

Binding of the J-Binding Protein to DNA Containing Glucosylated hmU (Base J) or 5-hmC: Evidence for a Rapid Conformational Change upon DNA Binding

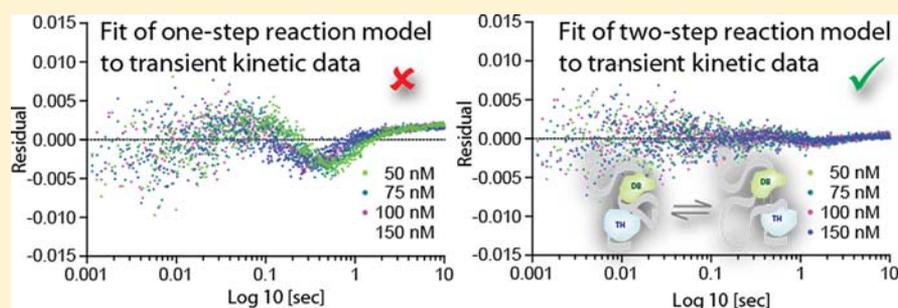
Tatjana Heidebrecht,[†] Alexander Fish,[†] Eleonore von Castelmur,[†] Kenneth A. Johnson,[‡] Giuseppe Zaccai,[§] Piet Borst,^{||} and Anastassis Perrakis^{*,†}

[†]Division of Biochemistry and ^{||}Division of Molecular Oncology, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands

[‡]Institute for Cellular and Molecular Biology, University of Texas, Austin, Texas 78712, United States

[§]Institut Laue-Langevin, 6 rue Jules Horowitz, BP 156, 38042 Grenoble Cedex 9, France

Supporting Information



ABSTRACT: Base J (β -D-glucosyl-hydroxymethyluracil) was discovered in the nuclear DNA of some pathogenic protozoa, such as trypanosomes and *Leishmania*, where it replaces a fraction of base T. We have found a J-Binding Protein 1 (JBP1) in these organisms, which contains a unique J-DNA binding domain (DB-JBP1) and a thymidine hydroxylase domain involved in the first step of J biosynthesis. This hydroxylase is related to the mammalian TET enzymes that hydroxylate 5-methylcytosine in DNA. We have now studied the binding of JBP1 and DB-JBP1 to oligonucleotides containing J or glucosylated 5-hydroxymethylcytosine (glu-5-hmC) using an equilibrium fluorescence polarization assay. We find that JBP1 binds glu-5-hmC-DNA with an affinity about 40-fold lower than J-DNA (~ 400 nM), which is still 200 times higher than the JBP1 affinity for T-DNA. The discrimination between glu-5-hmC-DNA and T-DNA by DB-JBP1 is about 2-fold less, but enough for DB-JBP1 to be useful as a tool to isolate 5-hmC-DNA. Pre-steady state kinetic data obtained in a stopped-flow device show that the initial binding of JBP1 to glucosylated DNA is very fast with a second order rate constant of $70 \mu\text{M}^{-1} \text{s}^{-1}$ and that JBP1 binds to J-DNA or glu-5-hmC-DNA in a two-step reaction, in contrast to DB-JBP1, which binds in a one-step reaction. As the second (slower) step in binding is concentration independent, we infer that JBP1 undergoes a conformational change upon binding to DNA. Global analysis of pre-steady state and equilibrium binding data supports such a two-step mechanism and allowed us to determine the kinetic parameters that describe it. This notion of a conformational change is supported by small-angle neutron scattering experiments, which show that the shape of JBP1 is more elongated in complex with DNA. The conformational change upon DNA binding may allow the hydroxylase domain of JBP1 to make contact with the DNA and hydroxylate T's in spatial proximity, resulting in regional introduction of base J into the DNA.

INTRODUCTION

J-binding protein 1 (JBP1) is one of the two proteins^{1,2} involved in the conversion of specific T-residues in DNA into hydroxymethyluracil (hmU),^{3–5} the first step in the biosynthesis of base J⁶ (β -D-glucosyl-hydroxymethyluracil, Chart 1). Base J replaces about 1% of T in the nuclear DNA of kinetoplastid flagellates,⁷ including the pathogenic *Trypanosoma* and *Leishmania* species,⁸ but is absent from other eukaryotes, prokaryotes, and viruses (reviewed in ref 9). Base J is mainly present in the telomeric repeat sequence $(\text{GGGTTA})_n$ ^{7,10} but small amounts are also found in other repetitive sequences¹¹

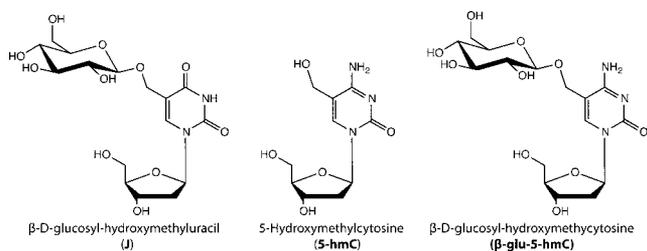
and in sequences between transcription units, called internal J (ij).^{12,13} This ij is important for proper transcription initiation in *Trypanosoma*^{13,14} and essential for correct transcription termination in *Leishmania*.¹⁵

We have shown that JBP1 contains at least two functional domains, a J-DNA-binding domain (DB-JBP1) and a thymidine hydroxylase (TH) domain.⁵ The TH domain catalyzes the first step in J biosynthesis, the hydroxylation of T in DNA. JBP2 is

Received: April 18, 2012

Published: July 9, 2012

Chart 1. Structures of J-base, 5-hmC, and glu-5-hmC



the only other kinetoplastid homologue of JBP1. It shares 34% identity with JBP1 in its N-terminal half² that contains the thymidine hydroxylase function.^{3,4} The only eukaryotic homologue of the JBP1/2 hydroxylase domain was identified in the mammalian protein TET1^{16,17} and in the related TET2 and TET3. All TET proteins catalyze the conversion of 5-methylcytosine (5-mC) in DNA to 5-hydroxymethylcytosine (5-hmC, Chart 1). This reaction eventually results in the removal of 5-mC from DNA and plays an important role in the epigenetic control of gene expression.^{18–24} JBP and TET proteins have been grouped together in the TET/JBP subfamily of iron- and 2-oxoglutarate-dependent dioxygenases.²⁵

The binding of JBP1 to J-DNA was first studied by biochemical methods^{26,27} using defined oligonucleotides containing J. That work showed that JBP1 only binds to duplex DNA, that at least five base pairs are required flanking base J for optimal binding, and that J-DNA recognition only requires base J itself and the base immediately 5' of J. Creating bases where the sugar residue in J was varied, and incorporating them in oligonucleotides, allowed it to be deduced that the hydrogen bonds between the 2- and 3-hydroxyl groups of the sugar are important for JBP1 binding.²⁸ Molecular modeling in the same study suggested that the sugar moiety is in an edge-on conformation, with the β -D-glucose lying within the major groove of J-containing DNA. The edge-on conformation is likely maintained by hydrogen bonds of the essential 2- and 3-hydroxyl groups and the phosphoryl-oxygen of the nucleotide upstream of J.²⁸

JBP1 has recently been used as a biotechnological tool for the identification of 5-hmC-containing DNA regions.^{29,30} This was based on the observation that JBP1 binds to DNA containing glucosylated 5-hmC (glu-5-hmC, Chart 1), with an affinity that is only about 15-fold lower than that for J-DNA.²⁷ After glucosylation of hmC residues in DNA with T4 glucosyl transferase, JBP1 could be used to enrich 5-hmC-containing regions in human DNA.³¹ This is not unexpected in view of the crucial role of the glucose moiety in the binding of JBP1 to J-DNA.^{28,32}

We recently showed that JBP1 contains a 160-residue autonomous folding unit (domain), the DNA-Binding JBP1 domain (DB-JBP1), located in the C-terminal half of the protein.³² DB-JBP1 binds to J-DNA with approximately the same affinity and specificity as full-length JBP1. In agreement with this, Hydrogen/Deuterium exchange mass spectrometry (HDX-MS) experiments showed that the most significant changes upon binding of full-length JBP1 to J-DNA are indeed all within specific regions of the DB-JBP1 domain. We have determined the crystal structure of DB-JBP1, revealing a novel "helical bouquet" fold. By site-directed mutagenesis and DNA binding studies we found that a single aspartate residue in JBP1 (Asp-525) is essential for the specific recognition of base J in DNA *in vitro* and for the function of JBP1 *in vivo*. Replacement

of this aspartate by an alanine completely abolishes specific binding of JBP1 to J-DNA. These data allowed us to construct a structural model for the binding of DB-JBP1 to J-DNA, which is supported by small-angle X-ray scattering measurements of the complex between DB-JBP1 and J-DNA.

To better understand the molecular mechanism of JBP1 binding to J-DNA and also to rationalize and extend its application as a tool to recognize hmC-DNA, we present here the characterization of the binding of JBP1 and DB-JBP1 to glu-5-hmC-DNA compared to J-DNA. Using fluorescence polarization measurements in a stopped-flow device, we establish that JBP1 binds both to J-DNA and to glu-5-hmC-DNA in a two-step reaction. In contrast, DB-JBP1 binds to both substrates in a one-step reaction. The kinetics of binding and the results of small-angle neutron scattering (SANS) experiments suggest that JBP1 undergoes a conformation change upon DNA binding. We infer that this conformational change, induced by the binding of the DB-JBP1 domain to DNA, moves the thymidine hydroxylase domain of JBP1 closer to the DNA, facilitating hydroxylation of T-residues.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification. Both JBP1 and DB-JBP1 from *Leishmania tarentolae* have been produced as soluble protein in *E. coli* as previously described.³² Briefly, vectors were transformed to BL21(DE3)T1R cells and protein production was induced with IPTG at 15 °C for 16–18 h. Cells were lysed in 20 mM Hepes/NaOH pH 7.5, 350 mM NaCl, 1 mM tris(2-carboxyethyl)phosphine (TCEP), and 10 mM imidazole (Buffer A), bound in batch to Ni-chelating sepharose beads, and eluted in Buffer A containing 400 mM imidazole. Purified proteins were cleaved with 3C protease overnight at 4 °C and applied to a S75 16/60 gel filtration column.

Preparation of Double-Stranded Labeled Oligonucleotides. All oligonucleotides were based on the sequence of the 14-mer J-oligo we have previously described (5'-GGCAGCJGCAACAA-3').^{28,32} An hmC-containing oligonucleotide with the same sequence (5'-GGCAGC(hmC)GCAACAA-3') was purchased from Glen Research Corporation and checked by mass spectrometry for purity. Control oligonucleotides with T and C at the place of the modified base (5'-GGCAGCTGCAACAA-3' and 5'-GGCAGCCGCAACAA-3', respectively) were purchased from Invitrogen. Two complementary oligonucleotides (with A for the J/T and G for the C/hmC oligonucleotides) were purchased from Invitrogen, and were labeled in the 5' with tetramethylrhodamine (TAMRA). Each complementary pair of oligonucleotides was dissolved in water and heat-annealed to the corresponding duplex DNA that was purified over a Superdex-75 10/30 Hi-Load (GE Healthcare) gel filtration column.

Creation of β -D-Glucosyl-hydroxymethylcytosine Oligos. The hmC-containing oligonucleotide and the complementary strand were dissolved in water to a concentration of 100 μ M and then heat-annealed. The double-stranded oligonucleotide was then glucosylated by the T4 Phage β -glucosyltransferase (T4-BGT) from New England BioLabs according to manufacturer's instructions.

Fluorescence Polarization Anisotropy Assays. This experiment was performed as we previously described.¹⁵ Briefly, the maximum amount of all JBP1 protein variants that was needed for each experiment was mixed with 1 nM of each duplex DNA, and subsequent dilutions were achieved by serial 1:1-dilutions in three repeats, in 96-well optiplates (Perkin-Elmer). All plates were read in a Perkin-Elmer or a BMG Pherastar fluorimeter, with an excitation filter with a CWL of 531 nm, and P and S emission filters with a CWL of 579 nm, at room temperature. The analysis of the equilibrium data has been performed in Graphpad/Prism by nonlinear regression analysis, using the formula:³³

$$A = A_f + (A_b - A_f) \frac{(L_t + P_t + K_D) - \sqrt{(L_t + P_t + K_D)^2 - 4P_t L_t}}{2L_t}$$

where A_f is the anisotropy of the free ligand, A_b the anisotropy of the bound ligand, L_t the concentration of the ligand (labeled DNA), P_t the protein concentration (JBP1 or variants), and K_D the dissociation constant. For the T-DNA and hmC-DNA affinities were estimated between 90 and 200 μM when constraining A_b to be the same as for the specific binding; the exact values would need higher concentration and would add limited insight in the context of this work.

Stopped-Flow Experiments. All stopped-flow experiments were carried out with use of a TgK Scientific stopped-flow system (model SF-61DX2). For fluorescence excitation we used monochromatic light at 546 nm and for the readout an OG 570 filter. The light was polarized by using a calcite prism for the incident beam and dichroic sheet polarizers in front of each of two photomultiplier detectors arranged in a T-configuration. The experiments were performed at 20 °C in 20 mM Hepes/NaOH pH 7.5, 140 mM NaCl, and 1 mM TCEP. For the association experiments equal volumes of various protein concentrations and 2 nM of double-stranded TAMRA labeled oligonucleotides were rapidly mixed in a 1:1 ratio. For the dissociation experiments, JBP1:DNA complex containing varying amounts of protein and 4 nM double-stranded TAMRA-labeled oligonucleotides were rapidly mixed in a 1:1 ratio with buffer. The measurements were repeated 10 times for JBP1 experiments and 20 times for DB-JBP1 experiments and recorded for 10–20 s.

Analysis of Pre-Steady State Kinetic Data. For the preliminary analysis of the pre-steady state kinetics stopped-flow data we used Graphpad/Prism to fit the data with two possible models. A single exponential was used for representing the one-phase association model:

$$A_{\text{obs}}(t) = A_0 + (A_{\text{max}} - A_0)(1 - e^{-kt}) \quad (1)$$

where A_{obs} is the observed anisotropy as a function of time (t), k is the apparent association constant, A_0 is a background measurement at the start and A_{max} is the anisotropy at the plateau at the end of the reaction.

A double exponential was used to represent the two-phase association model:

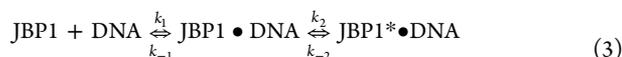
$$A_{\text{obs}}(t) = A_0 + \text{fraction}_{\text{fast}}(A_{\text{max}} - A_0)(1 - e^{-k_1 t}) + \text{fraction}_{\text{slow}}(A_{\text{max}} - A_0)(1 - e^{-k_2 t}) \quad (2)$$

with a similar notation as above, but where $\text{fraction}_{\text{fast}}$ and $\text{fraction}_{\text{slow}}$ each represent the signal accounted for by each of the two phases, under the constraint that

$$\text{fraction}_{\text{fast}} + \text{fraction}_{\text{slow}} = 1 \quad (2.1)$$

The fit was performed by using standard nonlinear regression methods as implemented in the software. The extra sum-of-squares F test, as implemented in the same software, was used to judge which of the two nested models fits best the data.

Global Analysis of Pre-Steady State and Equilibrium Kinetics Data. For the global analysis of kinetic (association and dissociation) and equilibrium data together we used the KinTek Explorer software.³⁴ The methods implemented in that software allow defining a model for the reaction at hand, and then relating all the experimental observations to that model, without making any specific assumptions or simplifications.³⁵ The model for the two-phase mechanism, presuming a conformational change that follows the initial binding, was described as:



The stopped-flow association data were related to that model according to the equation:

$$A_{\text{obs}}(t) = \alpha(\text{DNA}_t + \beta(C1_t + C2_t)) \quad (3.1)$$

where DNA_t , $C1_t$, and $C2_t$ are the concentrations of free DNA, the initial complex between DNA and JBP1, and the complex with the changed conformation, respectively, at time t . The parameters that related the concentration of these fluorescent components to the observed anisotropy are α (that is used to relate anisotropy to concentration) and β (that is used to express the difference in anisotropy between the free DNA and the DNA in complex with JBP1); as the two complexes are expected not to be dramatically different, it was assumed that both complexes change the anisotropy of the polarized signal in a similar fashion.

Each association curve was allowed a scaling factor to accommodate small errors in the protein concentration; these scaling factors never differed more than 5% (typically they were in the range of 2%) between curves, a reasonable error for diluting and mixing ingredients upon sample preparation. All association curves, each in a different starting JBP1 concentration, were fit according to this equation, with a common value for α and β .

To fit the dissociation curves we used the same function, with a different value of α ; using a different value for α has no physical meaning in the context of the two experiments, but we allowed it to vary as a scaling factor that would accommodate errors in the sample concentrations between association and dissociation experiments, like the scaling factor between concentrations.

To describe the equilibrium titration curves the same conceptual framework was used to derive the equation

$$A_{\text{obs}}(c) = \alpha_1(\text{DNA}_c + \beta_1(C1_c + C2_c)) \quad (3.2)$$

where the parameters have the same meaning as in eq 3.1 above, expressed as a function of concentration (c) instead of time (t). Different values for α and β were used as the two experiments were performed in different instruments, with different optics, sample holders, and sample volumes.

Association, dissociation, and equilibrium data were fit simultaneously. Estimates for σ values for all data points in each pre-steady state kinetics curve were obtained from calculating the deviation of the data from a double exponential model, while for the equilibrium data σ values were obtained from a triplicate experiment. These σ values were used as weights for fitting the residual with nonlinear regression algorithms. The value of $\chi^2/(\text{degrees of freedom})$ was used as a measure of global fit and for individual fits of each experiment.

From eq 3, if we define $K_1 = k_1/k_{-1}$ and $K_2 = k_2/k_{-2}$, it can be derived that an apparent K_D can be calculated according to the formula

$$K_D = 1/K_1(1 + K_2) \quad (4)$$

For the DB-JBP1 association experiment, where no evidence for a two-step mechanism was available, the global fit was performed by using the simplified equations:



$$A_{\text{obs}}(t) = \alpha(\text{DNA}_t + \beta C1_t) \quad (5.1)$$

$$A_{\text{obs}}(c) = \alpha_1(\text{DNA}_c + \beta_1 C1_c) \quad (5.2)$$

SANS Data Collection and Analysis. The same JBP1 protein and glu-5-hmC-DNA as used for all assays were also used for the SANS experiments. To create the JBP1:glu-5-hmC-DNA complex, concentrated JBP1 and double-stranded glu-5-hmC-DNA were mixed in a 1:1.2 molar ratio. Both JBP1 alone (10 mg mL⁻¹) and the complex (8 mg mL⁻¹) were purified by size exclusion chromatography (Superdex 200 10/30 Hi-Load gel filtration column) in 20 mM Hepes pH 7.5, 140 mM NaCl, and 1 mM TCEP buffer, either in 100% H₂O or 65% D₂O. The center of the peak fractions were collected and used without any further concentration for SANS measurements.

Measurements were done in rectangular Hellma quartz cuvettes of 1.0 mm optical path at the ILL instrument D22. Scattering data for all samples, for buffers alone, and for the empty cuvette were collected with a neutron wavelength of 8 Å at detector distances of 5.6 and 2 m

to cover the desired scattering vector modulus range $0.01 \text{ \AA}^{-1} < Q < 0.2 \text{ \AA}^{-1}$ ($Q = (4\pi \sin \theta)/\lambda$, where 2π is scattering angle and λ is wavelength). Since the neutron beam transmission of a sample depends on its $\text{D}_2\text{O}/\text{H}_2\text{O}$ ratio, it was verified that the different samples corresponded to the predicted values of 100% H_2O or 65% D_2O . Data from the two distances were averaged in the overlap region to improve counting statistics and angular resolution. Data reduction, calibration, and analysis were performed by the Ron Ghosh SANS suite of programs (<http://www.ill.eu/en/instruments-support/computing-for-science/cs-software/all-software/sans/>). The main analysis was performed in the Guinier approximation applied to the low Q region (see ref 36 for an overview), using the formula

$$\ln I(Q) = \ln I(0) - \frac{1}{3} R_G^2 Q^2 \quad (6)$$

where $I(0)$, the forward scattered intensity, is related to scattering particle concentration, molecular mass, and scattering density contrast with the solvent; R_G is the radius of gyration of scattering contrast in the particle. The approximation is considered valid for $R_G Q \approx 1$; typical fitting ranges for the current work were $0.5 < R_G Q < 1.5$. Note that both $I(0)$ and R_G are obtained and analyzed on an absolute scale.

RESULTS

JBP1 and DB-JBP1 Recognize glu-5-hmC-DNA with 200-Fold Preference over hmC-DNA or Normal DNA.

JBP1 has been shown to recognize DNA containing glucosylated cytosine with an affinity about 15 times lower than DNA containing glucosylated thymidine (J).²⁷ JBP1 could therefore be used in vitro to enrich glucosylated DNA to identify sequences that contain 5-hmC.³⁰ We wanted to better quantify the affinity of JBP1 for these two substrates. We have previously shown that the affinity of JBP1 for J-DNA is about 10 000 times higher than that for normal DNA (Table 1 and ref 32).

Table 1. Equilibrium Binding Constants (K_D in nM) for JBP1 and DB-JBP1 Binding to J-DNA and glu-5-hmC-DNA

protein	DNA	
	J	glu-5-hmC
JBP1	11.3 ± 0.6	401 ± 26
DB-JBP1	33.0 ± 1.3	945 ± 85

We extended these measurements using the same oligonucleotides, including now hmC and glu-5-hmC in the same position and sequence context as J or T. The affinity of JBP1 for glu-5-hmC-DNA is 401 nM (Table 1 and Figure 1), about 40 times lower than that for J-DNA (11.3 nM), and in good agreement with previous data based on gel shift assays.²⁷ Binding to hmC-DNA or T-DNA was practically indistinguish-

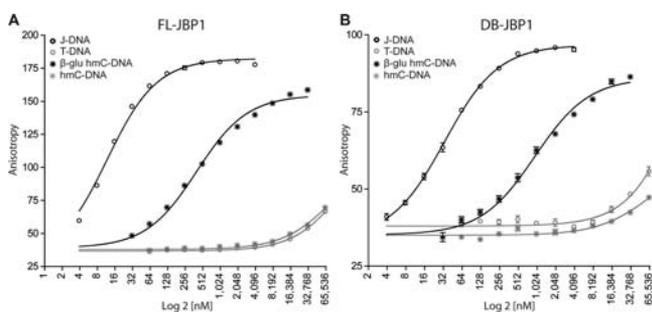


Figure 1. Equilibrium fluorescence polarization assay of JBP1 and DB-JBP1 binding to J-DNA, glu-5-hmC-DNA, and 5-hmC-DNA.

able, in the order of $\sim 100 \mu\text{M}$ (Figure 1) as we have previously reported.³² Hence, the discrimination of JBP1 between glu-5-hmC-DNA and unmodified or hmC-DNA is very good as the K_D values differ by a factor of about 200. This explains the utility of JBP1 for preferentially isolating glu-5-hmC-DNA from a pool of different DNA fragments.

Next we checked whether the difference in affinity for J-DNA and glu-5-hmC-DNA can be directly attributed to the DNA-binding module of JBP1 and DB-JBP1, or if regions outside the DB-JBP1 domain influence the affinity toward these two substrates. FP assays using DB-JBP1 show that the affinity toward glu-5-hmC-DNA is about 900 nM (Table 1 and Figure 1). The 2-fold loss of affinity of DB-JBP1 compared to JBP1 is comparable to its 3-fold loss of affinity toward J-DNA.

To find out whether glu-5-hmC-DNA recognition proceeds with the same residue as the recognition of J-DNA, we used our D525A mutant that has lost nearly all specificity toward J-DNA.³² As expected, the DB-JBP1-D525A mutant has also largely lost its ability to differentiate between J-DNA ($2.9 \mu\text{M}$), glu-5-hmC-DNA ($5.3 \mu\text{M}$) (Supporting Information, Figure 1), and normal DNA ($9.7 \mu\text{M}$ ³²), indicating that the Asp-525 residue recognizes the glucose moiety in the context of both J- and of glu-5-hmC-DNA.

JBP1 Binds DNA in a Two-Step Reaction. To delineate the binding mechanism of JBP1 to its substrates in more detail, we performed pre-steady state kinetics experiments. For these experiments, we adapted the FP assay to work in a stopped-flow device that allowed monitoring the initial stages of binding with a dead time of fewer than 5 ms. The same oligonucleotides used in the equilibrium FP assays were rapidly mixed in the flow cell with various concentrations of protein under the conditions we describe in Experimental Procedures.

As shown in Figure 2, the binding of JBP1 to J-DNA and glu-5-hmC-DNA is very fast and completes within the first second after mixing the reagents. We initially presumed a one-step reaction, the protein associating with DNA, which can be approximated by a single exponential function (see the Experimental Procedures for details). Although a single exponential did fit the binding data reasonably well over a wide range of JBP1 concentrations, we observed that the residuals of the fit (the difference between the model and the data) were appreciable and nonrandom (Figure 2).

We therefore tried to fit the data with a two-step association model that can be approximated by a double exponential function. The fit with a double exponential showed no systematic deviation of residuals; moreover an extra sum-of-squares F test, using the one-phase model as the null hypothesis and the two-phase model as the alternative hypothesis, suggests that our alternative hypothesis (two-step association) gives a significantly better fit to the data (p -value < 0.0001 for the alternative hypothesis to be incorrect). We therefore conclude that JBP1 association to J-DNA or glu-5-hmC DNA takes place in a two-step reaction.

We next examined the kinetics of the dissociation of JBP1 from J-DNA. For these experiments we performed a complex of labeled J-DNA or glu-5-hmC-DNA with various concentrations of JBP1 and rapidly diluted the complex. The dissociation data showed the same behavior (Supporting Information, Figures 2 and 3), albeit due to the more noisy character of the data the preference for the two-step dissociation model was less pronounced, but nevertheless statistically significant (p -values of 0.0447, 0.0063, and < 0.0001 for J-DNA and < 0.0001 for glu-5-hmC DNA). We therefore conclude that JBP1 dissociation

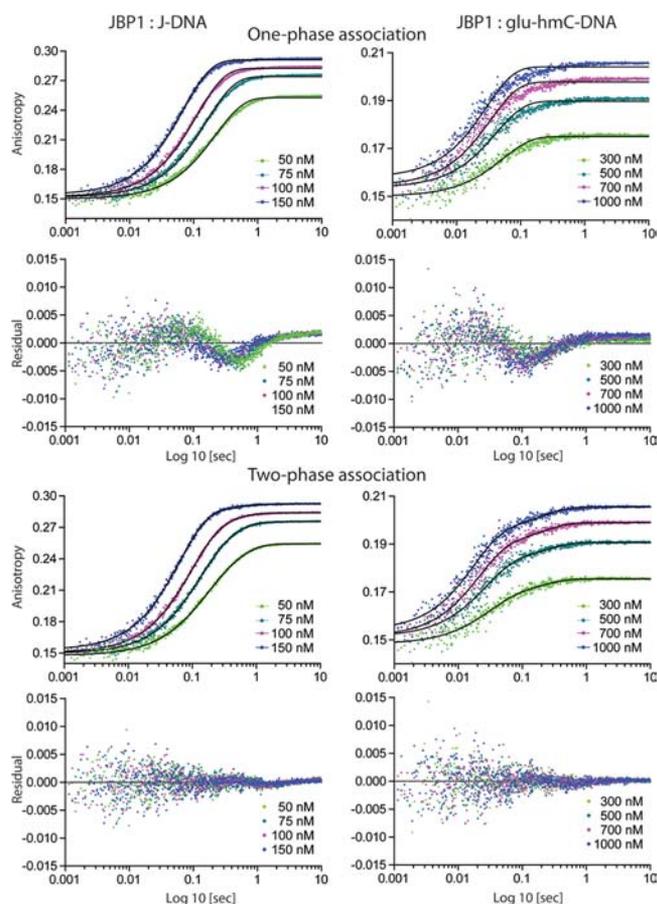


Figure 2. Pre-steady state kinetics association experiments of JBP1 with J-DNA and glu-5-hmC-DNA followed by polarization anisotropy are interpreted with either a one-phase or two-phase model. The fit to the associated data and the residuals are shown for both models.

from J-DNA or glu-5-hmC DNA also proceeds in a two-step reaction.

To address whether the same mechanism also holds for nonspecific binding of JBP1 to DNA, we used our D525A mutant that loses specificity toward J-DNA over normal DNA. Because of the replacement of the acidic D by a neutral A, this mutant JBP1 has a 10-fold higher affinity for T-DNA than wild-type JBP1, and therefore less mutant protein is required to study low-affinity binding. For the binding of JBP1-D525A to normal T-DNA we also find a preference for the two-step mechanism (Supporting Information, Figure 4; p -value < 0.0001). It should be noted, however, that the fit of the two-step mechanism to that data is still imperfect, implying that a more complex model might be applicable to T-DNA binding, at least in the context of the JBP1 D525A mutant.

To get more insight into the molecular implications of this two-step mechanism we next examined whether this mechanism is specific for full-length JBP1, or also applies to the DNA-binding domain alone, DB-JBP1.

DB-JBP1 Binds DNA in a One-Step Reaction. DB-JBP1 is a 160-residue domain with a helical bouquet fold,¹⁵ and an affinity toward J-DNA that is similar to full-length JBP1. The DB-JBP1 binding experiments were performed with the same oligonucleotides as for JBP1. As the DB-JBP1 is much smaller than JBP1, the change in the fluorescence anisotropy upon binding is small; the fluorescence anisotropy measurements were therefore repeated 20 times for each protein concen-

tration and the average was used for data analysis (Figure 3). In this case testing the two models resulted in p -values between

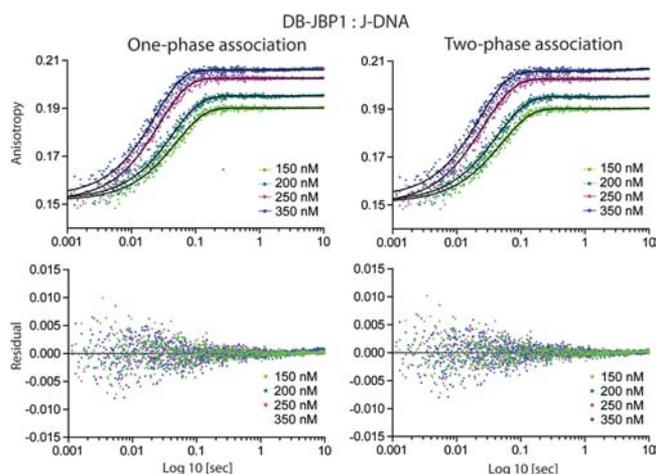


Figure 3. Pre-steady state kinetics association experiments of DB-JBP1 with J-DNA followed by polarization anisotropy are interpreted with either a one-phase or two-phase model. The fit to the associated data and the residuals are shown for both models.

0.04 and 0.64 for the two-step association model to be incorrect (as opposed to p -values less than 0.0001 in other association experiments). The same was observed for DB-JBP1 binding to glu-5-hmC-DNA (p -values between 0.02 and 0.74, Supporting Information, Figure 5). That strongly suggests that there is no compelling evidence to choose the more complicated model. We therefore conclude that DB-JBP1, in contrast to JBP1, binds with a single-step mechanism to J-DNA. Moreover, the clear observation of a one-step binding event excludes the possibility that our results indicating a two-step reaction model are due to some kind of instrumental or other experimental artifact.

JBP1 Binding to DNA Involves a Conformational Change. The law of mass action states that the apparent association rate should be proportional to the concentration of the participating molecules, i.e., the rate of association is linearly dependent on the concentration of components. If an apparent binding rate is concentration independent, however, this is suggestive of a conformational change that takes place upon binding. We therefore analyzed the two apparent association rates derived from the double exponential fit of the data obtained for JBP1 binding to J-DNA. While the first, fast, component shows a linear trend as a function of JBP1 concentration, we found that the second, slow, component remains the same regardless of concentration (Figure 4). That suggests a model for DNA binding involving two steps, according to eq 3: JBP1 and DNA first interact in a concentration dependent manner, and subsequently the complex formed undergoes a structural transition.

To analyze our data for their agreement with that model we used the KinTek explorer software.³⁴ This software allows “global fitting”, including experiments designed to show the association, dissociation, and equilibrium simultaneously. Combining an association experiment with a dissociation experiment provides complementary information about the system. In the association experiment protein and DNA are rapidly mixed and the amount of complex formed is measured as a function of time. In the dissociation experiment a

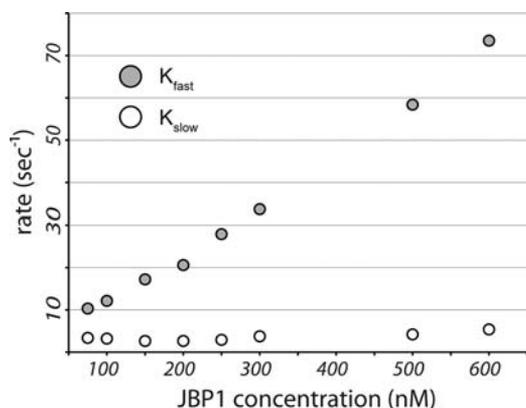


Figure 4. Apparent binding rates of JBP1 to J-DNA deduced from a double exponential fit for each concentration. The fast rate increases linearly with concentration, while the slow rate remains constant, suggesting a conformational change that takes place upon DNA binding.

performed complex with buffer is rapidly diluted and we measure how it falls apart. Combining both experiments at several concentration ratios of protein to DNA is necessary for determining precise association and dissociation rates, especially for a complex two-step system. Inclusion of binding equilibrium data (indicating the amount of complex after a long time, for different ratios of protein to DNA concentration) makes the parameter estimation in global modeling more robust. The fit of the model to the J-DNA and glu-5-hmC-DNA data is presented in Figure 5. The model is well constrained and the kinetic parameters obtained (see Table 2) are not

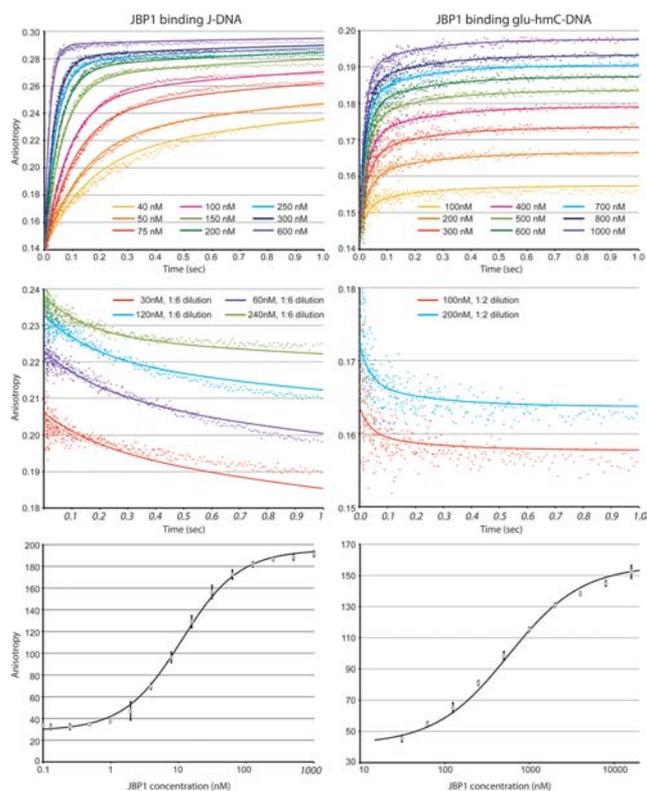


Figure 5. Global fitting of association and dissociation data and equilibrium binding data to the two-step model for JBP1 binding to J-DNA and glu-5-hmC-DNA.

correlated, as shown by the error analysis using KinTek Explorer³⁷ presented in the Supporting Information, Figures 6 and 7.

The initial rate of binding of JBP1 to J-DNA is relatively fast ($k_1 = 73 \mu\text{M}^{-1} \text{s}^{-1}$). Comparison of the rate for the dissociation of this initial complex ($k_{-1} = 2.7 \text{s}^{-1}$) and the rate for the formation of the final complex ($k_2 = 1.1 \text{s}^{-1}$) suggests that in about one of four initial binding events the conformational change takes place. The reverse conformational change event ($k_{-2} = 0.4 \text{s}^{-1}$) takes place at a frequency of about half of that of the change that leads to the final complex formation. Thus, the majority of the complex, about two-thirds, is in the conformation changed configuration at equilibrium conditions, when the protein concentration is significantly higher than the K_D .

Comparing the rates for binding to J-DNA and glu-5-hmC-DNA sheds some light on the mechanistic reasons for the reduced affinity to glu-5-hmC. The first binding event ($k_1 = 28 \mu\text{M}^{-1} \text{s}^{-1}$) is about three times slower than J-binding, but is not the major determinant for the reduced affinity. The rate of dissociation of this complex ($k_{-1} = 22 \text{s}^{-1}$) is about eight times higher than that of the J-DNA complex. Interestingly, the rate at which the conformational change takes place in the glu-5-hmC-DNA complex ($k_2 = 1.5 \text{s}^{-1}$) is similar to that observed upon binding to J-DNA ($k_2 = 1.1 \text{s}^{-1}$), but the rate at which this conformational change is reversed is about ten times as high ($k_{-2} = 4.0 \text{s}^{-1}$ instead of 0.4s^{-1}). That implies that only about one-quarter of the total complex is in the second conformation, at concentrations well above the K_D . In summary, both steps of the binding are affected during binding to glu-5-hmC-DNA compared to J-DNA. The equilibrium constant for the initial complex formation (K_1) is about 20 times lower, and the equilibrium constant for the conformational change (K_2) about 7 times lower for glu-5-hmC-DNA than for J-DNA. The apparent dissociation constant ($K_{D,app}$) calculated from the global fitting of all data is for glu-5-hmC binding $\sim 560 \text{ nM}$, about 50 times lower than that for J-DNA (10 nM). This is in good agreement with the equilibrium measurements alone (Table 1). The interesting implication is that the JBP1:J-DNA complex is mainly in the conformation-changed configuration ($K_2 = 2.5$), whereas the reverse is the case for the glu-5-hmC-DNA complex ($K_2 = 0.4$).

Since we have shown that DB-JBP1 recognition of J-DNA proceeds in a single step, we can also conclude that the conformation change takes place in the full-length JBP1 protein, and not in the DNA. We therefore tested whether this conformational change can be visualized by SANS.

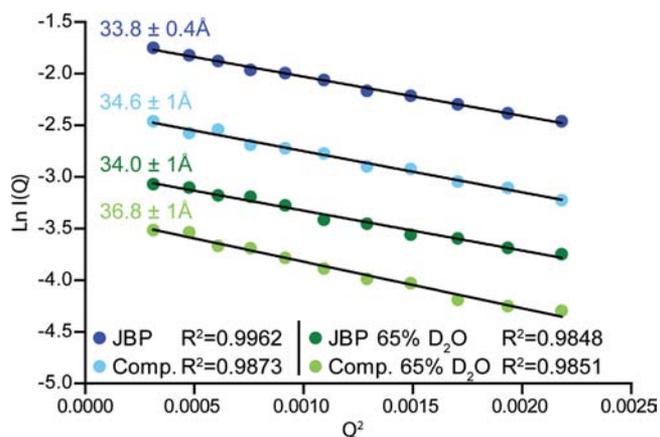
Free JBP1 and JBP1 Bound to glu-5-hmC-DNA Have Different Conformations in Solution. Small-angle neutron scattering (SANS) can provide information about the shape of proteins.^{38,39} By adjusting the $\text{D}_2\text{O}/\text{H}_2\text{O}$ ratio in the solvent it is possible to match the neutron scattering density of different parts of a protein–nucleic acid complex. The contrast match point of a particle is determined from the linear variation of the $\sqrt{I(0)}$ with percent D_2O . Thus protein is matched close to 40% and DNA close to 65% D_2O .³⁸ We used glu-5-hmC-DNA for these experiments, as this is more readily available than J-DNA. By using the $I(0)$ values from the H_2O and 65% D_2O samples (plots not shown), the JBP1 alone sample was found to have a measured match point of 42.5% D_2O ; the JBP1:DNA complex matched at 45% D_2O , as expected from the composition of the 1:1 complex of the corresponding molecular weight of protein (93 kD) and nucleic acid (8.7 kD), with a

Table 2. Kinetic Parameters for JBP1 and DB-JBP1 Binding to J-DNA and glu-5-hmC from the Global Analysis of Transient and Equilibrium Binding Experiments

protein	DNA	k_1 ($\mu\text{M}^{-1} \text{s}^{-1}$)	k_{-1} (s^{-1})	k_2 (s^{-1})	k_{-2} (s^{-1})	K_1 (μM^{-1})	K_2	$K_{D,\text{app}}$ (nM)
JBP1	J	72.80 ± 1.12	2.68 ± 0.09	1.07 ± 0.11	0.42 ± 0.03	27.16	2.52	10.45
JBP1	glu-5-hmC	28.31 ± 1.00	21.85 ± 0.57	1.55 ± 0.22	4.04 ± 0.40	1.29	0.38	560.2

protein match of 42.5% D_2O and a DNA match at 65% D_2O . We define the scattering mass of a component as the scattering density contrast multiplied by the volume of the component.

The Guinier plots (eq 6) of JBP1 alone and in complex with glu-5-hmC-DNA in 100% H_2O and 65% D_2O solvents are shown in Figure 6. In 100% H_2O , the measured radii of

**Figure 6.** Guinier plots for JBP1 and the JBP1:glu-5-hmC-DNA complex in water and in 65% D_2O .

gyration correspond to the scattering mass distributions of protein and DNA. In 65% D_2O the scattering mass of the DNA is close to zero, so that SANS is selectively “blind” to this component of the complex and the measured radius of gyration corresponds to the protein moiety alone. The JBP1 conformation and DNA binding position were determined from the contrast variation data, as exemplified in ref 40.

The measured radii of gyration for JBP1 are 33.8 ± 0.4 Å in H_2O and 34 ± 1 Å in 65% D_2O , for a molecular mass of 93 kD. As, for example, the radius of gyration of tetrameric malate dehydrogenases, which presents a “globular” shape, which is 30 Å for a molecular mass of 130kD,⁴¹ we expect that JBP1 has an elongated conformation (with an approximate axial ratio of about 2:1:1). The radius of gyration of JBP1 in the complex, obtained from the 65% D_2O data where DNA is invisible, was found to be 36.8 ± 1 Å, significantly larger than that of free JBP1, indicating an even more extended conformation.

The radius of gyration value of the complex in H_2O is 34.6 ± 1 Å, but that of JBP1 within the complex (obtained from the 65% D_2O data) is significantly larger, 36.87 ± 1 Å. The significant reduction of radius of gyration value when the DNA is added to the extended protein conformation is a clear indication that the DNA binding takes place close to the center of scattering mass.

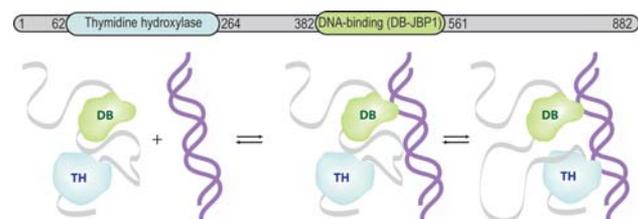
We finally note that even at the high protein and DNA concentration of the SANS experiment ($\sim 10 \mu\text{M}$), 26.5% of the complex is expected to be in the conformationally different configuration (from a simulation using Kintek Explorer). This suggests that the “averaged” changes observed in the SANS experiments are probably an underestimate of the actual conformational changes involved.

DISCUSSION

JBP1 not only binds very tightly to J-DNA, with an affinity of about ten nanomolar, but also has a remarkable ability to discriminate it from normal DNA by about 4 orders of magnitude (7000-fold).^{26,28,32} JBP1 retains some of this discriminatory power in the recognition of glu-5-hmC-DNA, about 200-fold, explaining its successful use as a biotechnological tool.^{29,30} This result is in line with the glucose recognition model that was suggested based on the substitution of the sugar in base J with different O-linked glycosides,²⁸ re-enforcing the notion that the specific recognition of the sugar edge-on conformation is central to base J recognition. The reduced specificity for glu-5-hmC-DNA could be explained by a different conformation of the glucose when linked to 5-hmC. We also note, that as the DNA-binding domain of JBP1, DB-JBP1, retains the high affinity and some of the discriminatory power of its full-length counterpart, it might be a useful tool to exploit in biotechnological applications, as it is easier to produce (more than 20 mg can be obtained from a bacterial culture of 1 L) and is more stable in solution.

The analysis of the pre-steady state kinetics of JBP1 binding to J-DNA (and to other DNA oligonucleotides) revealed a two-step reaction mechanism. As the second apparent binding rate did not vary with protein concentration, we interpreted that as a conformational change that JBP1 undergoes after DNA binding. The conformational change is specific for full-length JBP1, as the DNA-binding domain DB-JBP1 alone did not exhibit that behavior. Moreover, SANS analysis of the shape parameters of JBP1 alone or in complex with DNA supports the concept of a conformational change.

On the basis of these data, we put forward a model (Figure 7) in which the DNA-binding domain of JBP1, DB-JBP1, binds

**Figure 7.** A graphical depiction of the model suggested for the association of JBP1 with J-DNA; DNA-recognition by the DB-JBP1 domain causes a conformational change that allows the TH domain to come in contact with DNA.

to the DNA first; that binding triggers a conformational change in the remainder of the protein that allows the N-terminal part, which harbors the thymidine hydroxylase domain of JBP1 (TH-JBP1), to come in proximity to the DNA. As the two conformations appear to be in a relative equilibrium with several events per second leading from one conformation to another, we envisage that this allows the TH-JBP1 domain to “probe” DNA bases in spatial proximity to the pre-existing J-base, where JBP1 is “anchored” by the DB-JBP1 domain. If the base in contact happens to be a T, then the TH domain could

hydroxylate it. This would explain the known ability of JBP1 to insert base J preferentially in proximity of pre-existing J-bases.

As the DB-JBP1 domain only exists in JBP1, the two-step reaction and the model where the hydroxylase activity is promoted by the presence of pre-existing J-bases only applies to JBP1, but not to JBP2 or to the TET family proteins, which lack the DB-JBP1 domain. As the residence time of the second conformation is shorter in the case of binding to glu-5-hmC-DNA than to J-DNA, specific binding to J-DNA might stimulate the thymidine hydroxylase function of JBP1. Binding to J-DNA is not essential for the conformational change, however, as shown by our observation that binding of JBP1 to T-DNA also occurs by a two-step reaction. Thus the known ability of JBP1 to hydroxylate T's in DNA without J,¹³ albeit suboptimally, is also in-line with this model.

Finally, it needs to be emphasized that the model we put forward does not imply that JBP1 hydroxylates the nearest T-base neighbor in sequence or in space. It is likely that a stochastic component is present and local chromatin structure or other factors might be crucial to translate the bind-and-probe event to productive hydroxylation. We are currently determining the exact position of base-J in the *Leishmania* genome. Once such data become available, possibly together with a crystal structure of full-length JBP1, they can help explain the exact mechanism of base-J insertion by JBP1.

CONCLUSION

JBP1 binds both J-DNA and glu-5-hmC-DNA rapidly and with high affinity and specificity, rationalizing its use as a tool for detecting 5-hmC-DNA in the human epigenome, and offering important hints on how to enhance its usage as a biotechnological tool. JBP1 recognizes DNA with a two-step reaction mechanism that involves a conformational change that was confirmed by shape measurements in solution. This change in conformation might allow the thymidine hydroxylase domain to come into contact with DNA and allow it to hydroxylate specifically bases that are in the neighborhood of pre-existing J, in line with the physiological function of JBP1.

ASSOCIATED CONTENT

Supporting Information

Figure 1 showing the equilibrium binding of DB-D525A-JBP1 to J-DNA and glu-5-hmC-DNA, Figure 2 giving the one- and two-step dissociation models for the binding of JBP1 to J-DNA, Figure 3 giving the one- and two-step dissociation models for the binding of JBP1 to glu-5-hmC-DNA, Figure 4 showing the one- and two-step association models for the binding of JBP1 D525A to T-DNA, Figure 5 giving the one- and two-step association models for the binding of DB-JBP1 to glu-5-hmCDNA, Figure 6 showing the fitspace calculations for the parameters in J-DNA binding, and Figure 7 showing the fitspace calculations for the parameters in glu-5-hmC-DNA binding. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

a.perrakis@nki.nl

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We would like to thank Carlo Petosa and Joanna Timmins for making their FPLC at the IBS, Grenoble available to prepare fresh samples for SANS experiments and for their help; Lioner Porcar for his assistance in using the BL21 beamline at the ILL, Grenoble; and Titia Sixma and the B8 lab members for discussions and for critically reading the manuscript.

REFERENCES

- (1) Cross, M.; Kieft, R.; Sabatini, R.; Wilm, M.; de Kort, M.; van der Marel, G. A.; van Boom, J. H.; van Leeuwen, F.; Borst, P. *EMBO J.* **1999**, *18*, 6573.
- (2) DiPaolo, C.; Kieft, R.; Cross, M.; Sabatini, R. *Mol. Cell* **2005**, *17*, 441.
- (3) Kieft, R.; Brand, V.; Ekanayake, D. K.; Sweeney, K.; DiPaolo, C.; Reznikoff, W. S.; Sabatini, R. *Mol. Biochem. Parasitol.* **2007**, *156*, 24.
- (4) Vainio, S.; Genest, P. A.; ter Riet, B.; van Luenen, H.; Borst, P. *Mol. Biochem. Parasitol.* **2009**, *164*, 157.
- (5) Yu, Z.; Genest, P. A.; ter Riet, B.; Sweeney, K.; DiPaolo, C.; Kieft, R.; Christodoulou, E.; Perrakis, A.; Simmons, J. M.; Hausinger, R. P.; van Luenen, H. G.; Rigden, D. J.; Sabatini, R.; Borst, P. *Nucleic Acids Res.* **2007**, *35*, 2107.
- (6) Gommers-Ampt, J. H.; Van Leeuwen, F.; De Beer, A. L. J.; Vliegthart, J. F. G.; Dizdaroglu, M.; Kowalak, J. A.; Crain, P. F.; Borst, P. *Cell* **1993**, *75*, 1129.
- (7) Van Leeuwen, F.; Taylor, M. C.; Mondragon, A.; Moreau, H.; Gibson, W.; Kieft, R.; Borst, P. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2366.
- (8) Toaldo, C. B.; Kieft, R.; Dirks-Mulder, A.; Sabatini, R.; van Luenen, H. G.; Borst, P. *Mol. Biochem. Parasitol.* **2005**, *143*, 111.
- (9) Borst, P.; Sabatini, R. *Annu. Rev. Microbiol.* **2008**, *62*, 235.
- (10) van Leeuwen, F.; Wijsman, E. R.; Kuyl-Yeheskiely, E.; van der Marel, G. A.; van Boom, J. H.; Borst, P. *Nucleic Acids Res.* **1996**, *24*, 2476.
- (11) Van Leeuwen, F.; Kieft, R.; Cross, M.; Borst, P. *Mol. Biochem. Parasitol.* **2000**, *109*, 133.
- (12) Cliffe, L. J.; Kieft, R.; Southern, T.; Birkeland, S. R.; Marshall, M.; Sweeney, K.; Sabatini, R. *Nucleic Acids Res.* **2009**, *37*, 1452.
- (13) Cliffe, L. J.; Siegel, T. N.; Marshall, M.; Cross, G. A.; Sabatini, R. *Nucleic Acids Res.* **2010**, *38*, 3923.
- (14) Ekanayake, D. K.; Minning, T.; Weatherly, B.; Gunasekera, K.; Nilsson, D.; Tarleton, R.; Ochsenreiter, T.; Sabatini, R. *Mol. Cell. Biol.* **2011**, *31*, 1690.
- (15) van Luenen, H.; Farris, C.; Jan, S.; Genest, P. A.; Tripathi, P.; Velds, A.; Kerkhoven, R. M.; Nieuwland, M.; Haydock, A.; Ramasamy, G.; Vainio, S.; Heidebrecht, T.; Perrakis, A.; Pagie, L.; van Steensel, B.; Myler, P. J.; Borst, P. *Cell* **2012**, in press.
- (16) Kriacionis, S.; Heintz, N. *Science* **2009**, *324*, 929.
- (17) Tahiliani, M.; Koh, K. P.; Shen, Y.; Pastor, W. A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Iyer, L. M.; Liu, D. R.; Aravind, L.; Rao, A. *Science* **2009**, *324*, 930.
- (18) Gu, T. P.; Guo, F.; Yang, H.; Wu, H. P.; Xu, G. F.; Liu, W.; Xie, Z. G.; Shi, L.; He, X.; Jin, S. G.; Iqbal, K.; Shi, Y. G.; Deng, Z.; Szabo, P. E.; Pfeifer, G. P.; Li, J.; Xu, G. L. *Nature* **2011**, *477*, 606.
- (19) Guo, J. U.; Su, Y.; Zhong, C.; Ming, G. L.; Song, H. *Cell* **2011**, *145*, 423.
- (20) Iqbal, K.; Jin, S. G.; Pfeifer, G. P.; Szabo, P. E. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 3642.
- (21) Ito, S.; D'Alessio, A. C.; Taranova, O. V.; Hong, K.; Sowers, L. C.; Zhang, Y. *Nature* **2010**, *466*, 1129.
- (22) Ito, S.; Shen, L.; Dai, Q.; Wu, S. C.; Collins, L. B.; Swenberg, J. A.; He, C.; Zhang, Y. *Science* **2011**, *333*, 1300.
- (23) Ko, M.; Bandukwala, H. S.; An, J.; Lamperti, E. D.; Thompson, E. C.; Hastie, R.; Tsangaratou, A.; Rajewsky, K.; Koralov, S. B.; Rao, A. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 14566.
- (24) Ko, M.; Huang, Y.; Jankowska, A. M.; Pape, U. J.; Tahiliani, M.; Bandukwala, H. S.; An, J.; Lamperti, E. D.; Koh, K. P.; Ganetzky, R.;

Liu, X. S.; Aravind, L.; Agarwal, S.; Maciejewski, J. P.; Rao, A. *Nature* **2011**, *468*, 839.

(25) Iyer, L. M.; Tahiliani, M.; Rao, A.; Aravind, L. *Cell Cycle* **2009**, *8*, 1698.

(26) Sabatini, R.; Meeuwenoord, N.; van Boom, J. H.; Borst, P. *J. Biol. Chem.* **2002**, *277*, 958.

(27) Sabatini, R.; Meeuwenoord, N.; van Boom, J. H.; Borst, P. *J. Biol. Chem.* **2002**, *277*, 28150.

(28) Grover, R. K.; Pond, S. J.; Cui, Q.; Subramaniam, P.; Case, D. A.; Millar, D. P.; Wentworth, P., Jr. *Angew. Chem., Int. Ed. Engl.* **2007**, *46*, 2839.

(29) Robertson, A. B.; Dahl, J. A.; Ougland, R.; Klungland, A. *Nat. Protoc.* **2012**, *7*, 340.

(30) Robertson, A. B.; Dahl, J. A.; Vagbo, C. B.; Tripathi, P.; Krokan, H. E.; Klungland, A. *Nucleic Acids Res.* **2011**, *39*, e55.

(31) Robertson, J.; Robertson, A. B.; Klungland, A. *Biochem. Biophys. Res. Commun.* **2011**, *411*, 40.

(32) Heidebrecht, T.; Christodoulou, E.; Chalmers, M. J.; Jan, S.; Ter Riet, B.; Grover, R. K.; Joosten, R. P.; Littler, D.; van Luenen, H.; Griffin, P. R.; Wentworth, P., Jr.; Borst, P.; Perrakis, A. *Nucleic Acids Res.* **2011**, *39*, 5715.

(33) Stricher, F.; Martin, L.; Barthe, P.; Pogenberg, V.; Mechulam, A.; Menez, A.; Roumestand, C.; Veas, F.; Royer, C.; Vita, C. *Biochem. J.* **2005**, *390*, 29.

(34) Johnson, K. A.; Simpson, Z. B.; Blom, T. *Anal. Biochem.* **2009**, *387*, 20.

(35) Johnson, K. A. *Methods Enzymol.* **2009**, *467*, 601.

(36) Serdyuk, I. N.; Zaccai, N.; Zaccai, G. *Methods in Molecular Biophysics: Structure, Dynamics, Function*; Cambridge University Press, Cambridge, UK: 2007.

(37) Johnson, K. A.; Simpson, Z. B.; Blom, T. *Anal. Biochem.* **2009**, *387*, 30.

(38) Jacrot, B. *Rep. Prog. Phys.* **1976**, *39*, 911.

(39) Stuhmann, H. B. *Acta Crystallogr., Sect. A* **2008**, *64*, 181.

(40) Dessen, P.; Blanquet, S.; Zaccai, G.; Jacrot, B. *J. Mol. Biol.* **1978**, *126*, 293.

(41) Bonneté, F.; Ebel, C.; Zaccai, G.; Eisenberg, H. *J. Chem. Soc., Faraday Trans.* **1993**, *89*, 2659.