^2\theta(1+\cos\theta)^2\{\log[(1+\cos\theta)/2] + (1-\cos\theta)/2\}/[2(\cos\theta-1)] + (1-\cos\theta)(6 + 8\cos\theta - \cos^2\theta - 12\cos^3\theta - 7\cos^4\theta)/24\}/\{1-S_s^2)\tau_s\}$$

where D_w is the rotational diffusion coefficient of domain i, S_s^2 is the squared-order parameter for slow interdomain motion of domain i, and θ is the semi-angle of the cone in which domain i wobbles given by $\cos^{-1}\{[-1+(1+ 8S_s)^{1/2}]/2\}$. For KH3, $S_s^2 = 0.67$; for KH4, $S_s^2 = 0.70$. θ is therefore approximately the same for both domains (~30°). The calculated value of D_w for each domain is therefore ~1.7×10^7 s^{-2}.

3. Distance L from center of each domain to the hinge point of the interdomain motion[19]

$L_i = (D_{trans}/D_w)^{1/2}$

Since D_{trans} and D_w have essentially the same values for each domain (see above), the value of L for each domain is ~33 Å.

4. Total average length of the FBP3/4-M29 ssDNA complex

The total average length of the complex is given by the sum $R_{KH3} + L_{KH3} + R_{KH4} + L_{KH4}$. Since $R_{KH3} \sim R_{KH4} \sim 16.5$ Å and $L_{KH3} \sim L_{KH4} \sim 33$ Å, the overall average length of the complex is ~100 Å. If the overall complex is treated as a cylinder, the length to diameter ratio is 100 Å/33Å ~ 3. A body with these dimensions is expected to have a diffusion anisotropy of ~1.8, which is very close to the value of 1.85 derived from the model free analysis of the ^{15}N relaxation data. I.e. The data are self-consistent and the interdomain motion for each domain is centered about the same hinge point.

5. Average distance between the C-terminus of KH3 (residue 74) and the N-terminus of KH4 (residue 104)

From the dimensions calculated above and the coordinates of the two domains, one can estimate the average distance between C-terminus of KH3 (residue 74) and the N-terminus of KH4 (residue 104) to be ~35 Å. This value is very close to the expected average end-to-end distance of ~40 Å for a random-coil polypeptide of n=30 residues given by $(C_n n l^2)^{1/2}$ where l is the Cα-Cα distance (3.8 Å) and C_n the characteristic ratio (~3.7 for a copolyeptide with 37% Gly content)[20].

D. Further details regarding sample preparation

Residues 278-447 of human FBP (referred to as FBP3/4 and numbered from residues 5-174) were cloned in the pET15b vector as fusion protein with an N-terminal His-tag and expressed in *E. coli* BL21 (DE3) cells (Novagen). Subclones of the individual KH domains (residues 5-77 and 101-174 for the KH3 and KH4 constructs, respectively) were also prepared. Nucleotide sequences for all clones were confirmed by DNA sequencing, and the masses of the expressed proteins were determined by mass spectrometry. The recombinant proteins were purified by affinity chromatography on a Ni-NTA agarose column (Qiagen), gel filtration (Superdex-75, Amersham-Pharmacia) and reverse phase HPLC (Poros 20 R2 resin, PerSeptive Biosystems). The His-tag was cleaved prior to gel filtration with thrombin and factor Xa, leaving four extra residues (GSHM) at the N-termini of FBP3/4 and KH3 (numbered 1-4) and no extra residues for KH4. Oligonucleotides (Fig. 1b of main paper) were purchased from Midland Certified Reagent Co. and purified by anion exchange chromatography. Equilibrium dissociation constants were measured by isothermal titration calorimetry at 25 °C (MicroCal Inc hardware and software). The FBP3/4-M29 ssDNA complex used for NMR studies was purified on a Superdex S200 column. Samples for NMR contained 1:1 complexes of

protein (^{15}N, $^{15}N/^{13}C$ or $^{15}N/^{13}C/^{2}H$/VLI-methyl protonated) and ssDNA in 50 mM sodium phosphate, 20 mM EDTA and 0.02% NaN_3, pH 6.8.