PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Insights into the DNA cleavage mechanism of human LINE-1 retrotransposon endonuclease

Kostas Repanas,[1] Gloria Fuentes,[2] Serge X. Cohen,[1] Alexandre M. J. J. Bonvin,[2] and Anastassis Perrakis[1]*

[1] Division of Molecular Cancer Biology, The Netherlands Cancer Institute, 1066 CX Amsterdam, The Netherlands
[2] Bijvoet Center for Biomolecular Research, NMR Research Group Utrecht University, 3584 CH Utrecht, The Netherlands

ABSTRACT

The human LINE-1 endonuclease (L1-EN) contributes in defining the genomic integration sites of the abundant human L1 and Alu retrotransposons. LINEs have been considered as possible vehicles for gene delivery and understanding the mechanism of L1-EN could help engineering them as genetic tools. We tested the *in vitro* activity of point mutants in three L1-EN residues—Asp145, Arg155, Ile204—that are key for DNA cleavage, and determined their crystal structures. The L1-EN structure remains overall unaffected by the mutations, which change the enzyme activity but leave DNA cleavage sequence specificity mostly unaffected. To better understand the mechanism of L1-EN, we performed molecular dynamics simulations using as model the structures of wild type EN-L1, of two βB6-βB5 loop exchange mutants we have described previously to be important for DNA recognition, of the R155A mutant from this study, and of the homologous TRAS1 endonuclease: all confirm a rigid scaffold. The simulations crucially indicate that the βB6-βB5 loop shows an anticorrelated motion with the surface loops βA6-βA5 and βB3-αB1. The latter loop harbors N118, a residue that alters DNA cleavage specificity in homologous endonucleases, and implies that the plasticity and correlated motion of these loops has a functional importance in DNA recognition and binding. To further explore how these loops are possibly involved in DNA binding, we docked computationally two DNA substrates to our structure, one involving a flipped-out nucleotide downstream the scissile phosphodiester; and one not. The models for both scenarios are feasible and agree with the hypotheses derived from the dynamic simulations. The reduced cleavage activity we have observed for the I204Y mutant above however, favors the flipped out nucleotide model.

Proteins 2009; 74:917–928.

Key words: protein-DNA recognition; protein flexibility; X-Ray crystallography; computational docking; MD simulations.

INTRODUCTION

L1 retrotransposons are mobile long interspersed nuclear elements (*LINE-1*) that lack retroviral-like long terminal repeats and are able to copy and paste themselves from one genomic location to the other with the help of an endonuclease-reverse transcriptase self-encoded activity. *L1*s propagate via an RNA intermediate and have created more than 17% of the mass of the human genome, whereas *Alu* elements that likely propagate by hijacking the *L1* retrotransposition mechanism, account for an additional 11% of the genome.[1–3] *L1* elements have a short loose genomic integration site consensus sequence 5′-TTTT/AA-3′[4–6] that is nicked by their targeting endonuclease,[7,8] to free a 3′ OH group and prime reverse transcription.[9] The human *LINE-1*s are semi-specific for the target they recognize and integrate into, which differentiates them from elements of other organisms that exhibit higher specificity. Mobile elements like *R1Bm* from *Bombyx mori*[10,11] and *Tx1L* from *Xenopus laevis*[12] encode highly specific targeting endonucleases.[13,14] Each of these elements integrates into unique genomic locations, a specific sequence within 28S rDNA for *R1Bm* and within DNA transposon *Tx1D* for *Tx1L*.[15] The respective endonuclease largely determines the integration site. Previously, we described the crystal structure of the human L1 endonuclease L1-EN,[16] a member of the family of metal-dependent phosphohydrolases.[17]

To focus our understanding of how the L1-EN structure infers its functional properties and the role of point-mutants that affect the actual catalysis step, we here chose to study further the robustness and plasticity of the L1-EN catalytic scaffold and surface elements. Flexibility in protein systems can constitute a major component in creating their specificity and catalytic power, hence leading to the creation of a functional entity.

Additional Supporting Information may be found in the online version of this article.
"Gloria Fuentes's current address is Spanish National Cancer Research Centre (CNIO), Structural Computational Biology Group, E-28029 Madrid, Spain."
Grant sponsor: Netherlands Organization for Scientific Research (NWO); Grant numbers: NWO-CW 700.51.012, 700.56.442.
*Correspondence to: Anastassis Perrakis, The Netherlands Cancer Institute, Division of Molecular Carcinogenesis, 1066 CX Amsterdam, The Netherlands. E-mail: a.perrakis@nki.nl
Received 26 February 2008; Revised 5 June 2008; Accepted 25 June 2008
Published online 2 September 2008 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/prot.22201

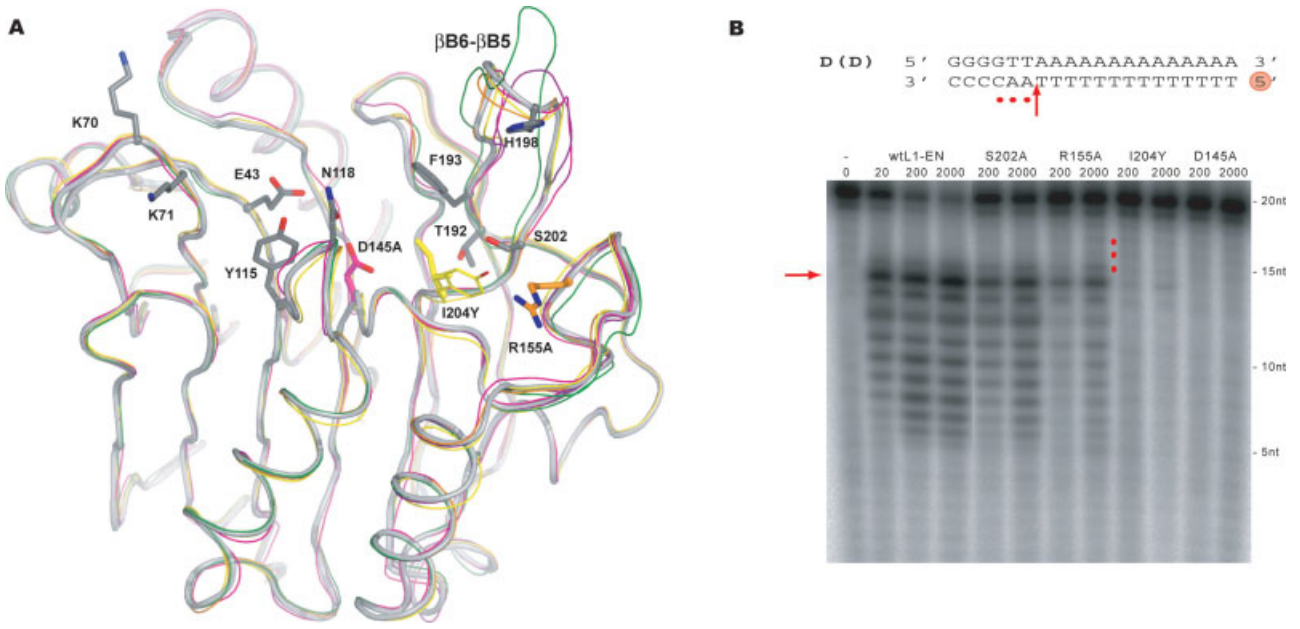

Figure 1

A: Ribbon representation of L1 endonuclease with key residues discussed in the text shown as ball and stick models. I204, R155, and S202 are thought to form a pocket for the accommodation of a flipped-out nucleotide. Residues that have been mutated are color coded: D145A-magenta, R155A-orange, and I204Y-yellow. Same colored ribbons of the respective variant structures are superposed on wtL1-EN. Ribbons of loop exchange variants LTx (green) and LR1 (purple) are also superposed for comparison. T192 and S202 are the anchoring residues of loop βB6-βB5 that also served as reference points in the loop exchange mutants (Table II). Oxygen: red, Nitrogen: blue. **B**: Activity and target recognition of L1-EN point mutants. 20mer DNA substrate duplex containing the poly(T)-A junction. Red Circle: 5′ end (labeled) of the substrate strand. DNA duplex (180 nM) was titrated with increasing concentrations of L1-EN and L1-EN variants. Products were analyzed on autoradiographs of denaturing polyacrylamide gels. Red arrow: poly(T)-A junction. Red dots: longer products downstream of the main nicking site.

Molecular dynamics simulations can reveal the role of protein flexibility in structure and mechanism, especially when carried out at the atomic level. Catalytic processes in enzymes exploit the energy available from local and global molecular motions just as much as from the particular chemical environment.[18] Many different theoretical methods exist for the description of protein flexibility, but molecular dynamics (MD) is probably the most powerful.[19,20] Since the very first applications to proteins in the late 1970s, MD has been successfully used to study the dynamic behavior of various proteins.[21,22] It is generally believed that protein behavior during MD simulations most likely reflects physically meaningful processes.[23]

In the enduring absence of an experimental L1-EN:DNA complex structure, and to promote our understanding of how L1-EN is recognizing target DNA, we use the computational docking program HADDOCK (High Ambiguity Driven DOCKing[24]). HADDOCK can incorporate available information, such as experimental and bioinformatics data, in order to drive the docking process.[25] Its use in a variety of cases predicting protein:DNA or protein:RNA complexes has been previously reported.[26–29] HADDOCK further allows for the introduction of protein flexibility, firstly on side-chains of the interface and then for both backbone and side-chains.

The structure of wild type L1-EN[16] comprises a rigid core that sculpts the surface of the likely DNA binding site and a prominent protruding loop [βB6-βB5, Fig. 1(A)]. We have speculated that catalysis may proceed with an APE1-like mechanism with a flipped out nucleotide. This was illustrated by a manually built model constructed using the positions of two sulfate ions and the scissile phosphate position as anchor points for the placement of substrate DNA. This model together with sequence conservation indicated that R155, S202, and I204 are forming a pocket that accommodates a flipped out adenine and was used as a guide for the mutagenesis experiments presented in this manuscript. In contrast to DNA repair enzymes APE-1[30] in humans and its' *Escherichia coli* counterpart Exonuclease III,[31] L1-EN can provide sufficient space for a flipped out nucleotide. The turned ribose could rest on I204 as in APE-1, whereas the semi-conserved R155 and S202 could provide base specific hydrogen bonds with the nucleotide. We previously exchanged the surface exposed βB6-βB5 hairpin loop of L1-EN for the ones of R1Bm and Tx1L endonucleases and determined the corresponding crystal structures [LR1 and LTx, Fig. 1(A), magenta and green ribbons]. Biochemical and structural analysis of the loop chimeras showed that the loop is a major specificity ele-

Table I
Summary of Mutants Activity and Specificity Profiles

Name	Mutation type	New residues	3D structure	Activity	Affects specificity
D145	Point	A	+	−	n/a
R155	Point	A	+	++	−
S202	Point	A	−	+++	−
I204	Point	Y	+	+	−
LTx	Loop-exchange	YVRVRDGHVSQ	+	++	+
LR1	Loop-exchange	FSTANGE	+	+	+
L1-EN	Wild-type	n/a	+	+++++	n/a

ment. We have tested the hypothesis that the prominent βB6-βB5 hairpin loop probably inserts into the DNA minor groove and may be a particularly important surface element involved in DNA target recognition. Loop grafting was indeed shown to alter the sequence specificity of L1-EN *in vitro* and of L1 integration specificity in cell cultures.[32] The readout of the sequence appears to be "indirect" and L1 most likely recognizes structural features inferred by the sequence and not the sequence itself.

Here we report the *in vitro* function and crystal structures of point mutants of EN-L1: D145A, the main catalytic residue that renders the protein inactive; I204Y that has implications to the accommodation of a speculated flipped-out DNA base during cleavage; and that of R155A and S202A that are in a position to fix the flipped out nucleotide. Studying the dynamic properties of various L1-EN mutant structures via MD simulations (of wild type L1-EN, two L1-EN loop graft mutants LTx and LR1, the R155A mutant, and that of retrotransposon-encoded TRAS1-EN[33]), we propose how the L1-EN activity is likely to be modulated by dynamic structural properties. To utilize and explain further observations from the dynamic simulations and the new structural results we describe, we construct two computational models for "normal" and "flipped-out" DNA complexes and discuss their likelihood taking into account all available data.

RESULTS AND DISCUSSION

Active site and peripheral L1-EN mutants affect cleavage activity but not specificity or structure

A summary of functional and structural data for all mutants used in this study is available as Table I. Structural comparisons are summarized in Supplementary Table I.[34]

Aspartate 145 is a totally conserved residue in the family of metal-dependent phosphohydrolases. This mutation totally abolishes activity [Fig. 1(B)] while no structural differences occur [Figs. 1(A) and 2; purple model] despite being situated at the heart of the L1 endonuclease. These observations confirm that D145 is absolutely required for the actual catalytic event.[30]

The isoleucine at position 204 was mutated to tyrosine to mimic the L1 homolog bovine pancreatic deoxyribonuclease I, DNaseI.[35] DNAseI-like enzymes form a small group of metal-dependent phosphohydrolases that display conservation in this tyrosine, which is usually a small hydrophobic residue among other family members. DNAseI binds DNA in a nonspecific manner and has not been shown to involve recognition of an extra-helical nucleotide for catalysis to proceed, like APE.[30] Based on the model we have previously manually constructed,[16] the isoleucine at position 204 provides enough space for accommodating the nucleotide adjacent to the scissile phosphate in a possible flip-out mechanism, as in the model for APE1 catalysis. Should the longer and bulkier tyrosine moiety assume a relative position similar to the one in DNAseI active site, it should obscure activity. Indeed the I204Y mutation has a dramatic effect in catalysis although residual activity can still be detected and compared to the totally inactive D145 mutation [Fig. 1(B)]. However, the overall structure is only moderately affected and the bulkier and more hydrophilic tyrosine is accommodated rather well in the L1-EN scaffold [Figs. 1(A) and 2; yellow model]. Because I204 is rather far from the actual catalytic residues and an isoleucine cannot participate in actual catalysis, the only reasonable ex-

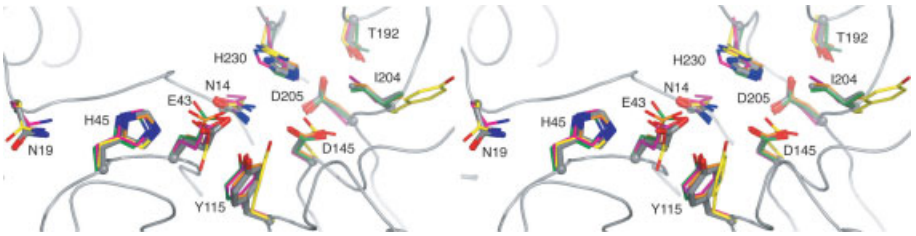

Figure 2
Stereo superposition of the L1-EN active site and of the five variants discussed in the text. Gray: L1-EN, purple: LR1, green: LTx, magenta: D145A, orange: R155A and yellow: I204Y. Oxygen: red, Nitrogen: blue.

planation is that indeed the tyrosine is not tolerated well due to steric hindrance upon DNA binding. Because wtL1-EN is only moderately specific demonstrating the characteristic laddering pattern in Figure 1(B), it is unlikely that the I204A mutation results in complete loss of specificity in our assay and subsequent loss of only apparent activity, that is further supported by cleavage of supercoiled plasmid DNA (data not shown).

Arginine 155 could hydrogen bond to an extra-helical nucleotide, if L1-EN follows the APE1 model and recognizes a flipped out base. The structure of the R155A mutant remains virtually unaffected [Figs. 1(A) and 2; orange model], while activity is reduced [Fig. 1(B)]. The R155A is significantly more active than the I204Y mutant: that is compatible with the hypothesis that isoleucine is crucial for allowing space for the flipped out nucleotide, whereas the arginine would only be needed to stabilize the base for more efficient catalysis. That also explains why sequences with nucleotides other than adenine downstream of the scissile bond can be accommodated during cleavage at that position, although less efficiently[32]: possibly, bases other than adenine cannot form efficient hydrogen bonds with the arginine side chain and catalysis is thus less efficient. Finally, the R155A mutation affects only nicking activity but has little apparent effect on specificity; if anything the observation of the slightly relatively more pronounced longer cleavage products outside the A-tract are compatible with our hypothesis for its role [Fig. 1(B), red dots].

The activity of the S202A mutant is less affected [Fig. 1(B)], which is consistent with a lesser role in fixing the flipped out nucleotide. The hydrogen bond possibly contributed by S202 to fixing the flipped-out nucleotide is thus important, but not as crucial as the hydrogen bond contributed by R155. This also can explain why R155 is important but not strictly essential for catalysis to proceed, since its role in possibly fixing the flipped out base can be partially realized by S202.

Structures of L1-EN mutants provide a structurally robust scaffold with conserved binding features

Crystals of R155A diffracted X-rays to 2.0 Å, those of I204Y to 2.2 Å, and the ones of D145A to 2.3 Å. All three crystals belong to space group $P2_1$, contain four copies of each protein in the asymmetric unit and are twinned (twin law −h, −k, l). Details for data collection, treatment for twinning, and refinement statistics are summarized in Table II. The structures were solved by molecular replacement using L1-EN as a search model; the mutated residues could be clearly confirmed and identified in the electron density maps. The three new structures presented here, combined with the wild type L1-EN structure and those of the two hairpin variants LTx and LR1 we have previously reported,[32] enable us to take a much more detailed look in and around the active site of the L1 endonuclease.

Table II
Data Collection and Refinement Statistics

Data collection	I204Y	R155A	D145A
Resolution, Å	2.2	2.0	2.4
Cell a,b,c (Å)	42.6, 93.8, 126.6	43.1, 93.6, 125.9	43.2, 93.5, 125.7
α,β,γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Space group	$P2_1$	$P2_1$	$P2_1$
R_{merge}, %	7.6 (48.2)	8.0 (55.3)	9.2 (43.7)
Completeness, %	86.2 (88.6)	97.3 (96.5)	84.1 (87.0)
I/σ(I)	6.8 (1.5)	12.1 (2.5)	9.3 (2.0)
No. of reflections			
Unique observed	46416	65720	37199
Total measured	62476	248082	78085
Twinning			
Operator	−h, −k, l	−h, −k, l	−h, −k, l
Twin fraction	0.45	0.1	0.29
Refinement			
R_{cryst}, %	14.2	17.5	17.3
R_{free}, %	20.8	24.7	25.6
Number of:			
molecules in AU	4	4	4
atoms	8198	8082	7930
ions	6	9	4
water molecules	435	389	430
Ramachandran plot			
Most favored regions, %	89.6	91.6	82.1
Allowed regions, %	10.4	8.2	17.1
Generously allowed, %	–	0.2	0.8
Rmsd			
Bond lengths, Å	0.003	0.006	0.007
Bond angles, °	0.55	0.77	0.86

Looking at the active site regions of all structures collectively (see Fig. 2), most of the induced differences apart from the mutations themselves seem to concentrate on E43 and H230. The glutamate has been shown to coordinate a Mg^{+2} ion essential for catalysis, in many metal-dependent phosphohydrolases. The histidine is likely to create the attacking nucleophile as shown for APE-1.[30] Comparing the overall positioning with respect to the wild type L1-EN structure, the structure of the I204Y variant is the one exhibiting the largest changes, besides the prominent substitution of isoleucine to tyrosine. It is not straightforward to explain these differences because the tyrosine that replaces the isoleucine points away from the active site.

It is notable that Mg^{+2} ions were necessary for crystallization to proceed in all cases; but Mg^{+2} could not be identified in any of these structures, with the exception of LR1. In the LR1 structure, we confirmed the metal position by substituting Mg^{+2} for Mn^{+2} and collecting anomalous X-ray diffraction data at 1.7 Å wavelength. The presence of a metal position was clearly confirmed, but despite extensive effort we were unable to model it accurately, due to dynamic or static disorder that manifested itself as a "double" anomalous peak. Because of this lack of conclusive data, we are not discussing this specific structure and possible implications of the metal positioning in catalysis.

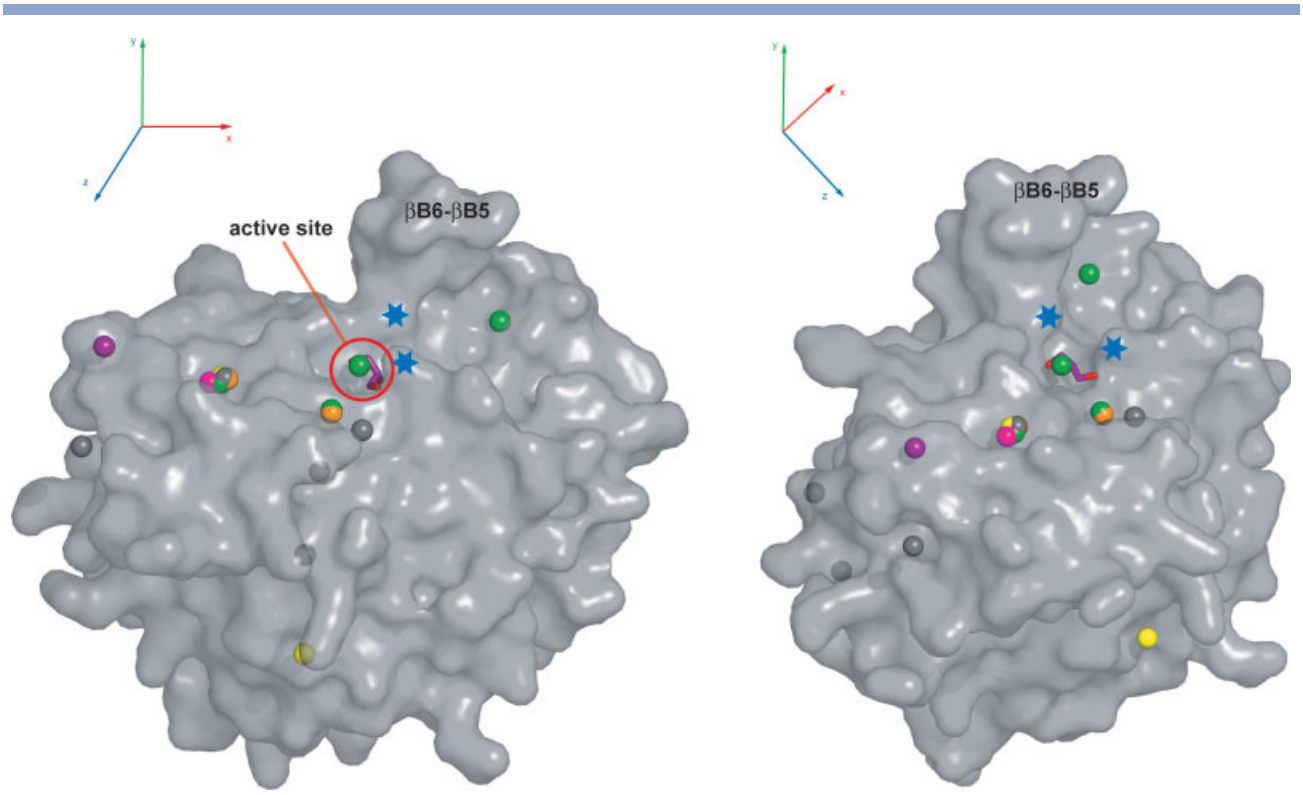

Figure 3

Surface representation of L1-EN with mapped superposed all sulfate ions as spheres and a glycerol as a stick model, from the six available crystal structures. Gray: L1-EN, purple: LR1, green: LTx, magenta: D145A, orange: R155A, and yellow: I204Y. The active site is indicated by a red circle. Blue asterisks mark the narrow channel formed by N118 and F193 above the active site. The βB6–βB5 hairpin is labeled to aid orientation and front and side views of the protein are shown.

Collectively, the superposition of the structures, with the various positions that residues can adopt, illustrates the plasticity of this region and hints to breathing motions while L1-EN is scanning and adapting the surface depending on the DNA substrate.

All crystallization conditions contained sulfate ions (ammonium sulfate was used together with polyethylene glycol as a precipitant) and the cryo-protection medium contained glycerol. In the three mutant structures we present here as well as in the previously solved structures of wtL1-EN, and the hairpin exchange variants LTx and LR1, both sulfate ions and glycerol molecules could be clearly located. After superimposition of all available structures (see Fig. 3), it can be seen that many of these ligands occupy common sites. The most frequently occupied position for a sulfate ion is between His45 and Asn19; all six available structures bind an ion at this position. Three structures (LTx, R155A, and L1-EN) have a sulfate ion between the "other" side of His45 and Tyr115. Other binding sites are clear but not consistent between the structures. In LTx, Arg155 binds a sulfate ion that could indicate a stabilizing role contacting a DNA phosphate, or the extra-helical nucleotide in the flipped-out base model, that are both discussed later. In LTx, we also find a sulfate ion in the active site of the protein, coordinated by His230, Asp145, Glu43, and Tyr193 (that replaces the wild type Phe193 present in L1-EN). LR1 also binds a glycerol moiety, which is located straight in the active site; the glycerol hydrogen bonds to His230, Tyr115, Asp145, Asn147; it is also in contact with Glu43 via the bound—albeit not clearly modeled—metal ion. The glycerol moiety superimposes well onto the DNA backbone phosphate of the APE1 complex and the LTx sulfate ion (Figs. 2 and 3).

Collectively the analysis of all biochemical assays and structures argue that the active site mutants presented here are in sharp contrast with the activity profile of the βB6-βB5 loop exchange variants LTx and LR1 we have previously described, which affect the specificity of cleavage (Table I). The dynamical character of this loop region was initially demonstrated with a normal mode analysis.[32] We use here a molecular dynamics approach to further analyze at atomic detail the role of flexibility in the structure and activity of L1-EN.

Molecular dynamics simulations: connecting flexibility and functionality

To further understand the functional role of loop plasticity and flexibility and its importance for target DNA

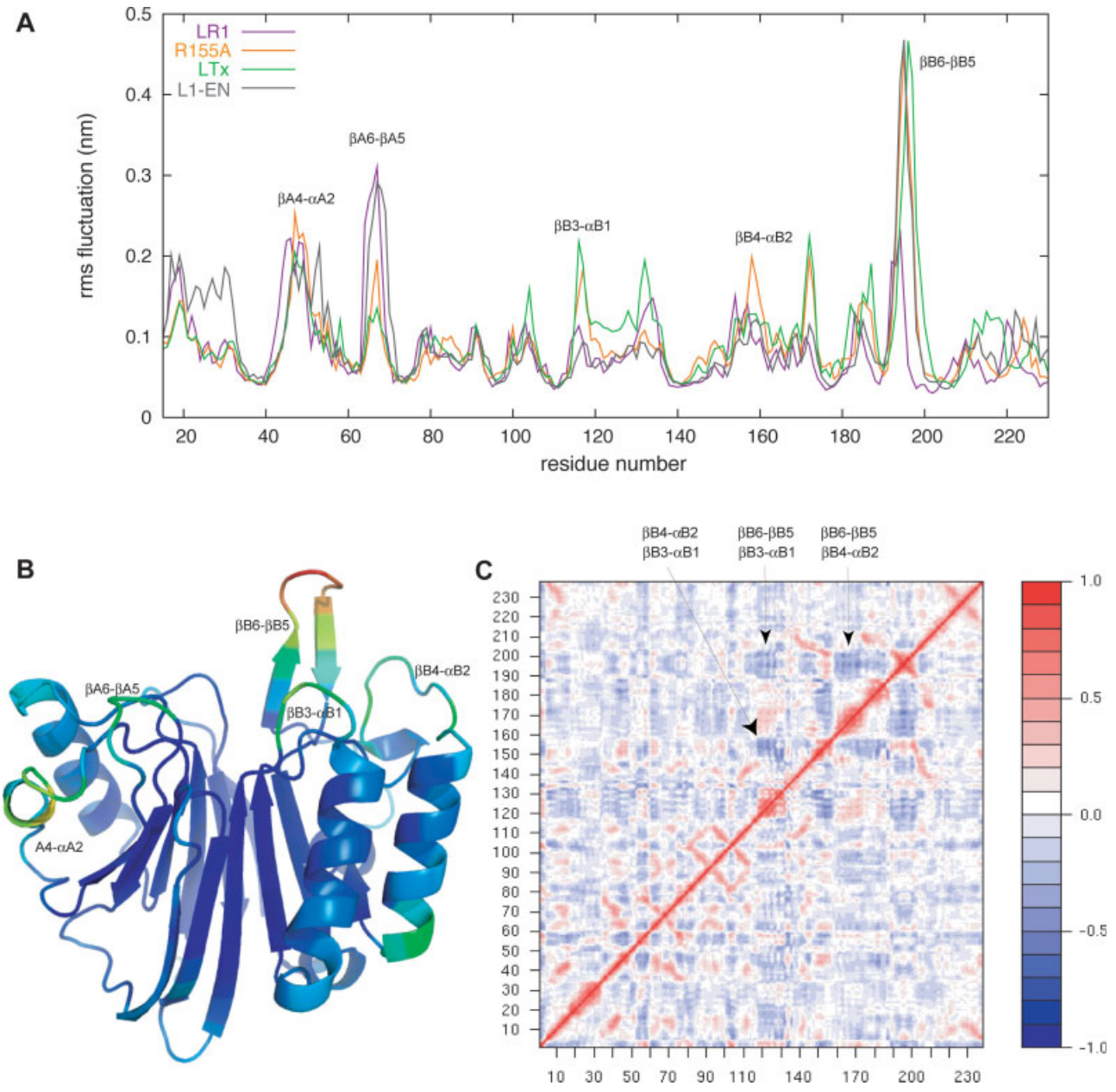

Figure 4

A: RMS fluctuations calculated over the last 6 ns of the 10 ns MD simulations of L1-EN, LTx, LR1, and R155A. **B**: Cartoon representation of L1-EN colored according to RMSF (0.0–0.5) values of R155A trajectory as an example. Flexible areas are named and correspond to the ones in A. **C**: The normalized covariance (C_{ij} cross-correlation) matrix for the Ca atomic fluctuations observed in the trajectory of LTx.

recognition, binding and cleavage, and at the same time confirm that the L1-EN scaffold is rigid as suggested by the mutant structures, we performed classical molecular dynamics simulations. A total of five 10 nanosecond simulations were performed on: wtL1-EN, LTx, LR1, R155A and, on the L1-EN homolog TRAS1-EN, the sequence-specific APE-type non-LTR retrotransposon endonuclease that cleaves telomeric repeats of insects.[33] Standard analysis tools indicate that our simulations are stable and show consistent results along all experiments. We show a plot of the average backbone root mean square fluctuation (RMSF) that illustrates the motion of each residue in the simulations [Fig. 4(A)]; the same fluctuations have been color-coded in a ribbon diagram of the structure of L1-EN [(Fig. 4(B)]. We have also performed principal component analysis (PCA) to reduce the dimensionality (Suppl. Fig. 2). In this type of analysis, motions that coincide on the same eigenvector indicate that given

loops move in a paired fashion. Additional information is revealed in the normalized covariance (C_{ij} cross-correlation) matrices for the Ca atomic fluctuations [the LTx matrix is shown in Fig. 4(C)], where red bands indicate correlated motion (e.g. along the diagonal for neighboring residues) while blue bands are indicative of anticorrelated motions.

The RMSF plots [Fig. 4(A)] show that the βB6-βB5 β-hairpin is much more rigid in the relatively inactive LR1 when compared with the flexible hairpins on L1-EN, LTx, and R155A. For L1-EN and R155A, the plots practically superimpose as expected because the hairpin residues are identical in both cases. These observations confirm and quantify our previous hypothesis about loop flexibility.[32] Additional regions in the presumed DNA binding region that have an elevated degree of flexibility are also evident: the βA4-αA2 loop near E43 and H45 (which are involved in ion coordination) and the βA6-βA5 loop near K70-K71 (which possibly contacts DNA) are most prominent in L1-EN and consistent in the other simulation.

The small βB3-αB1 loop is clearly mobile in the R155 and LTx simulations as well as in the TRAS1-EN simulation. That loop contains residue N118, which, when mutated to alanine in TRAS1-EN, retains activity but looses specificity for telomeric targets, possibly pointing towards major groove interactions.[33] N118 aligns with Y98 in the R1Bm-EN structure (PDB code: 2EI9) and was identified as the only residue outside the hairpin region altering target sequence specificity.[11] Moreover, the βB3-αB1 loop that harbors N118 shows anticorrelated motion to the βB6-βB5 hairpin loop [Fig. 4(C)]. The maximal distance between these loops varies between 5–10 Å. This is more pronounced in LTx (where the βB6-βB5 loop is more mobile) and less in R155A and wtL1-EN.

In the R155A mutant trajectory, the loop around the mutation (βB4-αB2) is more flexible than in other EN-L1 simulations. The salt-bridge between the nearby D154 and R182 (conserved and restricted to AP DNA repair endonucleases and mammalian-type L1 endonucleases[16]) appears to be destabilized by that mutation [Suppl. Fig. 1(A)]. The increased mobility could result in the reduced activity that this mutant exhibits, even if it is not directly responsible for binding a flipped out DNA base. In TRAS1-EN, the equivalent of the D154-R182 salt-bridge is missing and the βB4-αB2 loop is in a region of high mobility in the wild type TRAS1-EN [Suppl. Fig. 1(B)]. This further suggests that this peripheral loop, almost 17 Å away from the protein's active site, has an important role in aiding function and modulating activity, whereas as seen in both L1 and TRAS1 endonucleases it has no apparent effect on protein specificity.

Finally, as shown in Figure 4(C), the βB6-βB5 β-hairpin and the loops βB4-αB2 and βB3-αB1 that have been discussed earlier, all show anticorrelated motions suggesting a pattern of opening/closing that, when seen in the context of the structure [Fig. 4(B)], could indicate a "breathing" motion that can accommodate the DNA.

We proceeded by creating a computational model of L1-EN binding to DNA, using minimal knowledge bias, in an attempt to decipher how correlated or anticorrelated surface loop motions could influence DNA binding and to test the hypothesis of the flip-out mechanism.

Table III
Definition of the Ambiguous Interaction Restraints (AIRs) for the Two HADDOCK Runs

Residue number	
L1-EN	
Active	N19, H45, K70, R155, F193, H198, S202, I204
Passive	N16, S20, R53, K71, N118, T119, N147, T157, Q159, A196, P197, H199, D229
DNA	
Active	T6, A7
Passive	1–5, 8–24

Computational docking of an A-tract DNA substrate on L1-EN driven by mutational data

The L1 endonuclease could follow two main modes for binding and cleaving its DNA substrate. The first mode would be similar to DNaseI, where the enzyme cleaves DNA without the need to flip a nucleotide. The second mode, would be similar to the mode of action of APE1, and would involve a flipped-out nucleotide. We have previously constructed models of L1-EN:DNA complexes manually.[16] In the absence of an experimentally determined structure of an L1-EN:DNA complex, we decided to extend our manual docking studies and model the complex computationally, using the HADDOCK software[24] with suitable DNA substrates.[36]

The first DNA target we use for modeling, dT_4A_4 (pdb id: 1RVI) resembles the preferred target substrate of the L1-EN, having a unique geometry that includes a narrow minor groove that widens at the TpA step, making it a good candidate for the docking study. L1-EN residues defined as "active" and "passive" in the HADDOCK formalism were selected based on mutational data and are shown in Table III. The docking partner, dT_4A_4, was annotated with adenine and thymine at the TpA step as active and the rest of the molecule as passive. Complexes were scored as described in "Experimental procedures" and "Results" are summarized in Supp. Table III; our favored model obtained with HADDOCK is illustrated in Figure 5A (representative models of all model clusters are in Suppl. Fig. 3). The substrate DNA is docked on the predicted interface and the βB6-βB5 hairpin can be seen penetrating the wide minor groove. Major groove contacts are mediated by N118 and neighboring residues on

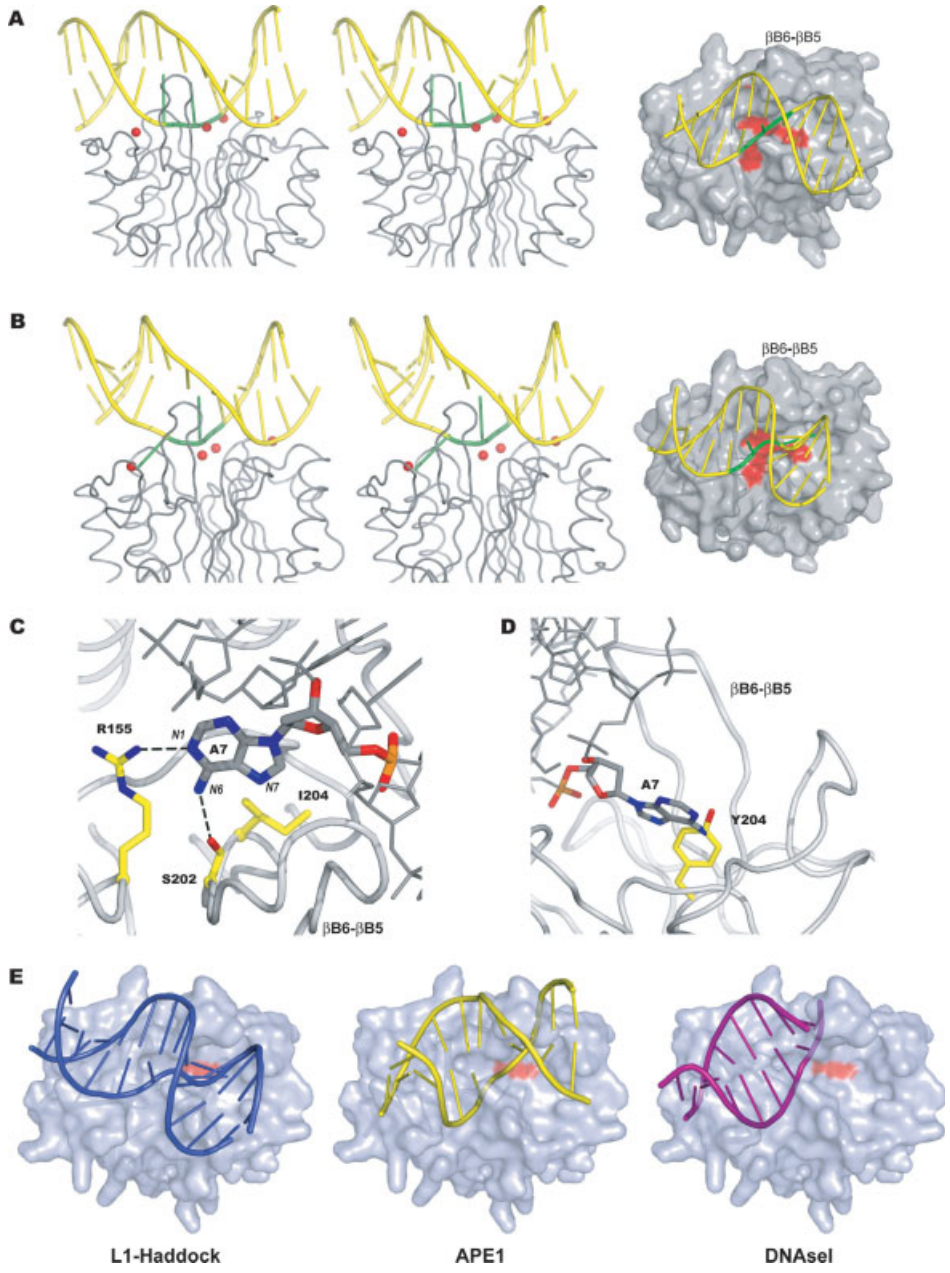

Figure 5

A: Stereo view of "nonflipped" DNA (wild-type) docked onto L1-EN with HADDOCK. The model depicted is the representative of the best cluster. The βB6-βB5 hairpin is inserting in the DNA minor groove. Red spheres represent sulfate ions from the LTx crystal structure. The DNA cartoon is shown in yellow, with the bases at the TpA step targeted for cleavage in green, Gray: L1-EN ribbon/Alternative representation: L1-EN is drawn as a gray surface and is viewed from the top, so that the βB6-βB5 hairpin is pointing towards the reader. The double-stranded A-tract DNA substrate is drawn as a yellow cartoon, with the TpA step that is the supposed target for cleavage highlighted in green. The red colored surface indicates the position of active site residues of L1-EN involved in catalysis. **B**: Same as in A but with "flipped-out" DNA as substrate. **C**: Close-up of the flipped-out HADDOCK model to illustrate the accommodation of adenine 7 by R155, S202 and I204. L1-EN is depicted as light-gray ribbon with the hairpin labeled and the DNA as dark-gray sticks. Oxygen: red, Nitrogen: blue, Phosphorus: orange. **D**: Alternative view of same model, illustrating the clash between I204Y mutant (superposed in yellow) and adenine 7 of the TpA step, which are shown as ball and stick models. Coloring as in C. **E**: Three surface representations of L1-EN with superposed DNA cartoons from flipped-out HADDOCK (blue), APE1 (yellow, pdb code: 1DE8) and DNAse I (magenta, pdb code: 1DNK). The position of the I204Y mutation is marked as red surface.

the βB3-αB1 loop. Residues N118 and F193 form a narrow passage of less than 6.5 Å, where the DNA can be locked in position while catalysis of the phosphodiester bond takes place (Fig. 3, blue asterisks). Four of the sulfate ions we previously described coincide with backbone phosphates of the DNA. The docked substrate is placed

on an orientation suitable for cleavage, with the target phosphodiester bond between Thy6 and Ade7 in the proximity of the active site.

The second DNA model used in this study contains an extra-helical adenine at the T-A junction. To construct it we chose to use angles from two known structures that contain flipped out nucleotides: that of APE1 bound to abasic DNA (PDB code: 1DE9): that of β-glucosyltransferase bound to a 13-mer DNA (PDB code: 1SXP). The model of L1-EN complex with this *in silico* "manipulated" DNA was created following exactly the same protocol and using identical restraints as earlier, to enable unbiased comparison of the two models. A number of complexes were selected for further inspection (Suppl. Table II and Suppl.Fig. 3) and the best resulting complex is shown in Figure 5(B). The overall binding of this DNA appears similar [illustrated in Fig. 5 (A,B)] but this time the extra-helical adenine at position 7 is within hydrogen bonding distance from R155, S202, and close to I204 that create a pocket for accommodating the flipped-out nucleotide [Fig. 5(C)]. Arginine 155 makes a single hydrogen bond to the N1 of the adenine base in this model; however one can easily imagine that a slight rearrangement can result to R155 contacting both N1 and N6, whereas the hydroxyl of S202 (now hydrogen bonding N6) contacts N7 instead, as we actually modeled it manually in the past.[16] This rearrangement would imply a more important role of R155 than S202 in catalysis, consistent with the activity experiments [Fig. 1(B)]. Similar to the nonflipped DNA docking, the bond to be cleaved at the TpA step is positioned near the active site, where direct or water mediated contacts could take place. In the case of a nucleotide flipping-out, its Watson-Crick partner would need to be stabilized during catalysis in a way similar to that of APE1, which provides two loops contacting the orphan base from both minor and major DNA grooves. These contacts in L1-EN could be achieved by F193 of the main hairpin βB6-βB5 in the minor groove and the βB3-αB1 loop containing the important N118 in the major groove, stabilizing the orphan thymidine in the HADDOCK model.

When the I204Y variant structure is superposed on the "flipped-out" complex, there is a very obvious clash between the extra-helical adenine and Y204 that would be present in DNaseI [Fig. 5(D)]. That would argue in favor of the base-flipping mechanism, illustrating that the bulky tyrosine instead of an isoleucine is obstructing the recognition of the extra-helical nucleotide and thus rendering the endonuclease inactive. On the contrary, superposition of the I204Y variant structure on the "nonflipped" complex does not seem to obstruct DNA binding in a major way. With the currently available models the behavior of the I204Y mutant is favoring the recognition of an extra-helical nucleotide during cleavage. However, given the available data, we cannot rule out the possibility that the tyrosine can rotate and point to different directions which would render any EN-L1:DNA complex unstable or inactive, even in the case that base-flipping would not be required.

When we superimpose the structures of the previously reported LTx and LR1 variants[32] on these new docking models, there is a striking difference on the way the two exchanged hairpins could interact with the DNA substrate. The morphology and positioning of the loops is quite different: the Tx1L hairpin in the LTx structure, penetrates and pushes the minor groove; the shorter R1Bm hairpin in the LR1 structure barely reaches the DNA backbone. These structural differences could explain the reduced activity and altered specificity these variants exhibit as we initially suggested based on the apo- structures and biochemical data.[32]

CONCLUSIONS

To understand the mechanism of cleavage of L1-EN, we have applied four methodologies on a series of mutants: biochemical assaying, crystal structure determination and analysis molecular dynamics simulations, and computational modeling studies. We were able to confirm that the protein consists of a rigid platform, and we have identified surface loops, in addition to the βB6-βB5 β-hairpin, that show increased mobility in the presumed DNA binding surface. Our simulations indicated that these loops move in an anticorrelated manner, suggesting the possibility of a DNA "breathing" motion that might prove important for target DNA recognition. It was however not possible to conclude if DNA cleavage proceeds with a flipped-out base mechanism or not. The two docking models we present yield complexes that resemble to a greater extent the way DNA is bound on APE1 structures, rather than how DNaseI binds its target [Fig. 5(E)]: they reveal a large binding interface with the DNA sitting on both sides of the hairpin loop. The positioning of the I204 residue in the flipped-out APE1-like DNA model, together with the reduction in EN activity when I204 is mutated to a tyrosine, favors the flipped-out DNA base possibility. Additional experiments and in particular the experimental structure of L1-EN bound to its target DNA are the only means that would allow to resolve this issue in a conclusive and deterministic manner.

METHODS

Expression and purification and crystallization of L1-EN variants

L1-EN variants were cloned, expressed, and purified as described previously.[16,32]

Crystallization

The D145A mutant was concentrated to 15 mg/mL and crystallized from 22% PEG4000, 0.2*M* Ammonium

Sulphate (AS), and 20 m*M* $MgCl_2$. The I204Y mutant was concentrated to 8 mg/mL and crystallized from 25% PEG4000, 0.2*M* Ammonium Sulphate (AS), and 20 m*M* $MgCl_2$. The R155A mutant was concentrated to 11 mg/mL and crystallized from 34% PEG4000, 0.2*M* Ammonium Sulphate (AS), and 20 m*M* $MgCl_2$.

Oligonucleotide nicking assay

Gel-purified synthetic oligonucleotides of a 20 base pair long A-T rich oligonucleotide that contains the preferred consensus L1-EN cleavage site 5′ TTTT/AA 3′ were labeled at the 5′ end with radioactive phosphate (^{32}P) using [g-^{32}P]ATP and T4 polynucleotide kinase and were repurified on gel. Equimolar amounts (450 n*M*) of unlabeled complementary and substrate strands were mixed with a trace amount of labeled substrate. The mixture was annealed in 5 m*M* Na-HEPES (pH = 7.5) by heating to 90°C and slow-cooled to room temperature. Nicking reactions (50 μL) were done in 50 m*M* Na-HEPES (pH = 6.5), 150 m*M* NaCl, 10 m*M* $MgCl_2$, 0.1 mg/mL bovine serum albumin, and 1 m*M* dithiothreitol. Final concentrations were 180 n*M* DNA (0.5–7.5 μg/mL) and 20–2000 n*M* protein, which had been diluted in protein buffer (5 m*M* Na-HEPES (pH = 7.5), 300 m*M* NaCl, 10 m*M* $MgCl_2$, 0.5 mg/mL bovine serum albumin, and 5 m*M* dithiothreitol) before. After 30 min at 37°C reactions were stopped by the addition of 175 μL of 380 m*M* Na-acetate (pH = 7.5), phenol extraction, and ethanol precipitation. Reaction products were separated on 10% denaturing polyacrylamide gels.

Crystallographic structure solution, refinement

Diffraction data were collected at beamline ID23-1 at the European Synchrotron Radiation Facility in Grenoble, France. Diffraction images were integrated by MOSFLM[37] and scaled in SCALA.[38] The structures were solved by molecular replacement using MOLREP[39] with L1-EN (PDB ID: 1VYB) as a search model. Automatic model building was done with ARP/wARP[40] and initial refinement was done using REFMAC5[41] and COOT[42] iteratively. After correcting for twining using X-triage[43] (see next paragraph) maps were inspected and models were finalized and validated using the Molprobity server.[44]

Detection and treatment of crystallographic twinning

Although electron density maps were of acceptable quality, refinement statistics failed to converge, with R_{free} ranging between 32.5% for the I204Y mutant, 30.9% for the D145A mutant, and 26.3% for the R155A mutant. Examining the diffraction data with the program X-triage,[43] we detected significant indications for twining, possible in this apparently monoclinic but metrically close to orthorhombic space group. We used the twin operator (−h, −k, l) to refine the twin fraction in the refinement program of the PHENIX package. The twin fraction refined to values ranging between 0.45, 0.29, 0.1 (for I204Y, D145A, R155A respectively) and in all cases resulted to considerably improved R_{free} values (20.8%, 25.6%, 24.7% for the I204Y, D145A, and R155A mutants respectively) and improved electron density maps. All detwinned maps were inspected and models were finalized and validated using the Molprobity server.[44]

Molecular dynamics simulations

A total of five molecular dynamics (MD) simulations in explicit solvent were performed using the GROMACS 3.3 package[45,46] and the GROMOS96 force field.[47] L1-EN (PDB code: 1VYB), LTx, LR1, R155A, and TRAS1-EN (PDB code: 1WDU) crystal structures were used as starting models for the MD simulations. Structures were solvated in a cubic box of SPC water[48] using a minimum distance of 14 Å between the protein and the box edges. After a steepest decent energy minimization step with positional restraints on the solute, negatively (Cl^-) or positively (Na^+) charged ions were introduced depending on the run, to obtain an electro-neutralized system. A second energy minimization was performed, followed by five successive 20 ps MD equilibration runs. During these, the position restraints force constant on the solute Kposre was decreased progressively (1000, 1000, 100, 10, 0 kJ/mol nm^2). After these equilibration stages, a 10 ns production run was performed.

Solute, solvent, and counterions were weakly coupled independently to reference temperature baths at 300 K (τ = 0.1 ps).[49] The pressure was maintained by weakly coupling the system to an external pressure bath at one atmosphere (1atm = 101,325 Pa). The LINCS algorithm was used to constrain bond lengths, allowing an integration time step of 2 fs to be used.[50] The nonbonded interactions were calculated with a twin-range cut-off of 0.8 and 1.4 nm.[51] The long-range electrostatic interactions beyond the 1.4 nm cut-off were treated with the generalized reaction field model, using a dielectric constant of 54.[52] Trajectory coordinates and energies were stored at 0.5 ps intervals. The analysis was performed routinely for the last 6 ns of each simulation, to allow for the system to start equilibrating, using the set of programs within GROMACS. Statistics for energy terms and RMSD are available in Suppl. Table III and Suppl. Figures 4 and 5.

To identify smaller scale motions that were possibly overshadowed by more dominant ones and examine possible correlations, we analyzed the projections of each trajectory on the eigenvectors of its covariance matrix. The fast, free and flexible GROMACS package is available at http://www.gromacs.org/.

HADDOCK protein DNA docking protocol

As starting structures for the protein partner we used our L1-EN crystal structure (PDB code: 1VYB), together

with two available additional structures of alternative space-group, in order to account for some degree of flexibility within the protein. For the DNA partner, we used an NMR ensemble consisting of nine structures (PDB code: 1RVI). Ambiguous Interaction Restraints (AIRs) are shown in Table III. As active residues, we choose those residues involved in the interaction interface using experimental information extracted from sequence, structural, and mutational data; for all of them a relative solvent accessibility of >50% was required, defined with NACCESS.[53] Neighbors of the active residues that could also be part of the interface were defined as passive residues using the same solvent accessibility criteria.

The docking protocol consists of three steps: (1) rigid-body docking, (2) semi-flexible simulated annealing, and (3) refinement in explicit solvent. During the initial rigid-body docking stage all combinations of different starting structures are used to create 1000 protein:DNA complexes. The best 200 of these in terms of intermolecular energy according to the HADDOCK score were selected for the simulated annealing step and the final water refinement as described in.[24] Default HADDOCK (version 2.0) parameters were used and the top scoring complexes were further visually inspected for contacts between DNA bases and residues in the protein involved in both catalysis and DNA stabilization. The HADDOCK package is freely available to academic users (http://www.nmr.chem.uu.nl/haddock).

Computer hardware

The GROMACS computations were performed on a cluster of 5 Apple Xserve dual G5 processors at 2.3 GHz having 1GB of RAM memory each and using Lam-MPI over gigabit ethernet as a inter-node communication implementation.

HADDOCK docking runs were performed on a Transtec (Tubingen, Germany) computer cluster operating with 32, 2.0 GHz, 64 bit Opteron processors.

ACKNOWLEDGMENTS

We thank O. Weichenrieder who originated the project at the NKI, and Titia K. Sixma, for helpful and critical discussions; beam line scientists at the ESRF for assistance with data collection; Peter H. Zwart for the help in twin detection and refinement. Coordinates and structure factors for the mutants have been deposited with the PDB.

REFERENCES

1. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, et al. Initial sequencing and analysis of the human genome. Nature 2001;409: 860–921.
2. Han JS, Boeke JD. LINE-1 retrotransposons: modulators of quantity and quality of mammalian gene expression? Bioessays 2005;27:775–784.
3. Babushok DV, Kazazian HH. Progress in understanding the biology of the human mutagen LINE-1. Hum Mutat 2007;28:527–539.
4. Szak ST, Pickeral OK, Makalowski W, Boguski MS, Landsman D, Boeke JD. Molecular archeology of L1 insertions in the human genome. Genome Biol 2002;3:research0052.
5. Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ, Parmigiani G, Boeke JD. Human l1 retrotransposition is associated with genetic instability in vivo. Cell 2002;110:327–338.
6. Gilbert N, Lutz-Prigge S, Moran JV. Genomic deletions created upon LINE-1 retrotransposition. Cell 2002;110:315–325.
7. Feng Q, Moran JV, Kazazian HH, Boeke JD. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. Cell 1996;87:905–916.
8. Cost GJ, Boeke JD. Targeting of human retrotransposon integration is directed by the specificity of the L1 endonuclease for regions of unusual DNA structure. Biochemistry 1998;37:18081–18093.
9. Luan DD, Korman MH, Jakubczak JL, Eickbush TH. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. Cell 1993;72:595–605.
10. Xiong Y, Eickbush TH. The site-specific ribosomal DNA insertion element R1Bm belongs to a class of non-long-terminal-repeat retrotransposons. Mol Cell Biol 1988;8:114–123.
11. Maita N, Aoyagi H, Osanai M, Shirakawa M, Fujiwara H. Characterization of the sequence specificity of the R1Bm endonuclease domain by structural and biochemical studies. Nucleic Acids Res 2007;35:3918–3927.
12. Garrett JE, Knutzon DS, Carroll D. Composite transposable elements in the Xenopus laevis genome. Mol Cell Biol 1989;9:3018–3027.
13. Feng Q, Schumann G, Boeke JD. Retrotransposon R1Bm endonuclease cleaves the target sequence. Proc Natl Acad Sci USA 1998;95: 2083–2088.
14. Christensen S, Pont-Kingdon G, Carroll D. Target specificity of the endonuclease from the Xenopus laevis non-long terminal repeat retrotransposon. Tx1L. Mol Cell Biol 2000;20:1219–1226.
15. Zingler N, Weichenrieder O, Schumann GG. APE-type non-LTR retrotransposons: determinants involved in target site recognition. Cytogenet Genome Res 2005;110:250–268.
16. Weichenrieder O, Repanas K, Perrakis A. Crystal structure of the targeting endonuclease of the human LINE-1 retrotransposon. Structure 2004;12:975–986.
17. Dlakić M. Functionally unrelated signalling proteins contain a fold similar to Mg^{2+}-dependent endonucleases. Trends Biochem Sci 2000;25:272–273.
18. Dodson G, Verma CS. Protein flexibility: its role in structure and mechanism revealed by molecular simulations. Cell Mol Life Sci 2006;63:207–219.
19. van Gunsteren WF, Berendsen HJ, Hermans J, Hol WG, Postma JP. Computer simulation of the dynamics of hydrated protein crystals and its comparison with x-ray data. Proc Natl Acad Sci USA 1983; 80:4315–4319.
20. McCammon JA, Gelin BR, Karplus M. Dynamics of folded proteins. Nature 1977;267:585–590.
21. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. Nat Struct Biol 2002;9:646–652.
22. Karplus M, Kuriyan J. Molecular dynamics and protein function. Proc Natl Acad Sci USA 2005;102:6679–6685.
23. Rueda M, Ferrer-Costa C, Meyer T, Pérez A, Camps J, Hospital A, Gelpí JL, Orozco M. A consensus view of protein dynamics. Proc Natl Acad Sci USA 2007;104:796–801.
24. Dominguez C, Boelens R, Bonvin AM. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J Am Chem Soc 2003;125:1731–1737.
25. van Dijk AD, Fushman D, Bonvin AM. Various strategies of using residual dipolar couplings in NMR-driven protein docking: application to Lys48-linked di-ubiquitin and validation against (15)N-relaxation data. Proteins 2005;60:367–381.
26. van Dijk M, van Dijk AD, Hsu V, Boelens R, Bonvin AM. Information-driven protein-DNA docking using HADDOCK: it is a matter of flexibility. Nucleic Acids Res 2006;34:3317–3325.

27. Kopke Salinas R, Folkers GE, Bonvin AM, Das D, Boelens R, Kaptein R. Altered specificity in DNA binding by the lac repressor: A mutant lac headpiece that mimics the gal repressor. Chembiochem 2005;6:1628–1637.
28. Kamphuis MB, Bonvin AM, Monti MC, Lemonnier M, Muñoz-Gómez A, van den Heuvel RH, Díaz-Orejas R, Boelens R. Model for RNA binding and the catalytic site of the RNase kid of the bacterial parD toxin-antitoxin system. J Mol Biol 2005;357:115–126.
29. Kalodimos CG, Biris N, Bonvin AM, Levandoski MM, Guennuegues M, Boelens R, Kaptein R. Structure and flexibility adaptation in nonspecific and specific protein-DNA complexes. Science 2004;305:386–389.
30. Mol CD, Izumi T, Mitra S, Tainer JA. DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination [corrected]. Nature 2000;403:451–456.
31. Mol CD, Kuo CF, Thayer MM, Cunningham RP, Tainer JA. Structure and function of the multifunctional DNA-repair enzyme exonuclease III. Nature 1995;374:381–386.
32. Repanas K, Zingler N, Layer LE, Schumann GG, Perrakis A, Weichenrieder O. Determinants for DNA target structure selectivity of the human LINE-1 retrotransposon endonuclease. Nucleic Acids Res 2007;35:4914–4926.
33. Maita N, Anzai T, Aoyagi H, Mizuno H, Fujiwara H. Crystal structure of the endonuclease domain encoded by the telomere-specific long interspersed nuclear element. TRAS1. J Biol Chem 2004;279:41067–41076.
34. Theobald DL, Wuttke DS. THESEUS: maximum likelihood superpositioning and analysis of macromolecular structures. Bioinformatics 2006;22:2171–2172.
35. Weston SA, Lahm A, Suck D. X-ray structure of the DNase I-d(GGTATACC)2 complex at 2.3 A resolution. J Mol Biol 1992;226:1237–1256.
36. Stefl R, Wu H, Ravindranathan S, Sklenár V, Feigon J. DNA A-tract bending in three dimensions: solving the dA4T4 vs. dT4A4 conundrum. Proc Natl Acad Sci USA 2004;101:1177–1182.
37. Leslie AG. The integration of macromolecular diffraction data. Acta Crystallogr D Biol Crystallogr 2006;62(Pt 1):48–57.
38. Evans P. Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr 2006;62(Pt 1):72–82.
39. Vagin A, Teplyakov A. An approach to multi-copy search in molecular replacement. Acta Crystallogr D Biol Crystallogr 2000;56(Pt 12):1622–1624.
40. Cohen SX, Morris RJ, Fernandez FJ, Ben Jelloul M, Kakaris M, Parthasarathy V, Lamzin VS, Kleywegt GJ, Perrakis A. Towards complete validated models in the next generation of ARP/wARP. Acta Crystallogr D Biol Crystallogr 2004;60(Pt 12 Pt 1):2222–2229.
41. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 1997;53(Pt 3):240–255.
42. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 2004;60(Pt 12 Pt 1):2126–2132.
43. Zwart PH, Grosse-Kunstleve RW, Lebedev AA, Murshudov GN, Adams PD. Surprises and pitfalls arising from (pseudo)symmetry. Acta Crystallogr D Biol Crystallogr 2008;64(Pt 1):99–107.
44. Lovell SC, Davis IW, Arendall WB, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by Calpha geometry: phi,psi and Cbeta deviation. Proteins 2003;50:437–450.
45. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. J Comput Chem 2005;26:1701–1718.
46. Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analisys. J Mol Model 2001;7:306–317.
47. Daura X, Mark AE, van Gunsteren WF. Parametrization of aliphatic CHn united atoms of GROMOS96 force field. J Comput Chem 1998;19:535–547.
48. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J. Interaction models for water in relation to protein hydration. In: Pullman B, editor. Intermolecular forces. Dordrecht: Reidel Publishing Company; 1981. pp 331–342.
49. Berendsen HJC, Postma JPM, van Gunsteren WF, Di Nola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
50. Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: a Linear constraint solver for molecular simulations. J Comput Chem 1997;18:1463–1472.
51. van Gunsteren WF, Berendsen HJC. Computer-simulation of molecular-dynamics—methodology. Applications, and perspectives in chemistry. Angew Chem Int Ed English 1990;29:992–1023.
52. Tironi IG, Sperb R, Smith PE, Vangunsteren WF. A generalized reaction field method for molecular-dynamics simulations. J Chem Phys 1995;102:5451–5459.
53. Hubbard SJ, Thornton JM. NACCESS. London: Department of Biochemistry and Molecular Biology, University College; 1993.