

# SCIENTIFIC REPORTS



OPEN

## Crystal structure and MD simulation of mouse EndoV reveal wedge motif plasticity in this inosine-specific endonuclease

Meh Sameen Nawaz<sup>1</sup>, Erik Sebastian Vik<sup>1</sup>, Mia Elise Ronander<sup>2</sup>, Anne Marthe Solvoll<sup>2</sup>, Pernille Blicher<sup>2</sup>, Magnar Bjørås<sup>1</sup>, Ingrun Alseth<sup>1</sup> & Bjørn Dalhus<sup>1,2</sup>

Endonuclease V (EndoV) is an enzyme with specificity for deaminated adenosine (inosine) in nucleic acids. EndoV from *Escherichia coli* (EcEndoV) acts both on inosines in DNA and RNA, whereas the human homolog cleaves only at inosines in RNA. Inosines in DNA are mutagenic and the role of EndoV in DNA repair is well established. In contrast, the biological function of EndoV in RNA processing is largely unexplored. Here we have characterized a second mammalian EndoV homolog, mouse EndoV (mEndoV), and show that mEndoV shares the same RNA selectivity as human EndoV (hEndoV). Mouse EndoV cleaves the same inosine-containing substrates as hEndoV, but with reduced efficiencies. The crystal structure of mEndoV reveals a conformation different from the hEndoV and prokaryotic EndoV structures, particularly for the conserved tyrosine in the wedge motif, suggesting that this strand separating element has some flexibility. Molecular dynamics simulations of mouse and human EndoV reveal alternative conformations for the invariant tyrosine. The configuration of the active site, on the other hand, is very similar between the prokaryotic and mammalian versions of EndoV.

The exocyclic amine group of adenosine (A) can be hydrolyzed, resulting in formation of inosine (I). This process occurs spontaneously in cells and is induced by certain types of stress, for example from nitrosative agents formed as a response to inflammation or infection, or from exposure from the environment<sup>1</sup>. Inosine is read as guanosine (G) and consequently cytosine (C) is incorporated opposite inosine by the DNA polymerases during replication<sup>2</sup>. The A-to-I conversion is thus mutagenic. Removal of inosines from genomic DNA is dealt with by the base excision repair pathway<sup>3–5</sup> as well as endonuclease V (EndoV), which is considered the major player<sup>6,7</sup>. In *Escherichia coli* and some other bacteria, EndoV initiates repair by cleavage of the second phosphodiester bond 3' to inosine in an Mg<sup>2+</sup> dependent reaction<sup>8–11</sup>. In contrast to DNA glycosylases, EndoV does not remove the deaminated nucleotide. After cleavage, EndoV stays tightly bound to the incised product<sup>12,13</sup> and additional, yet unknown, proteins are recruited to complete the repair<sup>14</sup>. Weak activity for inosines in DNA has been reported for mouse (m)EndoV<sup>15</sup> and human (h)EndoV<sup>16</sup>, however some controversy exists<sup>17</sup>.

Whereas inosine in DNA is considered as damage, inosine in RNA is introduced by specific enzymes in a highly regulated manner to increase transcriptomic diversity. The adenosine deaminases acting on RNA (ADARs) enzymes catalyze this A-to-I editing on mRNA and non-coding (nc)RNA including long ncRNA, micro (mi)RNA and small interfering (si)RNA. A-to-I editing is abundant in higher eukaryotes and edited sites amount to more than a hundred million and are spread over the majority of human genes<sup>18</sup>. Defect editing is linked to various human diseases including neurological disorders, infections and cancer<sup>19</sup>. Also some tRNAs undergo A-to-I editing and here the reaction is catalyzed by enzymes homologous to ADARs, namely the adenosine deaminases acting on tRNA (ADATs). This editing is essential for protein synthesis<sup>20</sup>. Unexpectedly, recently it was shown that both human and *E. coli* (Ec)EndoV could incise RNA substrates containing inosines. RNA cleavage was catalyzed with comparable efficiencies for the two enzymes, comparable to that of DNA for EcEndoV, suggesting that RNA

<sup>1</sup>Department of Microbiology, Oslo University Hospital HF, Rikshospitalet and University of Oslo, NO-0424 Oslo, Norway. <sup>2</sup>Department of Medical Biochemistry, Institute for Clinical Medicine, University of Oslo, NO-0424 Oslo, Norway. Correspondence and requests for materials should be addressed to B.D. (email: bjorn.dalhus@medisin.uio.no)

5' → 3' sequence	DNA/RNA	Description
CCGUAGAGCUAC[r]GAUCGGUCACCG	RNA	I-RNA
CGGUGACCGAUCUGUAGCUCUACGG	RNA	Complementary oligo for I-RNA
ACUGGACAAAUACUCCGAGG	RNA	ctr-RNA
CCUCGGAGUAUUUGUCCAGU	RNA	Complementary oligo for ctr-RNA
ACUGGACA[r][r]U[r]CUCCGAGG	RNA	IIUI-RNA
CCUCGGAGU[r]UUUGUCCAGU	RNA	Complementary oligo for IIUI-RNA
CCGTAGAGCTAC[dI][rG]ATCGGTCACCG	DNA	dI rG-DNA
CCGTAGAGCTAC[dI]GATCGGTCACCG	DNA	I-DNA
CGGTGACCGATCTGTAGCTCTACGG	DNA	Complementary oligo for dI rG/I-DNA

**Table 1. Oligonucleotide substrates used in the study.**

is the preferred substrate *in vivo* at least for hEndoV<sup>17,21</sup>. The biological significance of RNA incision at inosines by EndoV is yet not known<sup>22,23</sup>.

Some links between EndoV and RNA metabolism was already known from analysis of known 3D structures. For example, the crystal structure of EndoV from *Termotoga maritima* (TmEndoV) in complex with inosine-containing DNA has been solved<sup>12</sup>. This structure reveals that TmEndoV contains an “RNase H-like motif” resembling that in *E. coli* RNase H and the PIWI domain of *Pyrococcus furiosus* Argonaute, both being well characterised RNases<sup>12</sup>. As no robust DNA repair activity has been found for the mouse EndoV homolog, we speculated whether mEndoV also is an RNase with preference for inosines in RNA. Indeed mEndoV cleaved inosine-containing RNA, albeit less efficient than hEndoV. The crystal structure of mEndoV was solved and when compared to the structure of hEndoV, we observe a conformation of the strand-separating wedge, and a concurrent wide DNA/RNA binding cleft, different from all previous EndoV structures (prokaryotic and human). The new structure suggests that this conserved element is flexible and may switch between unproductive and productive conformations for RNA cleavage.

## Results

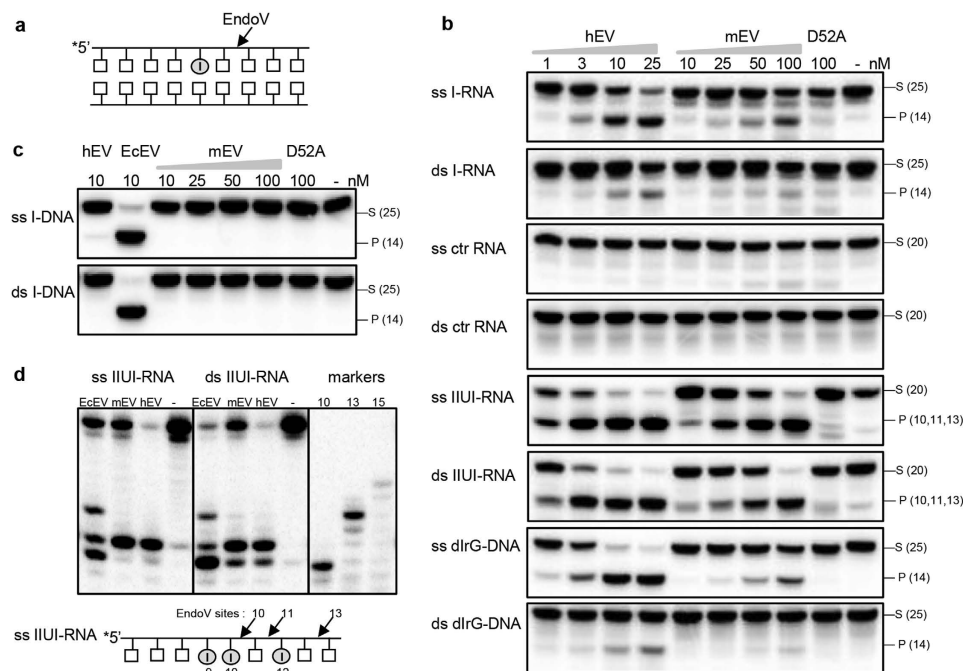
**Biochemical properties of mEndoV.** The mouse genome encodes a single EndoV homolog, mEndoV. The open reading frame of 1017 nucleotides translates into a protein of 338 amino acids with a calculated mass of 37.2 kDa. Mouse EndoV shares high sequence similarity with human EndoV except at the C-terminus which is also predicted to be structurally disordered and without evolutionary conservation<sup>24</sup>. Recent studies have revealed that hEndoV has a preference for inosines in RNA rather than DNA<sup>17,21</sup> and we tested whether this was true for mEndoV as well. Indeed, when incubated with a single-stranded RNA substrate with a central inosine residue (I-RNA, Table 1), mEndoV was active, albeit cleavage was less efficient than for hEndoV (Fig. 1a,b). The double-stranded I-RNA substrate was a weaker substrate for mEndoV. Neither mEndoV nor hEndoV incised RNA without inosine, excluding these enzymes as general ribonucleases (Fig. 1b).

In cells multiple inosines are often found clustered, a phenomena referred to as RNA hyperediting<sup>18</sup>. To test such a substrate, a synthetic RNA with multiple inosines (IIUI-RNA) corresponding to part of exon 2 of rat  $\alpha$ -tropomyosin gene was synthesized<sup>25</sup>. Mouse EndoV incised this substrate more efficiently both in a single- and double-stranded context (Fig. 1b). A site specific mEndoV mutant of the catalytic aspartate (D52A) totally lost the inosine-RNA incision activity for this as well as all other substrates tested (Fig. 1b).

Human EndoV has no activity on inosine-containing DNA substrates<sup>17</sup> which applied to mEndoV as well (Fig. 1c). However, hEndoV efficiently cleaves a DNA substrate with inosine when a ribonucleotide is present immediately 3' to the inosine residue<sup>17</sup>. This substrate (dI rG) was also incised by mEndoV confirming that both enzymes are critically dependent on a ribonucleotide for incision (Fig. 1b). Again the double-stranded substrate was less efficiently cleaved than the single-stranded version.

To map the exact cleavage position of the IIUI-RNA substrates, the reaction products were run on sequencing gels. The major incision products for both mEndoV and hEndoV corresponded to cleavage of the second phosphodiester bond 3' of the second inosine (from the 5' end) of the substrates (Fig. 1d). For the double-stranded substrate there was also some cleavage corresponding to the first inosine (from the 5' end). For EcEndoV three cleavage products were seen, suggesting that all three inosines may accommodate in the nucleotide binding pocket. As the substrates were 5'-labelled, these assays will not reveal potential cleavage at the 3' inosine prior to cleavage at the other inosines.

The catalytic properties of mouse and human EndoV were analyzed by testing different reaction conditions. Assays were performed at two pHs, 7.5 and 8.5, with different concentrations of Mn<sup>2+</sup> and Mg<sup>2+</sup> and with single- and double-stranded IIUI- and I-RNA substrates (Supplementary Fig. S1). The type and concentration of metal ion and pH did not change the activities of the two EndoVs notably, except that at pH 8.5, a Mn<sup>2+</sup> concentration of 2.5 mM and higher inhibited activity. For further analysis two conditions were chosen; pH 8.5 with 2.5 mM Mg<sup>2+</sup> (best activity) and pH 7.5 with 0.5 mM Mn<sup>2+</sup> (close to physiological pH). The two buffers were used to test the effect of various Na<sup>+</sup> and K<sup>+</sup> concentrations on EndoV activity with the IIUI-RNA substrates. NaCl concentrations of 50, 100 or 150 mM did not change EndoV activity (Supplementary Fig. S2), except for double-stranded IIUI at pH 7.5 and 0.5 mM Mn<sup>2+</sup> where inhibition was seen with increasing amounts of NaCl.



**Figure 1. Processing of inosine by EndoV enzymes.** (a) Schematic illustration of EndoV cleavage of the second phosphodiester bond 3' to an inosine residue. (b) Increasing amounts of human EndoV (hEV; 1–25 nM) and mouse EndoV (mEV; 10–100 nM) were incubated with 1 nM RNA or DNA substrates: ss/ds I-RNA, ss/ds ctr-RNA, ss/ds IIUI-RNA, ss/ds dIrG-DNA and (c) ss/ds I-DNA at 37 °C for 30 minutes using standard reaction buffer. 100 nM of the site specific mEndoV mutant D52A was included in all assays. Cleavage products were analyzed by 20% denaturing PAGE and visualized by phosphorimaging. (d) EndoV cleavage products for the IIUI-RNA substrates were run on 20% sequencing gels alongside with 10, 13 and 15 mer RNA markers. The illustration shows potential EndoV cleavage sites on the IIUI-RNA substrate. Abbreviations: - = no enzyme added, S = substrate, P = cleaved product, ss = single-stranded, ds = double-stranded, EcEV = *E. coli* EndoV. Enzyme amounts are shown in nM and sizes of RNA/DNA substrates and products are indicated in parentheses.

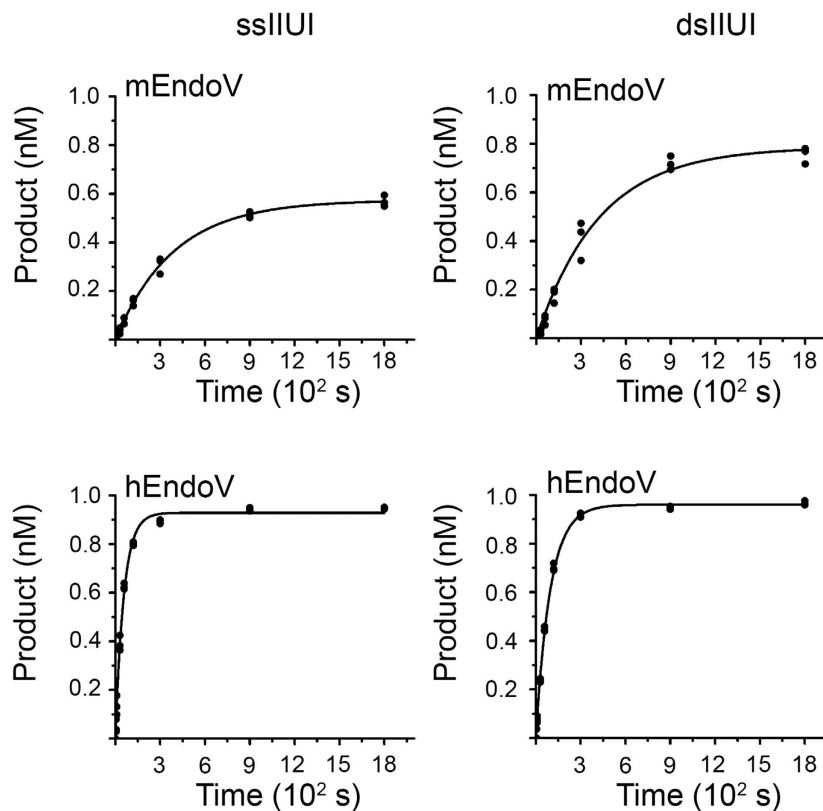
Enzyme	Substrate	$k_{obs}$ ( $s^{-1}$ )
Human EndoV	ss IIUI-RNA	$0.0177 \pm 0.0006$
	ds IIUI-RNA	$0.0104 \pm 0.0003$
Mouse EndoV	ss IIUI-RNA	$0.0026 \pm 0.0001$
	ds IIUI-RNA	$0.0024 \pm 0.0002$

**Table 2. Single-turnover kinetics rates ( $k_{obs}$ ) with standard deviations for mouse and human EndoV on ss IIUI-RNA and ds IIUI-RNA substrates.** Standard-deviations are calculated from curve fitting to three independent data sets.

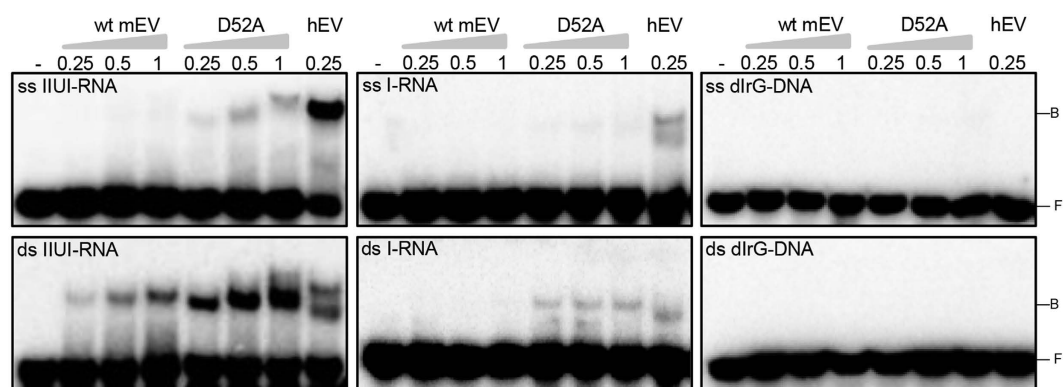
Higher KCl concentrations than 50 mM (100 and 150 mM) inhibited both mouse and human EndoV activity (Supplementary Fig. S1). In sum, the two enzymes respond very similarly to changes in reaction conditions, and under all conditions tested, mouse EndoV is less active than human EndoV.

To compare the catalytic efficiencies of mEndoV and hEndoV, single-turnover kinetic analyzes were performed on the two IIUI-RNA substrates. The results revealed a 4–7 fold higher turnover rate for these substrates for hEndoV over mEndoV (Table 2, Fig. 2). Human EndoV had about a 2 times higher turnover rate for the single- versus double-stranded substrate ( $k_{obs}$ ,  $0.0177 s^{-1}$  and  $0.0104 s^{-1}$ , respectively), a trend also previously shown for a single-inosine RNA substrate<sup>17</sup>, while mEndoV had the same turnover rate for cleavage of single- and double-stranded IIUI-RNA substrates ( $k_{obs}$ ,  $0.0026 s^{-1}$  and  $0.0024 s^{-1}$ , respectively; Table 2).

Human EndoV has been shown to interact and bind to different DNA and RNA substrates<sup>21,24</sup>. To test if mEndoV makes stable complexes with its substrates, electrophoretic mobility shift assays (EMSA) were done. Mouse EndoV barely made shifts with the single-stranded IIUI-RNA substrate, but bound stronger to the double-stranded IIUI-RNA substrate (Fig. 3). Under the conditions used, mEndoV bound much weaker than hEndoV, and no shift was observed for the I-RNA and dIrG-DNA substrates (Fig. 3). The catalytic inactive mEndoV mutant (D52A) made stronger shifts than the wild-type enzyme especially for the double stranded IIUI-RNA substrate, comparable in intensity to that of hEndoV. To verify that the bandshifts reflects substrate affinity and not product binding, the EMSA reaction products were analyzed by denaturing PAGE. Under the assay conditions used, no cleavage of the substrates was observed for none of the enzymes (data not shown).



**Figure 2.** Single-turnover kinetics of mouse and human EndoV on ss and ds IIUI-RNA substrates. 30 nM enzymes were incubated with 1 nM ss IIUI-RNA or ds IIUI-RNA substrates and samples were withdrawn at time points as indicated and analyzed by PAGE. A one phase association model (solid line) was fitted to three parallel, independent experiments for calculation of the turnover rate  $k_{\text{obs}}$ .



**Figure 3.** Binding properties of mEndoV to inosine containing substrates. Bandshift assays with wild type or mutant D52A mEndoV (0.25–1  $\mu\text{M}$ ) with ss/ds IIUI-RNA, ss/ds I-RNA and ss/ds dIrG-DNA substrates were performed. Human EndoV (0.25  $\mu\text{M}$ ) was included in all assays. After incubation on ice for 15 minutes, bound RNA/DNA (=B) were separated from free substrates (=F) by 10% native PAGE.

**3D structure of mEndoV.** The structure determination of mEndoV (Table 3) revealed a protein fold which is to a large extent similar to human and prokaryotic EndoVs (Fig. 4a)<sup>12,26</sup>. The first 9 residues on the N-terminal end are disordered with no specific conformation. The region around residues 167–174 contains a mammalian specific  $\alpha$ -helix similar to that of hEndoV (Fig. 4a). The prokaryotic EndoVs are 11 residues shorter in this region, without formation of a helix (Fig. 5). A noticeable difference in the structure between the present mEndoV and previous reported EndoV structures is the conformation of the strand-separating wedge motif. This difference manifests itself in a quite open lesion binding pocket in mEndoV compared with Tm and hEndoV (Fig. 4b,c). The wedge motifs in hEndoV and also EcEndoV (PDB entry 4xpu)<sup>27</sup> are retracted about 4–6 Å compared with the DNA bound TmEndoV, but the conserved Tyr91 in hEndoV<sup>26</sup> is still pointing towards the DNA/RNA binding

Data Collection	
Beamline	ID23-2 (ESRF, Grenoble)
Wavelength (Å)	0.8729
Space Group	C222 <sub>1</sub>
Unit-cell parameters (Å, °)	$a = 101.63, b = 114.03, c = 54.01$
	$\alpha = \beta = \gamma = 90$
Resolution (Å)	32.47–1.75 (1.78–1.75) <sup>a</sup>
Unique reflections	32035 (1744)
Completeness (%)	99.9 (99.7)
Multiplicity	4.0 (3.9)
Mean I/σI	11.0 (3.0)
R <sub>meas</sub> (all I <sup>+</sup> and I <sup>-</sup> ) <sup>b</sup>	0.098 (0.537)
Refinement statistics	
Resolution of data used in refinement (Å)	32.47–1.75
Reflections used in refinement	31483
Completeness (%)	98.2
R <sub>cryst</sub> /R <sub>free</sub> (%) <sup>c</sup>	16.5/19.0
R.m.s.d. bonds (Å)	0.007
R.m.s.d. angles (°)	1.19
Average B-factor (protein/solvent) (Å <sup>2</sup> )	24.3/38.4
Number of atoms	
Protein	1896
Solvent	302
Ramachandran plot (%) <sup>d</sup>	
Favorable region	98.4
Additionally allowed	1.2
Outliers	0.4

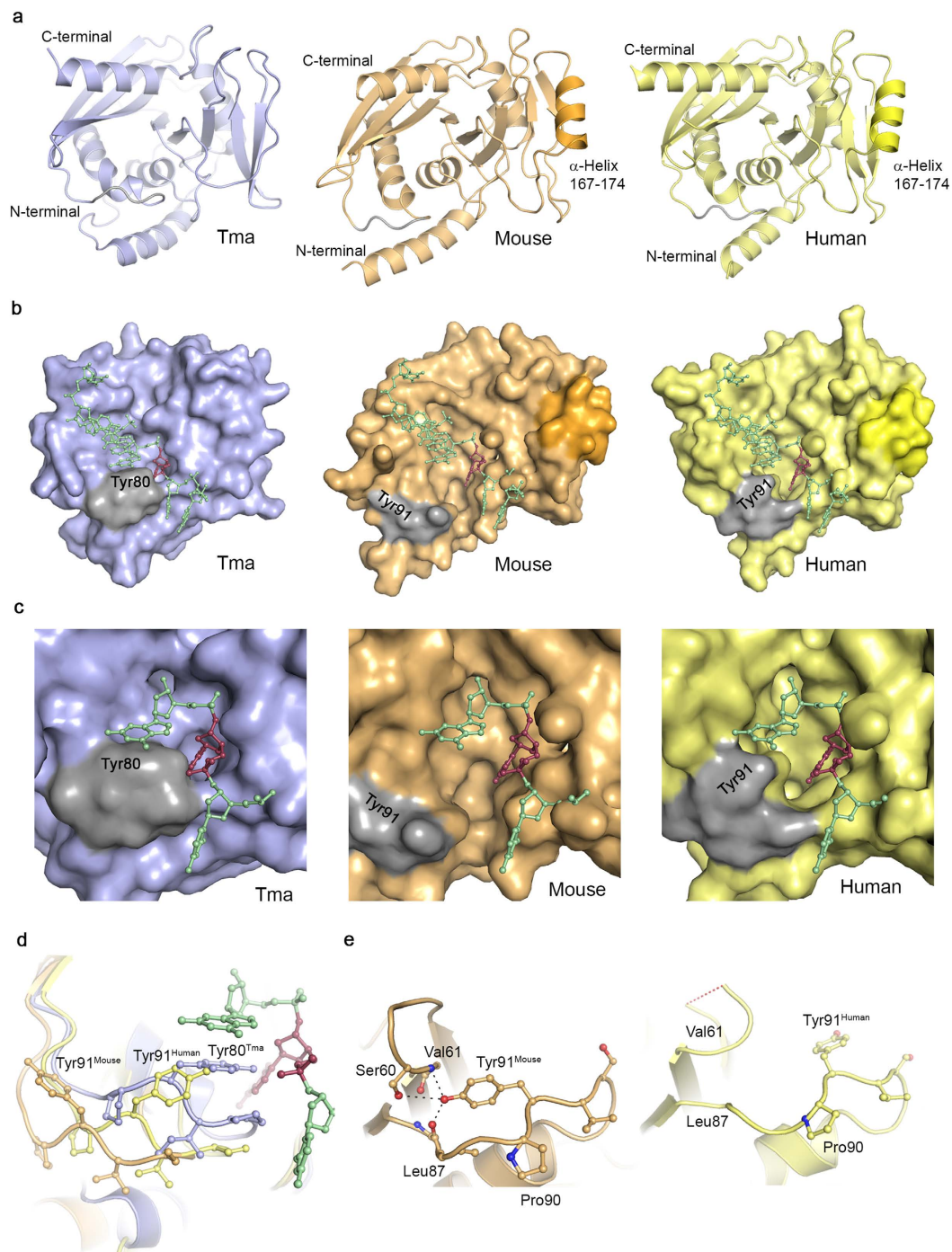
**Table 3. Crystal data, data-collection statistics and refinement data.** <sup>a</sup>Values in parentheses are for the highest resolution shells. <sup>b</sup>R<sub>meas</sub> defined by Diederichs & Karplus<sup>44</sup>. <sup>c</sup>R<sub>cryst</sub> =  $\sum hkl ||F_O| - |F_C|| / \sum hkl |F_O|$  where  $F_O$  and  $F_C$  are the observed and calculated structure factor amplitudes, respectively. R<sub>free</sub> is calculated from a randomly chosen 5.08% set of unique reflections not used in refinement. <sup>d</sup>Defined using MolProbity<sup>45</sup>.

groove, with side chain torsion angle  $\chi^1$  being  $\sim -170^\circ$  (*trans*) (Fig. 4d). Only a minor adjustment of the wedge segment is expected upon substrate binding. Contrary to this, in the structure of mEndoV, we find a different wedge conformation (Fig. 4d). The Tyr side chain points in the opposite direction, away from the DNA/RNA binding groove, with torsion angle  $\chi^1$  being  $\sim -40^\circ$  (*gauche*<sup>+</sup>). The mEndoV is the first example of an EndoV structure with a “flipped” wedge. In contrast to all other reported EndoV structures, the invariant Tyr91 in mEndoV is interacting with the backbone of the protein through hydrogen bonds to Ser60, Val61 and Leu87 (Fig. 4e). In order to stack with nucleotides in the active site pocket, the mEndoV Tyr91 must undergo a transition to the *trans* side chain conformation for  $\chi^1$ . To look further into the difference in wedge conformation we performed a molecular dynamics (MD) simulation of both mouse and human EndoV.

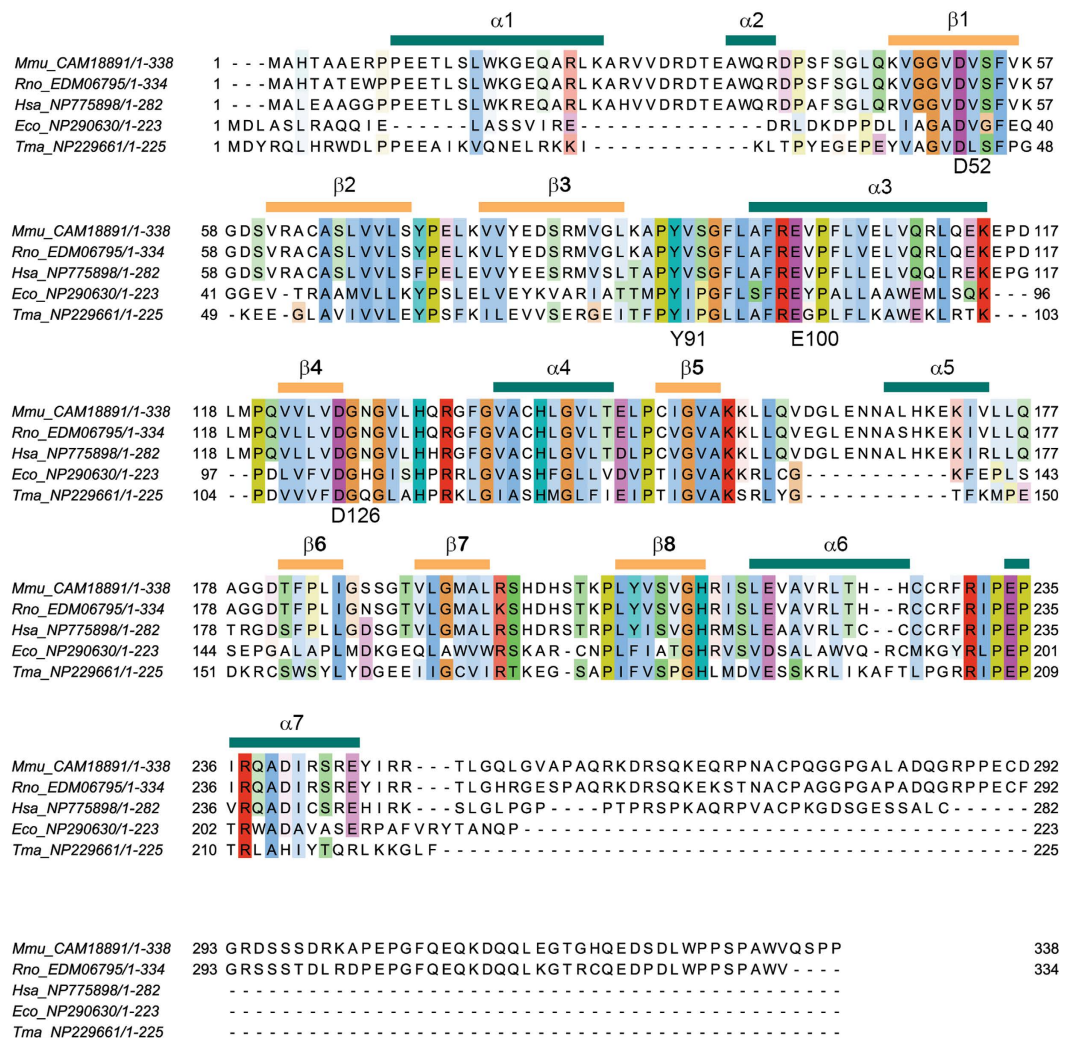
**Molecular dynamics simulation of the wedge flexibility.** The 100 nanoseconds simulation, with trajectory snapshots stored every 2.4 picosecond, reveal that Tyr91 in mEndoV moves between two side chain rotamers;  $\chi^1 \sim -60^\circ$  (*gauche*<sup>+</sup>) and  $\chi^1 \sim +60^\circ$  (*gauche*<sup>-</sup>), with a time distribution of about 3:1 (Supplementary Fig. S3). It never moves into the extended  $\chi^1 \sim 180^\circ$  (*trans*) conformation that is found in the crystal structure of human EndoV. For human EndoV, the Tyr91 side chain moves between the crystal structure starting point  $\chi^1 \sim 180^\circ$  (*trans*) and the  $\chi^1 \sim -60^\circ$  (*gauche*<sup>+</sup>), which corresponds to the conformation in the mouse EndoV structure. The side chain oscillates between the two conformations with a time distribution of about 1:1. These simulations support our structural data showing that the wedge motif has some plasticity and that transitions between different conformations can take place.

## Discussion

A series of prokaryotic and eukaryotic EndoV homologs have recently been studied with respect to catalysis, substrate specificity and structure<sup>12,21,26–28</sup>. The human homolog has activity for inosines in RNA, while the bacterial versions can cleave both RNA and DNA with inosines. Here we have characterized the activity profile of EndoV from mouse and present its 3D structure. Except for the structure of hEndoV, only bacterial EndoV enzymes have been structurally solved to date. Our data show that mEndoV shares substrate preferences with hEndoV by catalyzing cleavage at inosines in RNA and being inert towards inosines in DNA. Mouse EndoV seems to prefer



**Figure 4. Endonuclease V structures.** (a) Overall structure of *Thermotoga* (blue), mouse (orange) and human (yellow) EndoV. The mammalian versions used in crystallization have been truncated (mouse 1–250, human 13–250). The additional  $\alpha$ -helix common to the mammalian variants is colored in a stronger hue in mouse and human EndoV. (b) Surface display of *Thermotoga*, mouse and human EndoV. The wedge residues are colored gray; the surface contour of the mammal-specific helix is shown for mouse and human EndoV in a stronger hue. The lesion-strand in DNA in the complex with TmEndoV has been overlaid onto the mouse and human EndoV structures. The inosine residue is colored dark red. (c) Close-up view of the binding pocket regions of *Thermotoga*, mouse and human EndoV, illustrating the open character of mouse and human EndoV compared with the more closed TmEndoV. (d) The conformation of the wedge motif in *Thermotoga* (blue), mouse (orange) and human EndoV (yellow). The TmEndoV structure is that of the DNA bound form with the inosine lesion in dark red. (e) Close-up view of the Tyr91 conformation in mouse and human EndoV. The hydrogen bond between mouse Tyr91, Ser60, Val61 and Leu87 are shown (left) and compared with the similar open region in hEndoV (right). The red dashed line in the loop in hEndoV indicates disordered residues not present in the structural model.



**Figure 5.** Multiple sequence alignment of EndoV from mouse (*Mmu*), rat (*Rno*), human (*Hsa*), *E. coli* (*Eco*) and *T. maritima* (*Tma*), including a secondary structure element assignment calculated from the mouse structure. Helices are denoted  $\alpha 1$ – $\alpha 7$  and extended  $\beta$ -strands  $\beta 1$ – $\beta 8$ . The invariant Tyr91 and the three catalytic residues Asp52, Glu100 and Asp126 are labelled. The sequence reference codes for the NCBI protein database are listed. Sequences were aligned by Clustal Omega<sup>41</sup> and the figure was made with the JalView software. The secondary structure element assignment was calculated with DSSP/WhatIf<sup>42,43</sup>.

RNA substrates with clustered inosines, however the turnover rate for the best substrate tested, ss IIUI-RNA, is about 7-fold lower than that of hEndoV.

Both the activity and the bandshift analyses show that mEndoV favours substrates with multiple inosines. Recent publications of RNA sequencing data sets have revealed that also in the mouse transcriptome, the vast majority of A-to-I editing sites are found outside coding exons in intronic or untranslated regions consisting of inverted repeats<sup>29,30</sup>. It might be that mEndoV activity is adapted to hyperedited RNA rather than site-selective deamination sites. Under the reaction conditions used, both cleavage and substrate binding by mEndoV appeared less efficient than for hEndoV. This could be related to the relatively wide RNA/DNA binding groove and the orientation of the tyrosin (Tyr91) away from the active site as revealed by the mEndoV structure. In the TmEndoV structure this tyrosin is involved in base stacking with DNA, contributing to stabilisation of the enzyme-substrate complex. The much wider RNA binding cleft of mEndoV seems to allow more flexibility in the protein, possibly also with different substrate binding modes than TmEndoV. From the sequencing gel it appears that the second (middle) inosine in the hyperedited 5'...IIUI...3' substrate is preferred, probably by making more stable interactions with the enzyme using a binding conformation that positions the second inosine in the nucleotide binding pocket. In line with this, the strongest shift in EMSAs for mEndoV was found with this substrate.

The structure of mouse EndoV reveals a new conformation of the highly conserved strand-separating wedge motif. The invariant Tyr residue, normally pointing into the DNA/RNA binding groove and stacking with the DNA/RNA substrate during catalysis, is facing away from the groove. For this reason, the DNA/RNA binding groove in the crystal structure of mouse EndoV in the unbound state is wide and shallow. The MD simulation suggests that the wedge motif has some flexibility and that the invariant Tyr might flip between two conformations,

of which a probably unbound or unproductive state is shown in the structure of mEndoV. We cannot exclude that the observed difference in Tyr conformation and wedge backbone trace might be due to crystal packing effects, however, the MD simulation suggests that the wedge motif is more flexible than foreseen from previous structures, and that the present mouse EndoV structure has captured an ‘open’ state of the wedge.

The structure also unveils a surface-exposed  $\alpha$ -helix not present in the prokaryotic forms of EndoV but also found in the human homolog<sup>26</sup>. This element is on the outskirts of the DNA/RNA binding cleft, and may serve a regulatory or coordinative role in protein-protein interactions. Actually, an observed “head-to-shoulder” dimerization in the structure of EcEndoV is suggested to affect catalytic performance, as shown by reduced activity for the EcEndoV Ser144Ala single mutant and the Glu140Arg/Ser144Ala double mutant<sup>27</sup>. Intriguingly, the  $\alpha$ -helix unique to the mammalian variants corresponds to the proposed dimerization region in EcEndoV. In hEndoV, the quadruple-cysteine motif 225-CCCC-228 was shown to be important for catalysis, with serine mutations of Cys227 and Cys228 reducing the activity by 68 and 46%, respectively<sup>26</sup>. The molecular basis for the influence of these two cysteine residues on hEndoV catalysis is not known. The corresponding region in mEndoV is 225-HHCC-228, and our structure of mEndoV reveals that the first two Cys residues in hEndoV are replaced by histidines with no change in the local geometry. The latter two Cys residues are conserved between mouse and human EndoV, and with similar conformations and local environment in the two structures. Our present structure does not suggest a molecular explanation for the importance of these residues in catalysis, but it is tempting to suggest that the effect is due to protein stability in general, since EcEndoV and TmEndoV have strong activities even with Cys/Met and Phe/Thr in these positions, respectively.

In this study, we have shown that mouse EndoV is active on RNA substrates containing one or several inosines. In particular, our data revealed that hyperedited substrates, such as IIUI-RNA, are favored. The biological significance of the partition of activity between mammalian and prokaryotic EndoVs, operating on only RNA or both DNA and RNA respectively, is still to be investigated. The active site in mouse and human EndoV are rather similar to EcEndoV and TmEndoV, hence the difference in activity probably reside in substrate recognition (RNA versus DNA) or the specific need for a 2' hydroxyl group in the nucleotide 3' to the inosine aligned in the nucleotide binding pocket. Structures of EndoV in complex with RNA substrates will reveal detailed information on the molecular basis of RNA recognition and incision by mammalian EndoV.

## Methods

**Purification of mEndoV.** Codon optimized mouse EndoV cDNA (NP\_001164636.1) was cloned into pETM41 (EMBL, Heidelberg) in frame with the N-terminal His-maltose binding protein (His/MBP) tag (Genescript). Mouse EndoV proteins were expressed in a modified version of *E. coli* ER 2566 (ER2566  $\Delta nfi$   $\Delta rnhB$ ) where the endogenous *nfi* and *rnhB* genes had been inactivated by transformation with the kanamycin-deletion cassettes derived from the JW5547-1 (*nfi::kan*) and JW0178-1 (*rnhB::kan*) strains purchased from the Keio collection<sup>31</sup>. Correct integrations and subsequent excision of the cassettes by the use of the Flp recombinase encoded by pCP20 plasmid<sup>32</sup>, were verified by colony PCR. Cells were grown in Lysogeny broth (LB) medium, supplemented with 50 mg/l kanamycin at 37 °C until OD<sub>600</sub> reached 0.6–0.7. The temperature was lowered to 18 °C, protein expression induced with 1 mM IPTG and cells incubated at 18 °C over night. Cells were harvested by centrifugation, and the cell pellet dissolved in amylose buffer (50 mM Tris pH 8.0, 200 mM NaCl, 1 mM EDTA, 5% glycerol and 10 mM  $\beta$ -mercaptoethanol ( $\beta$ -ME)). Cells were sonicated (3  $\times$  30 s) followed by centrifugation (27 000 g, 30 min). The supernatant was incubated with amylose resin with gentle shaking at 4 °C and after 1 h transferred to a column. Fusion proteins were eluted with maltose-buffer (50 mM Tris pH 8.0, 200 mM NaCl, 1 mM EDTA, 10 mM  $\beta$ -ME, 5% glycerol, 10 mM maltose). Fractions were analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) and peak fractions pooled. After concentration (Amicon ultra-4, MWCO 10 kDa), proteins were applied to a Superdex 75 size-exclusion column using buffer 50 mM Tris pH 8.0, 50 mM NaCl, 10 mM  $\beta$ -ME. Fractions rich in EndoV were pooled and subjected to TEV protease cleavage (1:100 ratio in mg) by dialysis in 50 mM Tris pH 8.0, 1 mM DTT, 0.5 mM EDTA, 5% glycerol at 4 °C over night. To remove the released His-MBP tag, the protein solution was directly applied to a HiTrap Q ion exchange chromatography column. Free mEndoV did not bind to the HiTrap Q column and the flow through was then applied to a Resource S column after adjusting the final concentration of NaCl and MES pH 6.0 to 100 mM. Mouse EndoV was eluted in the flow through and stored at –80 °C in 20% glycerol. hEndoV was purified as previously described (ref 17).

For crystallisation, a site specific (D52A) truncated (residues 1–250) mutant of mEndoV fused to His/MBP in pETM41 was expressed in BL21 CodonPlus(DE3)-RIPL *E. coli* cells. The cells were grown in LB medium at 37 °C, and protein production was induced by the addition of 0.5 mM IPTG at OD<sub>600</sub> around 0.5–0.6 and incubation continued at 37 °C for 2 hours. Cells were harvested and the cell pellet was dissolved in cold buffer containing 50 mM Tris pH 8.0, 300 mM NaCl, 10 mM imidazole and 10 mM  $\beta$ -ME (buffer A). A cell-free extract was prepared by sonication and centrifugation. The extract was applied to a Ni-NTA agarose column and after washing extensively with buffer A, the His/MBP-mEndoV fusion protein was eluted using 50 and 300 mM imidazole in buffer A. The fractions rich in His/MBP-mEndoV were pooled, concentrated and supplemented with TEV protease (1:100 ratio) and dialyzed at 4 °C for 72 h against a buffer with 50 mM Tris pH 8.0, 10 mM  $\beta$ -ME and 0.5 mM EDTA. The protein solution was further dialysed against buffer A, and the free mEndoV was separated from TEV, His/MBP and uncleaved His/MBP-mEndoV by a second round of purification using Ni-NTA. The flow-through and wash fractions rich in mEndoV were concentrated as above and dialyzed to lower the pH in two stages using 50 mM MES pH 6.5, 100 mM NaCl and 10 mM  $\beta$ -ME, followed by 50 mM MES pH 6.0, 100 mM NaCl, and 10 mM  $\beta$ -ME. The protein was then applied to a Resource S column and eluted using a gradient in NaCl from 50 mM to 2M with 50 mM MES pH 6.0 and 10 mM  $\beta$ -ME. The pure mEndoV protein was concentrated to ~13 mg/ml for crystal screening.



**Cleavage and electrophoretic mobility shift assays.** Oligonucleotide substrates were from Eurofins (DNA) and Midland Certified (RNA) (Table 1). The oligonucleotides were end-labelled using T4 polynucleotide kinase (New England Biolabs) and [ $\gamma$ <sup>32</sup>P] ATP (3000 Ci/mmol, Amersham). Double stranded substrates were generated by annealing the labelled oligonucleotide to a complementary strand. 10  $\mu$ l reaction mixture contained 1 nM substrate, amount of enzymes as indicated and standard reaction buffer (10 mM Tris-HCl pH 7.5, 0.5 mM MnCl<sub>2</sub>, 50 mM KCl, 1 mM DTT and 5% glycerol), unless otherwise stated. Reactions proceeded at 37 °C for 30 min and stopped by adding 10  $\mu$ l formamide loading dye (80% formamide, 10 mM EDTA, 0.1% xylene cyanol and bromophenol blue). The oligonucleotides were denatured at 52 °C for 2 min, and the reaction products separated on 20% polyacrylamide/urea gels at 200 V for 1 h in 1x taurine buffer. The radiolabeled fragments were visualized by phosphorimaging (Typhoon 9410 Variable Mode Imager) and ImageQuant TL was used for quantification. All assays were performed 2–3 times and representative experiments are shown. Denaturing sequencing gels (20%) were run in 1x taurine at 35 W for 1 h.

For the single turnover assays, we first titrated 3–100 nM of mouse EndoV and the experiment showed that above ~25 nM enzyme concentration, we satisfy single-turnover conditions (data not shown). Thus, 30 nM EndoV was incubated with 1 nM substrate using the same reaction buffer as above. Samples were withdrawn as indicated in the figure and reactions stopped and analyzed as above. Samples were stored on ice until all time points were finished. For the calculation of the catalytic turnover rate  $k_{\text{obs}}$  (s<sup>-1</sup>), a one phase association model was fitted to three parallel data sets.

In electrophoretic mobility shift assays, enzymes (0.25–1  $\mu$ M) and substrates (1 nM) were mixed with band shift buffer (40 mM Tris-HCl pH 8.5, 10 mM CaCl<sub>2</sub>, 10 mM DTT, 6% glycerol) and tRNA (1 ng/ $\mu$ l) as a competitor. Samples were incubated on ice for 15 min, added DNA loading buffer (Thermo Scientific) prior to separation on 10% native polyacrylamide gels in 1x taurine with 5 mM CaCl<sub>2</sub> at 100 V for 40 min at room temperature. The enzyme-substrate complexes were visualized as above.

### Crystallization, diffraction data collection, structure determination and model refinement.

The D52A mutant of mEndoV was screened for crystallization using a Mosquito robot (TTP Labtech) and several commercial crystallization kits. Crystals appeared in condition 48 of the JCSG + kit, which contains 16% PEG 8000, 40  $\mu$ M KH<sub>2</sub>PO<sub>4</sub> and 20% by volume of glycerol. Crystals were transferred to a cryogenic solution made by mixing the reservoir solution with 10% ethylene glycol. After a short soak (<5 s) the crystals were flash-frozen in liquid nitrogen. X-ray diffraction images were collected at the ID23-2 beamline at the European Synchrotron Radiation Facility (ESRF). Data was collected at T = 100K using a wavelength of  $\lambda = 0.8729$  Å. The data was integrated, scaled and analyzed using MOSFLM<sup>33</sup>, Aimless<sup>34</sup> and CCP4i<sup>35</sup>. The structure was solved by molecular replacement using a poly-alanine version of the structure of *Thermotoga maritima* EndoV (PDB entry 4B20)<sup>36</sup>, as an input search template to Phaser<sup>37</sup>. The model was autobuilt into the mouse EndoV sequence using PHENIX<sup>38</sup>. The final refinement and model adjustments were done using PHENIX and Coot<sup>39</sup>. A few side chains have alternate positions, and these were modelled with 50% occupancy for each conformation. 8 residues in the N-terminus and 3 residues in the C-terminus are not visible in the electron density map, and are not included in the final model, and a few more residues at the C-terminal end have been modelled with 50% occupancy. Final atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession code 5aoy.

**Molecular dynamics simulation.** The preparation and simulations of mouse and human EndoV were carried out using the Desmond molecular dynamics (MD) program (version 4.4)<sup>39</sup> within the Maestro modelling software GUI from Schrödinger (release 2015-4). The crystal structures of mouse and human EndoV without solvent water molecules served as starting points for the MD simulation. For human EndoV, the missing residues Lys57 and Gly58 were filled in using Coot<sup>40</sup> and the loop was given an idealized geometry using the ‘Regularize zone’ tool of Coot. Each structure was pre-processed and energy minimized using the ‘Protein preparation’ tool in Schrödinger. Hydrogen atoms were added, the N- and C-terminal residues were capped and the hydrogen bond network optimized before energy minimization. The minimized protein structures were enclosed by solvated boxes filled with water molecules in addition to Na<sup>+</sup> and Cl<sup>-</sup> ions corresponding to a 150 mM buffer. The systems were again minimized and relaxed by short MD simulations using the default ‘Relax model system’ protocol in Desmond before starting 100 nanoseconds simulations with periodic boundary conditions. The simulation temperature was set to 300K, and both temperature and pressure were kept constant during the simulation (NTP ensemble simulation). Coordinates were stored every 2.4 picoseconds, forming simulation trajectories with more than 41000 structural snapshots. An inspection of the all-C <sub>$\alpha$</sub>  RMSD value for all structural snapshots relative to the starting point structure shows that both systems seem to equilibrate within the first 10–15 nanoseconds, with RMSD values fluctuating between 1.6–1.8 and 1.8–2.0 Å for mouse and human EndoV, respectively. There was a short period around 70 ns where the RMSD of human EndoV increased to ~2.8 Å over a 10 ns period, but the system subsequently returned to baseline. The trajectories were used to extract time-laps profiles of the side chain rotamers of Tyr91. The simulations were run on the Stallo computing cluster at the University of Tromsø, Norway, using 64 CPU-cores for each simulation.

### References

- Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715 (1993).
- Yasui, M. *et al.* Miscoding properties of 2'-deoxyinosine, a nitric oxide-derived DNA adduct, during translesion synthesis catalyzed by human DNA polymerases. *J. Mol. Biol.* **377**, 1015–1023 (2008).
- Engelward, B. P. *et al.* Base excision repair deficient mice lacking the Aag alkyladenine DNA glycosylase. *Proc. Natl. Acad. Sci. USA* **94**, 13087–13092 (1997).
- Hardeland, U., Bentele, M., Jiricny, J. & Schar, P. The versatile thymine DNA-glycosylase: a comparative characterization of the human, *Drosophila* and fission yeast orthologs. *Nucleic Acids Res.* **31**, 2261–2271 (2003).

5. Saparbaev, M. & Laval, J. Excision of hypoxanthine from DNA containing dIMP residues by the *Escherichia coli*, yeast, rat, and human alkylpurine DNA glycosylases. *Proc. Natl. Acad. Sci. USA* **91**, 5873–5877 (1994).
6. Guo, G., Ding, Y. & Weiss, B. *nfi*, the gene for endonuclease V in *Escherichia coli* K-12. *J. Bacteriol.* **179**, 310–316 (1997).
7. Yao, M., Hatahet, Z., Melamede, R. J. & Kow, Y. W. Purification and characterization of a novel deoxyinosine-specific enzyme, deoxyinosine 3' endonuclease, from *Escherichia coli*. *J. Biol. Chem.* **269**, 16260–16268 (1994).
8. Feng, H., Klutz, A. M. & Cao, W. Active site plasticity of endonuclease V from *Salmonella typhimurium*. *Biochemistry* **44**, 675–683 (2005).
9. Huang, J., Lu, J., Barany, F. & Cao, W. Multiple cleavage activities of endonuclease V from *Thermotoga maritima*: recognition and strand nicking mechanism. *Biochemistry* **40**, 8738–8748 (2001).
10. Kanugula, S., Pauly, G. T., Moschel, R. C. & Pegg, A. E. A bifunctional DNA repair protein from *Ferroplasma acidarmanus* exhibits O6-alkylguanine-DNA alkyltransferase and endonuclease V activities. *Proc. Natl. Acad. Sci. USA* **102**, 3617–3622 (2005).
11. Liu, J., He, B., Qing, H. & Kow, Y. W. A deoxyinosine specific endonuclease from hyperthermophile, *Archaeoglobus fulgidus*: a homolog of *Escherichia coli* endonuclease V. *Mutat. Res.* **461**, 169–177 (2000).
12. Dalhus, B. *et al.* Structures of endonuclease V with DNA reveal initiation of deaminated adenine repair. *Nat. Struct. Mol. Biol.* **16**, 138–143 (2009).
13. Yao, M. & Kow, Y. W. Interaction of deoxyinosine 3'-endonuclease from *Escherichia coli* with DNA containing deoxyinosine. *J. Biol. Chem.* **270**, 28609–28616 (1995).
14. Lee, C. C. *et al.* Endonuclease V-mediated deoxyinosine excision repair *in vitro*. *DNA Repair* **9**, 1073–1079 (2010).
15. Moe, A. *et al.* Incision at hypoxanthine residues in DNA by a mammalian homologue of the *Escherichia coli* antimutator enzyme endonuclease V. *Nucleic Acids Res.* **31**, 3893–3900 (2003).
16. Mi, R., ford-Zappala, M., Kow, Y. W., Cunningham, R. P. & Cao, W. Human endonuclease V as a repair enzyme for DNA deamination. *Mutat. Res.* **735**, 12–18 (2012).
17. Vik, E. S. *et al.* Endonuclease V cleaves at inosines in RNA. *Nat. Commun.* **4**, 2271 (2013).
18. Bazak, L. *et al.* A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes. *Genome Res.* **24**, 365–376 (2014).
19. Slotkin, W. & Nishikura, K. Adenosine-to-inosine RNA editing and human disease. *Genome Med.* **5**, 105 (2013).
20. Wolf, J., Gerber, A. P. & Keller, W. Tada, an essential tRNA-specific adenosine deaminase from *Escherichia coli*. *EMBO J.* **21**, 3841–3851 (2002).
21. Morita, Y. *et al.* Human endonuclease V is a ribonuclease specific for inosine-containing RNA. *Nat. Commun.* **4**, 2273 (2013).
22. Kuraoka, I. Diversity of Endonuclease V: From DNA repair to RNA editing. *Biomolecules* **5**, 2194–2206 (2015).
23. Alseth, I., Dalhus, B. & Bjoras, M. Inosine in DNA and RNA. *Curr. Opin. Genet. Dev.* **26**, 116–123 (2014).
24. Fladeby, C. *et al.* The human homologue of *Escherichia coli* Endonuclease V is a nucleolar protein with affinity for branched DNA structures. *PLoS. One.* **7**, e47466 (2012).
25. Scadden, A. D. & Smith, C. W. Specific cleavage of hyper-edited dsRNAs. *EMBO J.* **20**, 4243–4252 (2001).
26. Zhang, Z., Hao, Z., Wang, Z., Li, Q. & Xie, W. Structure of human endonuclease V as an inosine-specific ribonuclease. *Acta Crystallogr. D. Biol. Crystallogr.* **70**, 2286–2294 (2014).
27. Zhang, Z., Jia, Q., Zhou, C. & Xie, W. Crystal structure of *E. coli* endonuclease V, an essential enzyme for deamination repair. *Sci. Rep.* **5**, 12754 (2015).
28. Feng, H., Dong, L. & Cao, W. Catalytic mechanism of endonuclease V: a catalytic and regulatory two-metal model. *Biochemistry* **45**, 10251–10259 (2006).
29. Danecek, P. *et al.* High levels of RNA-editing site conservation amongst 15 laboratory mouse strains. *Genome Biol.* **13**, 26 (2012).
30. Porath, H. T., Carmi, S. & Levanon, E. Y. A genome-wide map of hyper-edited RNA reveals numerous new sites. *Nat. Commun.* **5**, 4726 (2014).
31. Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006 (2006).
32. Cherepanov, P. P. & Wackernagel, W. Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of FLP-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**, 9–14 (1995).
33. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 271–281 (2011).
34. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D. Biol. Crystallogr.* **69**, 1204–1214 (2013).
35. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.* **67**, 235–242 (2011).
36. Rosnes, I. *et al.* Structural basis of DNA loop recognition by endonuclease V. *Structure* **21**, 257–265 (2013).
37. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
38. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 213–221 (2010).
39. Kevin, J. *et al.* Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. *Proceedings of the ACM/IEEE Conference on Supercomputing (SC06)*. Tampa, Florida, November 11–17 (2006).
40. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501 (2010).
41. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**, 539 (2011).
42. Wouter, G. T. *et al.* A series of PDB-related databanks for everyday needs. *Nucleic Acids Res.* **43**, D364–D368 (2015).
43. Kabsch, W. *et al.* Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577–2637 (1983).
44. Diederichs, K. & Karplus, P. A. Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat. Struct. Biol.* **4**, 269–275 (1997).
45. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 12–21 (2010).

## Acknowledgements

We acknowledge financial support from The Norwegian Cancer Society (Grant No. 5739724) as well as the South-Eastern Norway Regional Health Authority (Grants No. 2012085 and 2015095; Regional Core Facility for Structural Biology, and 2014034 to BD). We are grateful for synchrotron travel support from The Research Council of Norway (216625/F50), The access to the computational resources at Stallo (UiT) was supported by Notur (project NN9309K). We would also like to thank the European Synchrotron Radiation Facility staff for help and beamtime at beamline ID23-2.

### Author Contributions

B.D., I.A. and M.B. conceived and designed the project; M.S.N., E.S.V., M.E.R., A.M.S., P.B., I.A. and B.D. performed the research; I.A. and B.D. analyzed the data; B.D. and I.A. wrote the paper. All authors reviewed the manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Nawaz, M. S. *et al.* Crystal structure and MD simulation of mouse EndoV reveal wedge motif plasticity in this inosine-specific endonuclease. *Sci. Rep.* **6**, 24979; doi: 10.1038/srep24979 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>