 Nucleic Acids Research, 1998, Vol. 26, No. 9 **2143–2149**

DNA binding properties of a chemically synthesized DNA binding domain of hRFX1

Fabrice Cornille*, Patrick Emery[1,+], Wolfgang Schüler, Christine Lenoir, Bernard Mach[1], Bernard P. Roques and Walter Reith[1]

Département de Pharmacochimie Moléculaire et Structurale, INSERM U266, CNRS URA D1500, UFR des Sciences Pharmaceutiques et Biologiques, 4 Avenue de l'Observatoire, 75270 Paris Cedex 06, France and
[1]Laboratoire L. Jeantet de Génétique Moléculaire, Département de Génétique et de Microbiologie, Centre Médical Universitaire (CMU), 9 Avenue de Champel, CH 1211 Genève 4, Switzerland

Received January 7, 1998; Revised and Accepted March 9, 1998

ABSTRACT

The RFX DNA binding domain (DBD) is a novel highly conserved motif belonging to a large number of dimeric DNA binding proteins which have diverse regulatory functions in eukaryotic organisms, ranging from yeasts to human. To characterize this novel motif, solid phase synthesis of a 76mer polypeptide corresponding to the DBD of human hRFX1 (hRFX1/DBD), a prototypical member of the RFX family, has been optimized to yield large quantities (~90 mg) of pure compound. Preliminary two-dimensional ^{1}H NMR experiments suggested the presence of helical regions in this sequence in agreement with previously reported secondary structure predictions. In gel mobility shift assays, this synthetic peptide was shown to bind in a cooperative manner the 23mer duplex oligodeoxynucleotide corresponding to the binding site of hRFX1, with a 2:1 stoichoimetry due to an inverse repeat present in the 23mer. The stoichiometry of this complex was reduced to 1:1 by decreasing the length of the DNA sequence to a 13mer oligonucleotide containing a single half-site. Surface plasmon resonance measurements were achieved using this 5′-biotylinated 13mer oligonucleotide immobilized on an avidin-coated sensor chip. Using this method an association constant (K_a = 4×10^5/M/s), a dissociation constant ($K_d = 6 \times 10^{-2}$/s) and an equilibrium dissociation constant (K_D = 153 nM) were determined for binding of hRFX1/DBD to the double-stranded 13mer oligonucleotide. In the presence of hRFX1/DBD the melting temperature of the 13mer DNA was increased by 16°C, illustrating stabilization of the double-stranded conformation induced by the peptide.

INTRODUCTION

The RFX family of proteins consists of highly homologous site-specific DNA binding proteins designated RFX1–5 in man and mouse (1,2) Sak1 in *Shizosaccharomyces pombe* (3), ScRFX in *Saccharomyces cerevisiae* and CeRFX in *Caenorhabditis elegans* (4). These proteins are characterized by several highly conserved sequences, among which the most prominent is a novel 76 amino acid DNA binding domain (DBD) designated the RFX DNA binding motif (1,2). The recent identification of several new RFX genes has permitted a precise definition of the consensus motif present in the DBD and has shed light on the evolutionary conservation and functional importance of RFX proteins (4). These proteins function as regulatory factors in a wide variety of unrelated systems, including regulation of the mitotic cell cycle in yeast fission, control of the immune response in mammals and promotion of infection by human hepatitis B virus. Thus RFX1 was shown to be a cellular transactivator used by the highly pathogenic hepatitis B virus (5,6), while RFX5, a 75 kDa subunit of a nuclear protein complex, is involved in transcription of MHC class II, Ii chain and DM genes (7–9). These genes play a key role in the immune system by controlling presentation of foreign antigenic peptides to $CD4^+$ helper T lymphocytes. The RFX family of proteins appears, therefore, as widespread and functionally important group of proteins as those containing other well-known DNA binding motifs, such as the well-known zinc finger, homeodomain, basic leucine zipper and basic helix–loop–helix proteins.

The DNA binding motif shared by RFX1–5, Sak1, ScRFX and CeRFX proteins has no significant homology with any other known DNA binding motif (4; Fig. 1A). It is characterized by 30 highly conserved amino acids clustered in the C-terminal half of the domain. Particularly prominent are four glycine residues and numerous aromatic (W, F and Y), basic (K and R) and hydrophobic (I, L and V) residues.

The human hRFX1 protein binds DNA as homodimeric or heterodimeric (RFX1/RFX2 and RFX1/RFX3) complexes whose specific binding sites are inverted repeats, referred to as either EF-C or MDBP, or MHC class II X box motifs (Fig. 1B). These motifs are present in several cellular genes as well as in the enhancers of polyomavirus, cytomegalovirus and hepatitis B virus (2,5). The domain responsible for dimerization of hRFX1 corresponds to a conserved region situated in the C-terminal part of the protein, 220 amino acids downstream of the DBD.

*To whom correspondence should be addressed Tel: +33 1 43 25 50 45; Fax: +33 1 43 26 69 18; Email: cornille@pharmacie.univ-paris5.fr

[+]Present address: Howard Hugues Medical Institute, National Science Foundation Science and Technology Center for Biological Timing, Department of Biology, Brandeis University, Waltham, MA 02254, USA

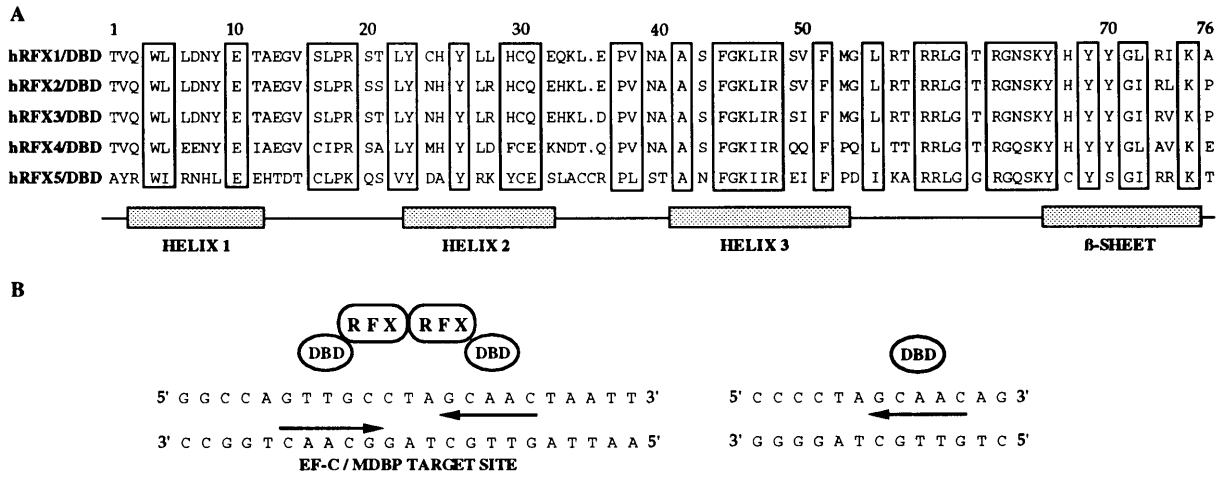

Figure 1. (**A**) Alignment between the DBD sequences of all known human RFX proteins. Amino acids that are identical or similar are boxed. The secondary structure prediction (20) is indicated below the sequences. Positions of the predicted α-helices and β-strand (extended) are indicated by boxes. (**B**) 23 bp DNA duplex containing the EF-C/MDBP target site of hRFX1/DBD. The presence of an inverted repeat in the sequence allows hRFX1 to bind as a homodimer (left). In order to simplify this model, our aim in this study was to obtain a 1:1 stoichiometric complex between the DBD of hRFX1 and one half-site of the EF-C/MDBP target site (right).

In order to characterize more precisely the mode of DNA recognition by proteins of the RFX family, the DBD of hRFX1 (hRFX1/DBD) was chosen as a model. Synthesis of this domain (76 amino acids) was achieved by the solid phase method, which required optimization of several coupling steps. The resulting peptide was obtained in large quantities (90 mg), allowing us to perform preliminary 600 MHz 2D ^{1}H NMR experiments, suggesting the occurrence of helical regions in the sequence. The synthetic peptide was also used to determine the minimum DNA binding site in order to simplify study of the interaction of hRFX1/DBD with DNA. The biotinylated 13mer DNA duplex characterized in this preliminary study was then immobilized on a streptavidin-coated biosensor chip in order to determine its kinetic binding parameters with hRFX1/DBD by surface plasmon resonance analysis and its selectivity for a double-stranded DNA. Finally, DNA duplex denaturation experiments were undertaken in order to investigate the effect of hRFX1/DBD on duplex stability.

MATERIALS AND METHODS

Peptide synthesis

Wang resin, piperidine, *N*-methylpyrrolidone, dichloromethane, dimethylaminopyridine, dicyclohexylcarbodiimide and 1-hydroxybenzotriazole were purchased from Perkin-Elmer (Saint Quentin en Yvelines, France). Trifluoroacetic acid was from SDS (Peypin, France) and triisopropylsilane and phenol were from Aldrich (Strasbourg, France).

Assembly of the protected peptide chain was carried out using the stepwise solid phase method of Merrifield (10) on an Applied Biosystems 431A automated peptide synthesizer. The different syntheses were run at 40 μmol scale on Wang resin (loading 1 mmol/g dry resin). Fmoc-protected amino acids were used with the following sidechain protection groups: t-butyl ether (Ser, Thr and Tyr); t-butyl ester (Glu and Asp); trityl (Cys, His, Asn and Gln); 2,2,5,7,8-pentamethylchroman-6-sulfonyl-5-(Arg), t-butoxycarbonyl (Lys and Trp). The hydroxy function on Wang resin was esterified by Fmoc alanine-activated as a symetrical anhydride with dimethylaminopyridine as the acylation catalyst. Deprotection of the Fmoc group during complete synthesis was obtained by three successive 3 min treatments with 20% piperidine in *N*-methylpyrrolidone. Successive couplings were performed with dicyclohexylcarbodiimide/hydroxybenzotriazole in *N*-methylpyrrolidone as coupling agent. Single couplings of 45 min were used (total cycle 1 h 15 min). *N,O*-bisFmoc derivatives of *N*-(2-hydroxy-4-methoxybenzyl)glycine was used for coupling of Gly53 in the second synthesis and for coupling of Gly44 and Gly53 in the third synthesis. In both cases coupling of the amino acid immediately following these glycines was performed using the symetric anhydride procedure with dicyclohexylcarbodiimide in dichloromethane as described in Johnson *et al.* (11). Fmoc deprotection in these syntheses was followed by spectrophotometric monitoring of the dibenzofulvene piperidine adduct at λ = 301 nm (Fig. 2). The total time to carry out complete synthesis was 102 h (4 days) in the optimized synthesis (Fig. 2C). The peptide resins were treated for 4 h with 40 ml trifluoroacetic acid containing phenol (3 g), H_2O (2 ml) and triisopropylsilane (1 ml) in order to cleave the peptide from the resin and to remove the protecting groups from the amino acid sidechains. Trifluoroacetic acid was removed under vacuum and the peptide precipitated with t-butylmethyl ether previously cooled to –20 °C. Fractions of 384, 235 and 705 mg crude peptide were obtained in the first, second and third syntheses respectively (Fig. 2). Purification of the peptide was carried out by reverse phase HPLC on a Vydac C4 10 × 250 mm column with a linear gradient of 20–70% B over 90 min (A, 0.1% trifluoroacetic acid in H_2O; B, 70% acetonitrile, 0.09% trifluoroacetic acid in H_2O) at a flow rate of 2 ml/min with detection at a wavelength of 234 nm. The fractions containing the pure peptide (~60 min retention time) were collected and freeze dried to afford 28, 12 and 88 mg purified HRFX1 in the first, second and third syntheses, corresponding to 7, 3 and 21% yield respectively (mol. wt 10 531 as the trifluoroacetate salt). The quality of each synthesis was determined using HPLC and electrospray mass spectrometry (Fig. 3).

NMR experiments

hRFX1/DBD was dissolved in water containing 30% TFE-d2 (SDS, Peypin, France) to a final sample concentration of 1 mM. Two-dimensional phase-sensitive ^{1}H Clean-TOCSY (70 ms spin

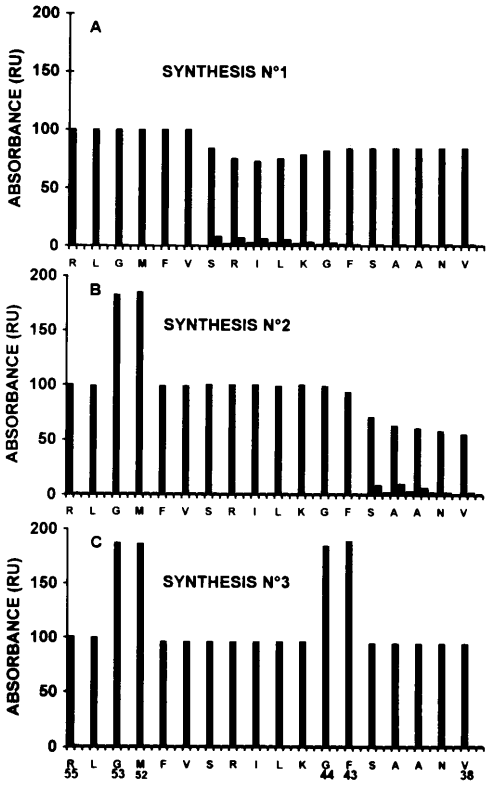

Figure 2. Ultraviolet monitoring ($\lambda_{301\ nm}$) of Fmoc deprotection during the three different syntheses of the DBD of hRFX1. The scheme is limited to the 38–55 region of hRFX1/DBD because the difficult couplings and deprotections were located only in this part of the synthesis. The sequence 38–55 of hRFX1 is represented in the reverse orientation, which corresponds to the order of incorporation of solid phase synthesis from the C- to the N-terminus. (**A**) Fmoc deprotection scheme of the first synthesis. Difficult couplings and deprotections are located from Ser49 to Gly44. (**B**) Incorporation of a 2-hydroxy-4-methoxy-benzyl as amino protecting group of Gly53 prevented any synthesis problem until Gly44, but new synthetic difficulties occurred at Phe43–Ala40. (**C**) The use of 2-hydroxy-4-methoxy-benzyl groups at Gly53 and Gly44 definitively solved the synthesis problems. The double intensity of peak heights corresponding to deprotection of Gly53, Met52, Gly44 and Phe43 is due to the presence of a *N,O*-bisFmoc on the glycine derivative, and the subsequent N,O double coupling of Fmoc methionine and Fmoc phenylalanine is due to the activation procedure of these residues as symmetrical anhydrides.

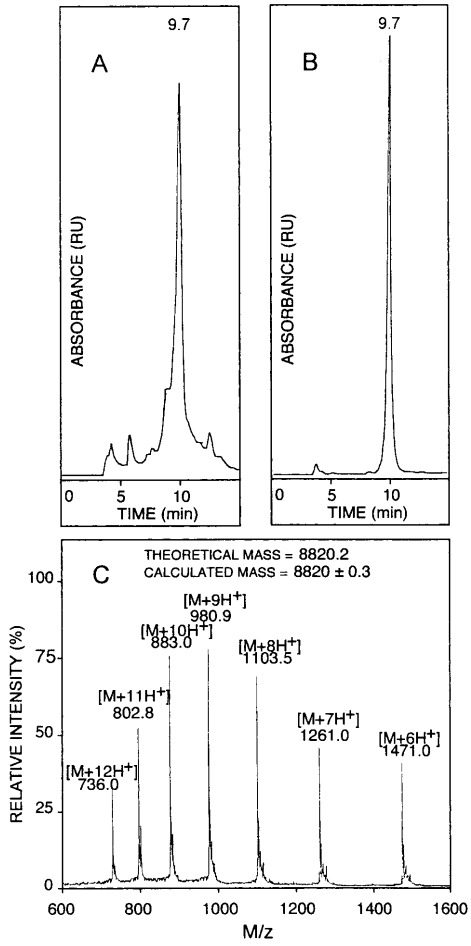

Figure 3. HPLC chromatograms of the crude (**A**) and purified (**B**) products from the third synthesis of hRFX1/DBD. The analyses were obtained with an acetonitrile gradient of from 50 to 80% B over 15 min at a flow rate of 1 ml/min on a capcell C_8 column (5 µm, 300 Å, 250 × 4.6 mm) with UV detection at λ = 214 nm; A, 0.1% TFA in H_2O; B, 70% CH_3CN, 0.09% TFA in H_2O. The retention time of the purified hRFX1 is 9.7 min under these conditions. (**C**) Ionisation electrospray mass spectrum of the purified hRFX1/DBD.

lock) (12) and NOESY (mixing times 100 and 200 ms) (13) spectra utilizing time-proportional phase incrementation (14) were recorded at 293 K and pH 4.0.

NMR spectra were recorded on a Brüker AMX 600 spectrometer operating at 600.14 MHz without sample spinning with 2K real t_2 points, with a spectral width of 7246 Hz, and 512 t_1 increments. The transmitter frequency was set to the water signal, which was suppressed by irradiation during the relaxation delay of 1.6 s between scans, and a water gate sequence before acquisition. The temperature was externally controlled using a SK107 unit (Haake GmbH, Karlsruhe, Germany). The data were processed with FELIX (Biosym/MSI). Phase-shifted sine bell (shifted by π/6) window functions were applied prior to Fourier transformation in both dimensions. The ^1H chemical shifts were referenced to hexamethyldisilazane (HMDS; SDS, France), used as internal standard.

Gel mobility shift assays

Binding conditions and gel electrophoresis were carried out essentially as described (2,5). Briefly, 100 000 c.p.m. (50 pg) double-stranded oligonucleotide probes (end-labeled with ^{32}P) were diluted with known amounts (see figure legends) of unlabeled oligonucleotide and incubated for 30 min on ice with 10 ng hRFX1/DBD in 20 µl binding buffer. Binding buffer was as described (2,5). Sequences of the duplex oligonucleotides were as follows: 23mer, 5′-GGCCAGTTGCCTAGCAACTAATT-3′; 16mer, 5′-CCCCTAGCAACAGATG-3′; 13mer, 5′-CCCCTAGCAACAG-3′; 12mer, 5′-CCCTAGCAACAG-3′. Only the upper strand is shown. Half-sites are underlined.

Surface plasmon resonance experiments

The BIAcore system used in these studies is from Pharmacia Biosensor AB (Uppsala, Sweden). Sensor chips and surfactant P20 were from Pharmacia Biosensor AB. *N*-Hydroxysuccinimide

(NHS), *N*-ethyl-*N*′-(3-diethylaminopropyl)carbodiimide (EDC), diaminopropane and streptavidin were purchased from Sigma. The complementary 13mers 5′-CCCCTAGCAACAG-3 and 5′-CTGTTGCTAGGGG-3′, biotinylated or not, were from Birsner & Grob-Biotech GmbH (Denzlingen, Germany).

The surface of the CM5 sensor chip was derivatized with streptavidin (2300 RU) using the EDC-NHS procedure according to the manufacturer's recommendations, followed by inactivation of the uncoupled dextran carboxylic groups with 1 M diaminopropane in H_2O.

Hybridization of 2.5 µg 13mer 5′-biotin(5′-CTGTTGCTAGG-GG-3′) was performed using 3.75 µg complementary 13mer (5′-CCCCTAGCAACAG-3′) in 500 µl 50 mM HEPES, pH 7.4, 0.5 M NaCl, 5 mM $MgCl_2$. After 1 h at 60°C, the temperature was allowed to decrease overnight to room temperature. Three different amounts of double-strand biotinylated oligonucleotide were immobilized on the sensor chip using 20 µl stock solution diluted 1:10 with prehybridized DNA which were injected at a flow rate of 5 µl/min over the three channels of the sensor chip. This led to immobilization of 120, 140 and 170 RU respectively.

Binding experiments were carried out at 25°C in 10 mM HEPES, pH 7.4, 0.15 M NaCl, 10 mM $MgCl_2$, 0.3 mM DTT, 3 mM EDTA, 0.005% P20. For kinetic studies 40 µl solutions of HRFX1 at concentrations ranging from 50 to 800 nM were injected at 20 µl/min. Concentration and ratio between peptide and immobilized nucleic acids were determined using the equation:

$$n = (RU_{peptide}/RU_{DNA}) \times (MW_{DNA}/MW_{peptide}) \quad \mathbf{1}$$

in which n is the stoichiometry of the complex, RU is the resonance unit and MW is the molecular weight.

After each injection of protein over the DNA, the sensor chip surface was regenerated by a 5 µl pulse (15 s) of 1 M NaCl in binding buffer.

Determination of kinetic constants

Apparent kinetic rates were calculated using the BIAcore 2.0 analytical software. To fit the dissociation curves, a single exponential dissociation rate equation

$$y = Re^{-K_d t} \quad \mathbf{2}$$

was used, where R is the observed response at the beginning of dissociation and K_d is the dissociation rate constant.

For determination of the association rate constant K_a, the binding curve of the sensorgram was fitted to the exponential function

$$R = R_{eq}(1 - e^{-K_s(t - t_O)}) \quad \mathbf{3}$$

where

$$K_s = K_a C + K_d \quad \mathbf{4}$$

C is the molar concentration of injected protein, K_a is the apparent association rate constant in per M/s units. K_s was plotted against concentration and the slope of the curve provided the apparent K_a value.

DNA duplex melting temperature

DNA denaturation experiments were achieved with a UVIKON 941 UV spectrophotometer piloted by KONTRON acquisition software for melting temperatures. Aliquots of 1.5 ml of a 1.5 µM solution of the 13mer oligodeoxynucleotide 5′-CCCCTAG-CAACAG-3′/5′-CTGTTGCTAGGGG-3′ solution and 2 µM hRFX1/DBD in 10 mM HEPES, pH 7.4, 1 mM β-mercaptoethanol were heated over a gradient of temperature ranging from 4 to 65°C at 20°C/h. Two other cells containing the same concentration of oligodeoxynucleotide or hRFX1/DBD respectively were used as controls. The absorbance at $\lambda_{260\ nm}$, which corresponds to maximum absorbance of nucleotide bases, was expressed in relative units and recorded. The absorbance of hRFX1/DBD alone at the same concentration ($D_o \sim 0.01$ RU) was recorded and substracted from the values obtained with DNA in the presence of hRFX1/DBD.

RESULTS

Peptide synthesis

Solid phase synthesis of the DBD of hRFX1 was carried out in a stepwise manner on Wang resin using the Fmoc/tBu strategy (15). Figure 2 shows UV monitoring of Fmoc deprotection during the different syntheses of hRFX1. The comparison of relative peak height at each deprotection step allows quantification of the amount of deprotection of the temporary protecting group Fmoc and consequently allows an estimation of the efficiency of the previous coupling. In Figure 2A, the Fmoc deprotection signal decreases from Ser49 to Gly44, indicating that incomplete coupling and deprotection are obtained under such conditions. This special feature is known as a 'difficult sequence' and is a consequence of interchain or intrachain association within the resin matrix (16). In order to circumvent this problem the use of *N*-(2-hydroxy 4-methoxybenzyl) derivatives of amino acids has recently been proposed (11,17). These *N*-protected amino acids prevent agregation through inter- or intramolecular hydrogen bonding (17). Gly53 was chosen to introduce a *N*,*O*-bisFmoc derivative in a second synthesis. As shown in Figure 2B, coupling and deprotection of this amino acid, as well as the subsequent sequence Met52–Gly44, proceeded correctly. Nevertheless, the pattern of a new 'difficult sequence' appeared after Gly44, as shown in Figure 2B. Therefore, the *N*,*O*-bisFmoc derivative of *N*-(2-hydroxy 4-methoxybenzyl)glycine was used twice, at Gly53 and Gly44, in a third assay. Figure 2C shows that in this case the synthesis proceeded correctly until the end. This last synthesis afforded 705 mg crude peptide, reverse phase HPLC analysis of which is represented in Figure 3A. This crude product was purified by HPLC on a Vydac C4 column with CH_3CN/H_2O containing 0.1% trifluoroacetic acid as eluent. Finally, 90 mg (21% yield) peptide HRFX1 were obtained. Purity was checked by analytical HPLC (Fig. 3B) and correct molecular mass was verified by electrospray mass spectrometry (Fig. 3C).

NMR experiments

In a preliminary assay, the NMR spectrum of synthetic hRFX1/DBD was achieved in 20 mM acetate buffer, 0.3 mM DTT, pH 5.4, H_2O/D_2O 9:1 at 293 K. A large broadening of the signals, indicating aggregation, was observed under these conditions (data not shown). In 30% trifluoroethanol in water, 5 mM DTT, pH 4, at 293 K (data not shown) the linewidths of the signals were sharper. Figure 4 shows the NH-NH region of the NOESY spectrum corresponding to this last sample. A large number of NH-NH correlations could be observed, indicating the presence

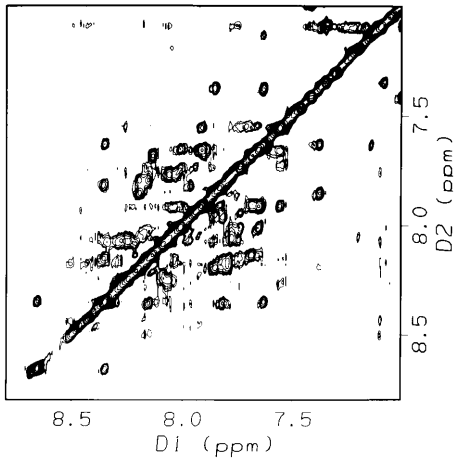

Figure 4. NH-NH region of a two-dimensional 600 MHz ^{1}H NMR NOESY spectrum corresponding to a solution of the synthetic hRFX1/DBD in H_2O/TFE 7:3, pH 4, 5 mM DTT at 293 K. The multiplicity of NH-NH NOE cross-peaks are indicative of helical structure.

of several helical regions. However, poor resolution of the signals in the NOESY/Clean-TOCSY spectra prevents sequence-specific assignment of the resonances by the chain tracing method. A more detailed analysis will require the use of $^{15}N/^{13}C$ enrichment of the hRFX1/DBD motif.

Gel mobility shift assays

To determine whether the synthetic hRFX1/DBD polypeptide is biologically functional, i.e. capable of binding specifically to DNA, we performed gel mobility shift assays (Fig. 5A). A 23 bp double-stranded oligonucleotide containing an inverted repeat corresponding to the EF-C/MDBP binding site was used. Depending on the ratio of protein to DNA in the binding reactions, two different protein–DNA complexes were formed. Previous studies (5) have demonstrated that the complex of greater mobility corresponds to a single hRFX1/DBD monomer bound to one half-site of the oligonucleotide, while the complex of slower mobility corresponds to binding of two hRFX1/DBD molecules to the two symmetrical half-sites of the oligonucleotide. When 10 ng hRFX1/DBD were incubated with 100 000 c.p.m. (50 pg) ^{32}P-labeled 23 bp DNA (Fig. 5A) only the 1:2 DNA–protein complex is observed. By adding increasing amounts of unlabeled 23 bp DNA an equilibrium appears between the 1:1 and 1:2 DNA–protein complexes, with only the 1:1 complex in the presence of excess DNA (50 ng, 3.3 equiv.). Interestingly, even in the presence of a slight excess of DNA (20 ng, 1.3 equiv.) to protein, the 1:2 DNA–protein complex was still as abundant as the 1:1 complex. This clearly indicates that binding of the first molecule of hRFX1/DBD to the 23 bp DNA favors binding of a second one in a cooperative manner.

Further gel shift experiments were performed to determine the minimum size of the DNA fragment required for binding of a monomer of hRFX1/DBD. A constant amount of hRFX1/DBD was incubated with double-stranded oligonucleotides of different sizes, each containing the same EF-C/MDBP half-site. Oligonucleotides of 20, 19, 16, 13, 12 and 10 bp were tested. Only the results for oligonucleotides of 16, 13 and 12 bp are shown in Figure 5B. Since these oligonucleotides contain only a single half-site, the

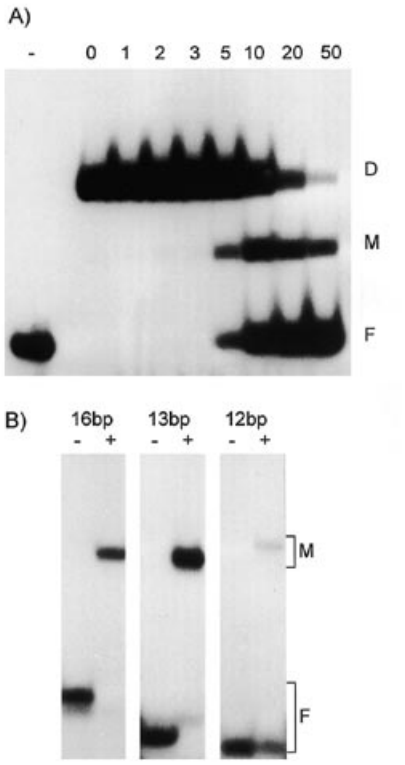

Figure 5. (**A**) Gel mobility shift assays were performed with 10 ng hRFX1/DBD and 100 000 c.p.m. (50 pg) 23 bp double-stranded oligonucleotide containing a palindromic binding site as shown in Figure 1B. The labeled probe was diluted with the quantities of unlabeled 23 bp oligonucleotide indicated above the lanes. In the left lane (–) no protein was added. Positions of free DNA (F) and protein–DNA complexes resulting from binding of a monomer (M) or a dimer (D) of hRFX1/DBD are indicated on the right. (**B**) Gel mobility shift assays were performed without (–) or with (+) hRFX1/DBD and double-stranded oligonucleotides of different lengths (16, 13 and 12 bp) containing a single half-site. Positions of free DNA (F) and the protein–DNA complex resulting from binding of an hRFX1/DBD monomer (M) are indicated on the right. Under the conditions used essentially all of the 16 and 13 bp oligonucleotides are complexed with hRFX1/DBD, while binding to the 12 bp oligonucleotide is severely impaired.

monomeric protein–DNA complex is exclusively formed. This complex forms efficiently on the 16 and 13 bp oligonucleotides. However, binding is largely abolished when the size of the oligonucleotide is reduced to 12 bp. The minimal size of oligonucleotide permitting efficient binding thus appears to be 13 bp.

Surface plasmon resonance experiments

In order to study the structural basis of the affinity and selectivity of the interaction between DNA and hRFX1 protein, surface plasmon resonance (SPR) measurements were used (18). The minimal sequence of DNA duplex able to interact with hRFX1 (5′-CCCCTAGCAACAG-3′/5′-CTGTTGCTAGGGG-3′) was biotinylated at the 5′-end of the first strand and immobilized onto streptavidin-coated chips of a BIAcore biosensor apparatus. Figure 6 shows SPR profiles obtained after introducing various concentrations of hRFX1 in aqueous buffer. Corresponding SPR signals, obtained in parallel by testing identical concentrations of hRFX1 on streptavidin-coated biosensor chips without oligo-

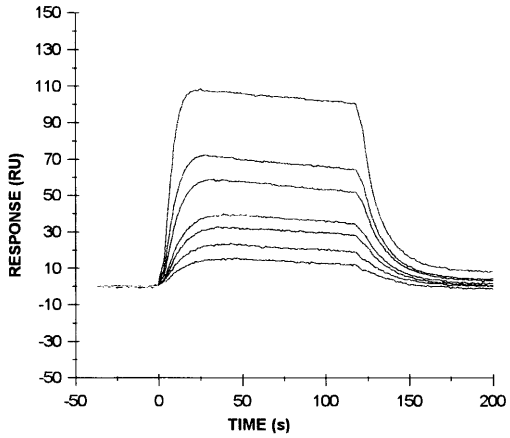

Figure 6. Determination of the kinetic parameters of the interaction of the DBD of hRFX1 to the DNA duplex using suface plasmon resonance (SPR). The response functions (sensorgrams) shown illustrate specific binding and dissociation profiles obtained after various concentrations of hRFX1 protein were added to the buffer flowing over a biosensor chip coated with the DNA duplex (CTGTTGCTAGGGG) immobilized via a streptavividin–biotin system. The data in these sensorgrams are expressed as corrected response units, since the contribution by the bulk refractive index (background) has been substracted. For clarity only the sensorgrams obtained using 80 (bottom), 140, 185, 250, 330, 445 and 600 nM (top) hRFX1 are shown. The association constant (K_a = $3.9 \pm 0.3 \times 10^5$/M/s), dissociation constant ($K_d = 6.0 \pm 0.24 \times 10^{-2}$/s) and the equilibrium dissociation constant ($K_D = 153 \pm 18$ nM) determined by this study are the average of three independent experiments at 25°C in 10 mM HEPES, pH 7.4, 0.15 M NaCl, 10 mM $MgCl_2$, 0.3 mM DTT, 3 mM EDTA, 0.005% P20 with immobilized amounts of DNA duplex ranging from 120 to 170 RU.

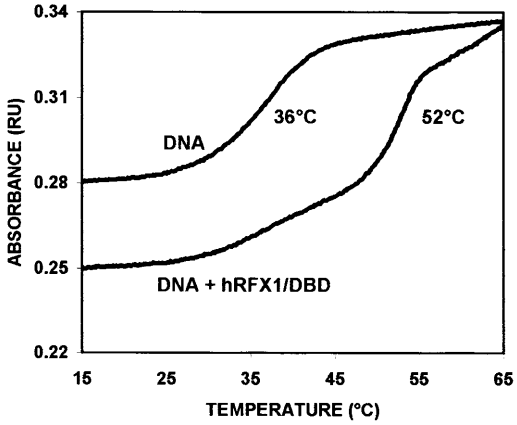

Figure 7. Melting temperature of a 1.5 μM solution of DNA duplex 13mer (5′-CTGTTGCTAGGGG-3′) in 10 mM HEPES, pH 7.4, 1 mM β-mercapto-ethanol obtained by UV monitoring at 260 nm with a gradient of temperature ranging from 6 to 70°C at 20°C/h. An elevation of 16°C in this melting temperature was obtained by adding 2 μM synthetic hRFX1/DBD to this solution, indicating increased stability of the DNA duplex.

nucleotide dimer (to control for bulk refractive index effects), have been substracted from these values. Therefore these profiles reflect the specific association between hRFX1 and the immobilized double-stranded oligonucleotide.

The association and dissociation rate constants were calculated for different protein concentrations (185, 250, 330, 445, 600 and 800 nM respectively) using the equations reported in Materials and Methods. Based on these measurements the data fitted a first order model well, yielding the kinetic parameters of binding. The values obtained for the association constant K_a ($3.9 \pm 0.3 \times 10^5$/M/s), the dissociation constant K_d ($6.0 \pm 0.24 \times 10^{-2}$/s) and the equilibrium dissociation constant K_D (153 ± 18 nM) are the average of three independent experiments (Fig. 6). Furthermore, when the association equilibrium reached a plateau, the resonance units were converted to a ratio of peptide bound to double-strand oligonucleotide using the equation given in Materials and Methods. The 1:1 stoichiometry observed by gel shift assays was thus confirmed by SPR experiments (not shown). In addition, hRFX1/DBD was unable to bind single-stranded biotinylated oligonucleotide (5′-CCCCTAGCAACAG-3′), immobilized on a streptavidin-coated sensor chip.

DNA duplex melting temperature

DNA denaturation experiments were performed on the 13mer duplex 5′-CCCCTAGCAACAG-3′/5′-CTGTTGCTAGGGG-3′. As shown in Figure 7, this oligonucleotide melts at 36°C in 10 mM HEPES, pH 7.4, 1 mM β-mercaptoethanol. Increased stabilization of this DNA duplex was observed in the presence of hRFX1/DBD (Fig. 7). Indeed, its melting temperature was raised to 52°C in the presence of 2 μM hRFX1/DBD (1.3 equiv.). The weak first transition around 36°C in the presence of protein corresponds to DNA not complexed with hRFX1/DBD. This data is consistent with the equilibrium dissociation constant determined by BIAcore, which indicates that under the conditions used for DNA denaturation the oligonucleotide would not be totally complexed with hRFX1/DBD.

DISCUSSION

With the aim of characterizing kinetic and physicochemical constants of a novel DNA binding motif present in the family of RFX proteins, solid phase peptide synthesis of a 76mer polypeptide, hRFX1/DBD, was performed using the Fmoc/tBu strategy. In the first synthesis difficulties in the deprotection steps and incomplete couplings were observed from Gly44 to Ser49 (Fig. 2A). This is a consequence of interchain or intrachain association occurring within the resin matrix in these so-called 'difficult sequences' (16). In order to overcome this problem, *N*-2-hydroxy 4-methoxy-benzyl derivatives of glycine (11,17) were used at Gly53 in the second synthesis, then at both Gly44 and Gly53 in the third synthesis. This led to disappearance of incomplete couplings and deprotections, as observed in Figure 2C. The fact that glycine is present two or three residues before the 'difficult sequence' and that this residue has a high propensity to be involved in β-turns (19) suggests that intramolecular β-sheet formation could be responsible for the synthesis problem encountered in our case.

Careful characterization of 76 amino acid synthetic hRFX1/DBD by HPLC and electrospray mass spectrometry (Fig. 3) indicated that purity of the compound was >95%. This allowed preliminary ^{1}H NMR experiments to be performed in order to investigate the secondary structure of this polypeptide in solution. The spectrum of synthetic hRFX1/DBD in aqueous buffer exhibited broad signals characteristic of aggregation (data not shown). Nevertheless, addition of trifluoroethanol (30%) to the aqueous solution (data not shown) improved resolution of the proton signals leading to the observation in the NOESY spectrum of a pattern characteristic of helical structure (Fig. 4). This is in agreement with secondary

structure predictions previously described for this sequence (see Fig. 1A) (20).

Gel mobility shift assays were performed with the synthetic peptide showing that it was able to bind to a double-stranded 23mer oligodeoxynucleotide containing the binding site of native hRFX1. Due to the presence of inverted repeats in the DNA, the peptide was shown to form a 2:1 complex in a cooperative manner with this oligonucleotide, as is observed with entire hRFX1 protein. Using oligonucleotides of different lengths, the optimum double-stranded oligonucleotide sequence leading to a 1:1 complex with hRFX1/DBD was determined to be the 13mer sequence 5′-CCCCTAGCAACAG-3′/5′-CTGTTGCTAGGGG-3′ (Fig. 5).

This 5′-biotinylated DNA 13mer was used in SPR experiments to determine its association (4×10^5/M/s) and dissociation constants (6×10^{-2}/s) for the hRFX1/DBD peptide, leading to an equilibrium dissociation constant K_D of 153 nM. These values are in agreement with those found for other nucleic acid binding motifs, such as the zinc finger domain of HIV-1 NCp7 (21). However, a high selectivity and a higher affinity, by at least three or four orders of magnitude, of hRFX1 proteins for the EFC/MDBP binding site are expected to result from a cooperative effect between the two subunits of the hRFX1 dimer. This was observed with several DNA binding proteins, such as *Escherichia coli* catabolite activator protein, which binds DNA as a dimer of two identical protomers and exhibits 2-fold symmetrical interactions with a 22 bp 2-fold symmetrical DNA site (22). The equilibrium dissociation constant of catabolite activator protein is in the subnanomolar range, which is in agreement with the calculated increase in affinity, as compared with a monomer, resulting from interaction of a dimer with two symmetrical binding sites on DNA (23).

Finally, DNA duplex denaturation experiments performed on the DNA 13mer allowed us to demonstrate significant stabilization of the DNA duplex in the presence of synthetic hRFX1/DBD peptide.

In conclusion, this study has allowed us to determine the minimum double-strand DNA binding site and the binding constants for this domain. The results of these experiments are essential for subsequent investigation of the three-dimensional structure in solution of $^{13}C/^{15}N$-labeled hRFX1/DBD by NMR and of its complex with the 13mer oligonucleotides by X-ray crystallography. These studies are now in progress.

ACKNOWLEDGEMENTS

We would like to acknowledge Dr J.Paoletti for providing UV spectrometry facilities and for helpful discussions and E.Barras for excellent technical assistance. We thank C.Dupuis for her invaluable help in drafting this manuscript.

REFERENCES

1 Reith,W., Herrero-Sanchez,C., Kobr,M., Silacci,P., Berte,C., Barras,E., Fey,S. and Mach,B. (1990) *Genes Dev.*, **4**, 1528–1540.
2 Reith,W., Ucla,C., Barras,E., Gaud,A., Durand,B., Herrero-Sanchez,C., Kobr,M. and Mach,B. (1994) *Mol. Cell. Biol.*, **14**, 1230–1244.
3 Wu,S.-Y. and McLeod,M. (1995) *Mol. Cell. Biol.*, **15**, 1479–1488.
4 Emery,P., Durand,B., Mach,B. and Reith,W. (1997) *Nucleic Acids Res.*, **24**, 803–807.
5 Siegrist,C.A., Durand,B., Emery,P., David,E., Hearing,P., Mach,B. and Reith,W. (1993) *Mol. Cell. Biol.*, **13**, 6375–6384.
6 Garcia,A.D., Ostaptchuk,P. and Hearing,P. (1993) *J. Virol.*, **67**, 3940–3950.
7 Steimle,V., Durand,B., Barras,E., Zufferey,M., Hadam,M.R., Mach,B. and Reith,W. (1995) *Genes Dev.*, **9**, 1021–1032.
8 Kern,I., Steimle,V., Siegrist,C. and Mach,B. (1995) *Int. Immunol.*, **7**, 1295–1299.
9 Mach,B., Steimle,V., Martinez-Soria,E. and Reith,W. (1996) *Annu. Rev. Immunol.*, **14**, 301–331.
10 Merrifield,R.B. (1963) *J. Am. Chem. Soc.*, **85**, 2149–2152.
11 Johnson,T., Quibell,M. and Sheppard,R.C. (1995) *J. Peptide Sci.*, **1**, 11–25.
12 Griesinger,C., Otting,G., Wüthrich,K. and Ernst,R.R. (1988) *J. Am. Chem. Soc.*, **110**, 7870–7872.
13 Jeener,J., Meier,B.H., Bachmann,P. and Ernst,R.R. (1982) *J. Chem. Phys.*, **71**, 4546–4553.
14 Marion,D. and Wüthrich,K. (1983) *Biochem. Biophys. Res. Commun.*, **113**, 967–974.
15 Atherton,E., Logan,C.J. and Sheppard,R.C. (1981) *J. Chem. Soc. Perkin Trans.*, **1**, 538.
16 Larsen,B.D. and Holm,A.(1994) *Int. J. Peptide Protein Res.*, **43**, 1–9.
17 Hyde,C., Johnson,T., Owen,D., Quibell,M. and Sheppard,R.C. (1994) In Epton,R. (ed.), *Innovation and Perspectives in Solid Phase Synthesis—Peptides, Proteins and Nucleic Acids.* Mayflower Worldwide, Birmingham, AL, Vol. 3, pp. 29–38.
18 O'Shannessy,D.J., Brigham,B.M., Soneson,K.K., Hensley,P. and Brooks,J. (1993) *Anal. Biochem.*, **212**, 457–468.
19 Hutchinson,E.G. and Thornton,J.M. (1994) *Protein Sci.*, **3**, 2207–2216.
20 Emery,P., Strubin,M., Hofman,K., Bucher,P., Mach,B. and Reith,W. (1995) *Mol. Cell. Biol.*, **16**, 4486–4494.
21 Mély,Y., de Rocquigny,H., Sornas-Jimeno,M., Keith,G., Roques,B.P., Marquet,R. and Gérard,D. (1995) *J. Biol. Chem.* **270**, 1650–1656.
22 Parkinson,G., Gunasekera,A., Vojtechovsky,J., Zhang,X., Kunkel,T.A., Berman,H. and Ebright,R.H. (1996) *Nature Struct. Biol.*, **3**, 837–841.
23 Le Pecq,J.B., Le Bret,M., Barret,J. and Roques,B.P. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 2915–2919.