

Identification of a novel motif in DNA ligases exemplified by DNA ligase IV

Article (Unspecified)

Marchetti, Caterina, Walker, Sarah A, Odreman, Federico, Vindigni, Alessandro, Doherty, Aidan J and Jeggo, Penny (2006) Identification of a novel motif in DNA ligases exemplified by DNA ligase IV. *DNA Repair*, 5 (7). pp. 788-798. ISSN 1568-7864

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/17805/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Identification of a novel motif in DNA ligases exemplified by DNA ligase IV.

Caterina Marchetti^{1,2}, Sarah A Walker¹, Federico Odreman², Alessandro Vindigni²,
Aidan J. Doherty^{1*} and Penny Jeggo^{1*}.

¹Genome Damage and Stability Centre,
University of Sussex,
BN9 1RQ UK

²International Centre for Genetic Engineering and Biotechnology,
Padriciano 99, I-34012
Trieste, Italy.

Keywords: DNA ligases; LIG4 syndrome, DNA ligase structure, DNA repair, human syndromes.

*Correspondence should be addressed to:

P. A. Jeggo P.A.Jeggo@sussex.ac.uk

Tel 00 44 1273 678482

Fax: 00 44 1273 678121

A. J. Doherty A.J.Doherty@sussex.ac.uk.

Tel: 00 44 1273

Fax 00 44 1273 678121

ABSTRACT

DNA ligase IV is an essential protein that functions in DNA non-homologous end-joining, the major mechanism that rejoins DNA double strand breaks in mammalian cells. LIG4 syndrome represents a human disorder caused by mutations in DNA ligase IV that lead to impaired but not ablated activity. Five conserved motifs in DNA ligases have been previously identified. We previously reported G469E as a mutational change in a LIG4 syndrome patient. G469 does not lie in any of the previously reported motifs. A sequence comparison between DNA ligases led us to identify residues 468-476 of DNA ligase IV as a further conserved motif, designated motif Va, present in eukaryotic DNA ligases. We carried out mutational analysis of residues within motif Va examining the impact on adenylation, double-stranded ligation, and DNA binding. We interpret our results using the DNA ligase I:DNA crystal structure. Substitution of the glycine at position 468 with an alanine or glutamic acid severely compromises protein activity and stability. Substitution of G469 with an alanine or glutamic acid is better tolerated but still impacts upon activity and protein stability. These findings suggest that G468 and G469 are important for protein stability and provide insight into the hypomorphic nature of the G469E mutation identified in a LIG4 syndrome patient. In contrast, residues 470, 473 and 476 within motif Va can be changed to alanine residues without any impact on DNA binding or adenylation activity. Importantly however, such mutational changes do impact upon double stranded ligation activity. Considered in light of the DNA ligase I:DNA crystal structure, our findings suggest that motif Va functions as part of a molecular pincer that maintains the DNA in a conformation that is required for ligation.

Introduction

DNA non-homologous end-joining (NHEJ) is the major mechanism for the rejoining of DNA double strand breaks (DSBs) in mammalian cells [1]. Five proteins have been identified that function in NHEJ. The heterodimeric protein, Ku, binds to double stranded DNA ends, recruits and activates the catalytic subunit of the DNA-dependent protein kinase, DNA-PKcs. Together this complex constitutes the DNA-dependent protein kinase (DNA-PK). DNA-PK then recruits Xrcc4 and DNA ligase IV, which tightly co-associate. DNA ligase IV carries out the final rejoining step of NHEJ. In addition to playing a major role in the rejoining of radiation and endogenously induced DSBs, NHEJ also effects the rejoining step during V(D)J recombination [2]. Although Xrcc4 and DNA ligase IV are essential in mice, LIG4 syndrome has been identified as a human disorder conferred by hypomorphic mutations in DNA ligase IV [3-6]. The disorder is associated with clinical radiosensitivity and immunodeficiency, consistent with the known functions of DNA ligase IV. LIG4 syndrome patients also display developmental delay and microcephaly suggesting that DNA ligase IV plays an important role during development.

DNA ligase IV contains a conserved ligase domain at its N-terminus and a tandem BRCT domain at its C-terminus (Fig. 1a) [7]. Interaction with Xrcc4 requires the region between the BRCT domains and likely part of the BRCT domain [8,9]. The DNA ligase IV/Xrcc4 complex is large and its crystal structure has not yet been solved. Structural analysis has relied in part on the use of a range of other ATP-dependent DNA ligases as model systems [10-15]. The first step of ligation involves the formation of a covalent AMP-enzyme intermediate with AMP being attached to the enzyme via a highly conserved lysine residue. The second step involves formation of a DNA-adenylate complex followed finally by rejoining. All ATP-dependent DNA

ligases have a modular structure of two domains, an adenylation domain (AdD or Domain 1) and an oligo-binding domain (OBD or Domain 2) [12]. Six conserved motifs, designated motifs I, III, IIIa, IV, V and VI, have been identified among covalent nucleotide transferases, of which 5 are found in the AdD (Fig. 1a). Motifs I, III, IIIa, IV and V are essential for ATP binding and the auto-adenylation reaction. Motif I, encompassing the conserved lysine residue, forms the active site loop of the enzyme and constitutes part of the ATP binding pocket. Based on the crystal structures of a number of DNA ligase complexes, it has been proposed that ligases undergo profound conformational changes upon ATP binding and/or DNA binding [10,12,13,16,17]. This is exemplified by the rotation of the OBD. Motif VI lies within the OBD, distant from the active site on AdD [10]. However, upon ATP-binding this face of OBD moves towards the active site and residues, including those from motif VI, participate in the adenylation reaction [13,14,16,17]. Subsequently, the OBD moves away from the active site and swivels around placing motif VI far from the AdD, orientating the DNA-binding surface of OBD towards the now adenylated AdD [12,13,16,17]. This switching is essential for the catalytic cycle as it most likely prevents the formation of non-productive complexes between non-adenylated ligase and unnicked/unbroken DNA.

The larger eukaryotic ligases, such as LigI and LigIV, also possess an additional N-terminal DNA-binding domain (DBD) that is required for efficient ligation (Fig. 1c) and enables these ligases to encircle DNA [17]. An equivalent helix-hairpin-helix domain is also present in the bacterial NAD-dependent ligases [13,18]. The toroidal structure of LigI maintains the nicked DNA in a distorted conformation and locates the catalytic domain over the site of the nick prior to ligation [17].

Since DNA ligase IV is essential, the mutations identified in LIG4 syndrome are hypomorphic, that is, they confer residual activity. The analysis of the impact of such mutations is important to help evaluate the clinical impact and potentially to help direct patient care. Additionally, such mutations have the potential to provide novel insight into domains or motifs important for function. In one LIG4 syndrome patient an arginine was mutated to histidine within motif 1 (R278H) close to the active site lysine residue [19,20]. More recently, another mutational change (G469E) was identified [5]. Although this mutation lies adjacent to a conserved residue (G468), it lies outside of the six core motifs and, when first identified, we questioned whether it would alter the activity of the enzyme [21]. Our initial studies demonstrated, however, that it was a mutational change that impacted upon function [5]. A sequence comparison between DNA ligases strongly suggests that residues 468-476 represent a further conserved motif present in eukaryotic DNA ligases, which we have designated motif Va (Fig. 1b). The recently reported structure of DNA ligase I also provides evidence that this region might be important for function (Pascal, 2004 #10951). Here, we have undertaken, a mutational analysis of residues within motif Va to determine whether it was also important for function in DNA ligase IV. Mutational change of either residue G469 or G468 markedly reduced both adenylation and ligation activity. Since the mutant proteins were difficult to express our findings suggest that these two residues are important for protein stability and we consider this in the light of the DNA ligase I/DNA structure. Additionally, we found that residues 470-476 do not affect either DNA binding or adenylation activity but significantly impair double-stranded ligation activity. We present a model that these residues act, along with another conserved structural motif, as a pincer to facilitate the conformational change of the DNA to enhance catalysis.

Materials and Methods

Expression of wild type and mutant Xrcc4/DNA ligase IV complexes in insect cells.

The pFastBac vector (Invitrogen) expressing wild type histidine tagged human DNA ligase IV was as previously described.[20,22] Human Xrcc4 was obtained from pCI-neo-Xrcc4 by digestion with *EcoRI* and *NotI*, and inserted into the multiple cloning site of pFasBac Dual vector (Invitrogen) generating pFastBacDual-Xrcc4. Transcription was under the control of the polyhedron promoter. pFastBac vectors expressing histidine tagged DNA ligase IV mutants were generated by site directed mutagenesis. The DNA ligase IV coding sequence of the wild type and mutant constructs was sequenced to verify the presence of the mutation and the absence of non-specific mutations. The two recombinant baculoviruses expressing histidine tagged wild type or mutant DNA ligase IV and untagged Xrcc4 were co-expressed using the Bac-toBac Baculovirus expression system (Gibco).

Purification of WT and mutant DNA Ligase IV-XRCC4 Complexes.

The Sf9 (*Spodoptera frugiperda* ovary) cells were maintained at 27°C in SF-900 II medium (Invitrogen). Sf9 cells, at 1.5×10^6 cells/ml density, were co-infected with DNA ligase IV and Xrcc4 baculoviruses at the same multiplicity of infection. Ninety-six hours after infection, cells were harvested by centrifugation, resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 0.5 mM β -mercaptoethanol, 1% Nonidet P-40) supplemented with protease inhibitors (Roche Molecular Biochemicals) and incubated for 10 minutes at 4°C. 10 ml of lysis buffer was used per 100 ml of cell culture. The lysate was cleared by centrifugation and incubated with TALON metal affinity resin (Clontech) for 2 h at 4°C. The resin was washed with lysis buffer, then buffer (20 mM

Tris-HCl, pH 8, 0.5 mM β -mercaptoethanol, 12.5 mM imidazole) containing 500 mM NaCl (three washes) and finally buffer containing 100 mM NaCl (two washes). The histidine-tagged ligase IV in complex with untagged Xrcc4 was eluted in E buffer (20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM β -mercaptoethanol, 100 mM imidazole). Purified complexes were dialyzed and concentrated in buffer (20 mM Tris-HCl pH 8.0, 50 mM NaCl) on 0.5 ml Vivaspin columns (10 kDa cut-off size). Aliquots of the purified complexes were snap-frozen in liquid nitrogen and stored at -80°C in the presence of 10% glycerol. The purity of the preparation was verified by SDS-polyacrylamide gel stained with Colloidal Coomassie blue G-250 (Sigma).

The concentration of the ligase IV/Xrcc4 complex was determined by UV absorption measurements using an extinction coefficient at 280 nm of $121,330\text{ M}^{-1}\text{ cm}^{-1}$ estimated from the amino acid sequence (ProtParam, available at www.expasy.ch). The value obtained was confirmed by comparing the Coomassie Blue staining of the complexes with that of a known quantity of BSA after SDS-PAGE. The concentration of the G468E and G469E ligase IV/Xrcc4 complexes was only quantified by Coomassie Blue staining relative to wild type complexes. All other mutant complexes were quantified by both UV absorption and quantitative Coomassie Blue Staining.

Adenylation Assay

Wild type and mutant protein complexes were pretreated in buffer (20 mM Tris-HCl , pH 8.0, 50 mM NaCl) containing 5 mM disodium pyrophosphate for 15 min at room temperature. A Vivaspin 0.5 ml column (Sartorius) was used to remove pyrophosphate. Varying amounts of wild type and mutant protein complexes were incubated with $[\alpha\text{-}^{32}\text{P}]\text{ATP}$ (0.5 μCi , Amersham) in adenylation buffer (60mM Tris-

HCl, pH 8.0, 10 mM MgCl₂, 5 mM dithiothreitol) in 10 µl of reaction volume. Each reaction was incubated for 10 minutes at room temperature. Reactions were stopped by the addition of 1 volume of 2X standard SDS-polyacrylamide gel electrophoresis loading buffer. The samples were boiled for 5 min and run on a 9% SDS-polyacrylamide gel. Fixed and dried gels were analyzed and quantified using a STORM PhosphorImager (Molecular Dynamics).

Double-Stranded ligation assay.

A 442bp ds DNA fragment with 4bp overhangs at each end was produced from the Bluescript plasmid (Stratagene) following digestion with *PstI* and *AflIII* (New England Biolabs) and purification by electrophoresis using a 0.8% agarose gel. Following gel-extraction with Qiaquick DNA extraction kit (Qiagen), the 442bp substrate was treated with CIP and was 5' end-labelled with [γ -³²P] ATP.

The indicated amounts of protein complexes were incubated for 30 minutes in 30 µl of reaction mixture (50 mM triethanolamine, pH 7.5, 2 mM Mg(OAc)₂, 2 mM dithiothreitol, 0.1 mg/ml bovine serum albumin, 12% polyethylene glycol) with 20 ng of labelled DNA fragment. After incubation, the reactions were stopped by the addition of 1.5 µl of 10% SDS. Following deproteinization using the Qiaquick purification columns (Qiagen), the DNA was eluted in 30 µl of water. 15 µl of DNA were added to 1 volume of 2X standard DNA loading buffer, heated to 65°C for 5 minutes then cooled rapidly to 4 °C on ice. The DNA was electrophoresed on a 0.8% agarose gel. Dried gels were analyzed and quantified using Cyclon Storage Phosphor System (Packard).

Electrophoretic Mobility Shift Assays

For electrophoretic mobility shift assays (EMSA) studies, the DNA substrate was a 54bp DNA fragment with 4bp overhangs prepared by annealing two complementary synthetic oligonucleotides (54-mer). One 54-mer was labelled at the 5' end prior to annealing using T4-PNK and [γ -³²P] ATP. LX was incubated with the DNA (15 fmol) in 20 μ l of reaction mixture (10mM triethanolamine pH 7.5, 0.5 mM dithiothreitol, 0.1 mg/ml bovine serum albumin) at room temperature for 20 minutes. Loading buffer was added (20% Ficoll, 0.4% xylene cyanol), and the samples were subjected to 8% PAGE in 0.5 x TAE buffer (20 mM Tris acetate, 0.5 mM EDTA), pH 8.1 at 4°C. Dried gels were analyzed using Cyclon Storage Phosphor System (Packard).

Results

Identification of motif Va in DNA ligases

G469E is a mutational change identified in a LIG4 syndrome patient. Further examination of the amino acid sequence in the vicinity revealed the presence of a motif well conserved between DNA ligases I, III and IV and less stringently conserved in more distantly related DNA ligases (Fig. 1a). G468 is highly conserved in all ATP-dependent ligases [21]. The tyrosine, glycine, glycine and arginine residues at positions 3, 5, 7 and 9 are also well conserved (Fig. 1b). Examination of the crystal structure of DNA ligase I revealed that this motif, designated motif Va (Fig. 1b), lies in a loop located on the surface of the OBD (Fig. 1c). Following DNA binding, motif Va makes close contact with the DNA (Fig. 1c). This motif may, therefore, be important for transactions of DNA ligase IV with DNA during catalysis. To examine this further, we created DNA ligase IV cDNAs expressing mutational changes as shown in Fig. 2a, expressed the mutant proteins in baculovirus (Fig. 2b) and examined their activity using *in vitro* assays for adenylation, double stranded ligation, and DNA binding.

Mutations at residues 470-476 do not impact upon adenylate complex formation.

Previously, we observed that Xrcc4 can be stably expressed and solubilised in insect cells in the absence of DNA ligase IV but not vice versa. Xrcc4 stimulates DNA ligase IV activity. Therefore, to ensure that we monitor the activity of DNA ligase IV complexes, we designed baculovirus constructs expressing histidine-tagged DNA ligase IV and non-tagged Xrcc4. After co-expression of wild type (WT) or mutant proteins in insect cells, we purified ligase IV/Xrcc4 complexes exploiting the histidine tag on DNA ligase IV. All complexes contained similar molar ratios of DNA ligase

IV: Xrcc4 demonstrating, as predicted, that mutations in motif Va do not impair interaction with Xrcc4 (Fig. 2). The WT DNA ligase/Xrcc4 complex and mutant Y470A, K473A, and R476A ligase IV/Xrcc4 complexes were expressed at similar levels whereas expression of the G468A and G469A DNA ligase IV/Xrcc4 mutant complexes was routinely slightly lower. The G468E and G469E DNA ligase IV/Xrcc4 mutant complexes were poorly expressed and had higher levels of non-specific bands, many of which represent breakdown products, relative to the more readily expressed proteins. Similar relative levels of mutant complexes were employed in the assays described below (see Materials and Methods). For the G468E and G469E complexes, quantification was assessed by Coomassie staining (as shown in Fig. 2). Insect cell expressed DNA ligase IV/Xrcc4 complexes are pre-adenylated. We, therefore, monitored the ability of the mutant ligase IV/Xrcc4 to form adenylate complexes following treatment with inorganic pyrophosphate (PPi) to generate non-adenylated complexes. Mutation of either G468 or G469 to glutamic acid completely abolished adenylate complex formation. Changing either residue to alanine also significantly impaired adenylate complex formation, although residual activity was readily detectable with the G469A mutant complex and at high protein concentration with the G468A mutant protein. In stark contrast, Y470A, K473A and R476A mutant complexes were proficient in adenylate complex formation (Fig. 3).

These findings suggest that a glycine at positions 468 and 469 is required for adenylate complex formation whereas the remaining residues in motif Va do not play critical roles in DNA ligase IV adenylate complex formation.

Mutational changes within motif Va impact upon DNA ligase IV double-stranded ligation activity

The same mutant proteins were examined for double-stranded (ds) ligation activity using a 442 bp double stranded DNA fragment with 4 bp overhangs as a substrate. Since adenylation is required before ligation can ensue, we used insect cell expressed complexes without pyrophosphate treatment to monitor ds ligation in the absence of ATP. The G468E and G469E mutant complexes had little or no detectable ligation activity (Fig. 4). Surprisingly, G469A despite having significantly impaired adenylation activity, showed normal ligation activity (see below). The ds ligation activity of the remaining mutant complexes, despite being competent for auto-adenylation, were significantly impaired although all proteins retained residual activity. Together these findings provide evidence that motif Va is required for efficient end-joining.

The observation that G469A has normal ligation activity despite reduced adenylation activity is most likely explained by the fact that DNA ligase IV/Xrcc4 complexes expressed in insect cell are pre-adenylated. In support of this we have observed that adenylation activity is only detectable after treatment with PPI suggesting that the majority of the DNA ligase IV/Xrcc4 complexes expressed in insect cells are pre-adenylated [22] (data not shown). Additionally, we have found that non-adenylatable DNA ligase IV complexes are expressed at low levels in insect cells. It seems likely, therefore, that pre-adenylated DNA ligase IV/Xrcc4 complexes have enhanced stability relative to non-adenylated complexes and accumulate in insect cells. Since none of the mutant complexes showed any adenylation activity in the absence of PPI treatment (data not shown), we consider it likely that even mutant complexes with low adenylation activity *in vitro*, accumulate as adenylate complexes

in insect cells [20]. G469A may, therefore, be expressed as an adenylated complex and be efficient for ligation despite showing impaired *in vitro* adenylation activity.

Motif Va is dispensable for DNA binding

To determine if the diminished ligation activity of the mutant complexes could be attributed to impaired DNA binding capacity, we used a 54 bp ds DNA fragment in electrophoretic mobility shift assays (EMSA) to examine the ability of the motif Va mutant proteins to bind to ds DNA. It has been shown previously that WT DNA ligase IV/Xrcc4 complex is able to bind efficiently to a DNA fragment of 54bp length [22]. We observed that Y470A, R476A and K473A mutant complexes were proficient for DNA binding (Fig. 5). Mutant complexes involving G468 and G469 were not analysed in this assay due to difficulties in obtaining sufficiently concentrated protein for analysis.

Discussion

Previously, we reported the identification of a mutation, G469E, in a LIG4 syndrome patient which severely impacts upon DNA ligase IV function [5]. Closer examination of the sequence in the vicinity of G469 revealed a motif, designated Va, that is well conserved between DNA ligases (Fig 1b). Mutational analysis of residues within motif Va highlights an important role in ligation. G468 represents a highly conserved residue, as noted previously [21] which, we show, is required for adenylate complex formation and for catalysis. In contrast, residues 470, 473 and 476 which lie within the loop region encompassing motif Va are dispensable for adenylate complex formation and for DNA binding but are required for efficient DNA ligation. During the course of this work, the crystal structure of the Ligase I:DNA complex was reported [17], enabling us to directly evaluate the observed mutational effects at the molecular level.

G468 is buried within the OBD at a position of close packing between β sheets (Fig. 6a). Glycine is an atypical amino acid lacking a side-chain. It is likely that substitution of this residue by other amino acids would severely compromise the stability of the OBD. Indeed, substitution of a glutamic acid, which has a large side-chain resulted in a protein that appeared to be poorly expressed, likely due to impaired stability. Substitution of the slightly smaller alanine was slightly better tolerated but protein expression was still low. The mutant proteins showed little functional activity either for adenylation or ligation activities consistent with a significant impact on conformation. G469, which is slightly less buried in the OBD, also appears to be important for protein conformation. Significantly, G469A is better tolerated than other substitutions at the 468/469 residues and yields a protein with measurable residual

adenylation activity. This is consistent with the less stringent conservation of this residue, and indeed, DNA ligase I has an alanine at this site. The near normal ligation activity of G469A suggests that once adenylate complex formation occurs, protein stability may be enhanced and the protein can carry out double-stranded ligation efficiently. The substitution by a residue with a bulkier side-chain (glutamic acid) is less well tolerated at this position. Together, this analysis suggests that G469E generates a protein with significant reduced conformational stability but retaining some residual activity providing an explanation for the viability but significant clinical features of the patient.

In contrast to G468 and G469, which are buried (Fig. 6a), residues 470-476 are equivalent to the surface loop L12 present in the crystal structure of the LigI complex (Fig. 6b) [17]. The OBD of DNA ligases adopts at least two distinct orientations during the catalytic cycle. In the adenylation step, the face of the OBD which contains motif VI swivels towards the AdD following ATP binding formation, acting as a “lid” over the active site, with a number of OBD residues directly participating in the adenylation reaction. In the DNA binding/ligation step, the OBD lid opens up (following adenylation) and the domain pivots around, orienting the DNA binding face including motif Va towards the active site, positioning motif VI away from the AdD. The OBD of LigI binds directly to DNA [17] interacting with the minor groove adjacent to the nicked DNA (Fig. 6b). Pascal *et al.* have reported that the OBD alters the curvature of the DNA backbone, conferring an A-form conformation (with an expanded minor groove) on the DNA upstream of the nick and maintaining a B-form on the downstream side [17]. A number of regions of LigI are critical for preserving the B-type DNA conformation including the loops L12, L45 and α -helix S. Surface loop L12 (motif Va) binds to the template strand of DNA

(Fig. 6b), making direct contacts via a number of residues, including Arg771 (R476 in Lig IV) that extends from L12 into the minor groove (Fig. 6b). Our findings reveal that residues in the loop region of motif Va do not impact upon adenylation but are required for double-strand ligation activity. Taken together with the crystal structure of Lig I, we suggest that the Motif Va loop functions in conjunction with helix S that binds to the opposite strand, as a molecular pincer that senses the dimensions of the minor groove (Fig. 6b) and actively maintains the B-form that is essential for efficient ligation. Consistent with this model, it has been shown that the presence of an A-form structure downstream of a nick severely affects ligation efficiency [17,23,24]. Although motif Va may be important for modulating the conformation of the DNA, it does not appear to be critical for DNA binding, which is likely due to the fact that the DNA makes many contacts with DNA ligase IV involving all three domains [17]. Interestingly, a mutation A771W in DNA ligase I has been described in a patient [25]. A771 lies within the equivalent motif in ligase I. This mutation does not affect adenylate complex formation but does affect nick ligation consistent with our model [26]. Notably, mice expressing this mutation exhibit an increased predisposition to cancer [27]. Interestingly, although DNA ligases have different substrate specificities and DNA ligase III and IV can rejoin DNA/RNA substrates in contrast to DNA ligase I, motif Va appears conserved between all DNA ligases, suggesting a role distinct from substrate specificity. Our findings suggest that a role in promoting DNA conformation may be a conserved function of motif Va although differences in this region between ligases might relate to substrate specificity.

In conclusion, our analysis provides insight into the impact of a mutational change observed in a LIG4 syndrome patient suggesting that G469E impacts upon protein conformation and/or stability in a manner conferring residual but impaired

function. We also demonstrate that residues 470-476 are important for efficient double-stranded ligation whilst being dispensable for adenylation and DNA binding. We discuss our results in light of the Ligase I:DNA crystal structure and suggest that Motif Va plays an important role in promoting conformational changes of the DNA that facilitate efficient catalysis.

Acknowledgements:

Work in the PAJ laboratory is supported by the Medical Research Council, the Leukaemia Research Fund, the Human Frontier Science Program, the International Agency for Cancer Research and the European Community Grant Number FIGH-CT-200200207. AJD is a Royal Society University Research Fellow and is supported by grants from the Biotechnology and Biological Sciences Research Council (BBSRC), Association for International Cancer Research (AICR) and Cancer Research UK (CR-UK). CM was supported by a Mobility Project and a D4 Project fellowship from the AREA Science Park Consortium for Research and Technology of Trieste (Italy). The work in AV laboratory is supported by a grant from the Human Frontier Science Program, by an FIRB grant of MIUR (Ministero dell'Istruzione dell'Universita' e della Ricerca), and by grant no. 02.00648.ST97 of Consiglio Nazionale delle Ricerche, Rome.

References

- [1] P.A. Jeggo The mechanism of DNA non-homologous end-joining: Lessons learned from biophysical, biochemical and cellular studies, in: K.W. Caldecott (Ed.), Eukaryotic DNA damage surveillance and repair, Eureka.com and Kluwer Academic/Plenum Publishers, 2004, pp. 146-158.
- [2] G.E. Taccioli, G. Rathbun, E. Oltz, T. Stamato, P.A. Jeggo and F.W. Alt Impairment of V(D)J recombination in double-strand break repair mutants, *Science* 260 (1993) 207-210.
- [3] K.M. Frank, J.M. Sekiguchi, K.J. Seidl, W. Swat, G.A. Rathbun, H.L. Cheng, L. Davidson, L. Kangaloo and F.W. Alt Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV, *Nature* 396 (1998) 173-177.
- [4] D.E. Barnes, G. Stamp, I. Rosewell, A. Denzel and T. Lindahl Targeted disruption of the gene encoding DNA ligase IV leads to lethality in embryonic mice, *Curr. Biol.* 8 (1998) 1395-1398.
- [5] M. O'Driscoll, K.M. Cerosaletti, P.-M. Girard, Y. Dai, M. Stumm, B. Kysela, B. Hirsch, A. Gennery, S.E. Palmer, J. Seidel, R.A. Gatti, R. Varon, M.A. Oettinger, K. Sperling, P.A. Jeggo and P. Concannon DNA Ligase IV mutations identified in patients exhibiting development delay and immunodeficiency, *Mol. Cell* 8 (2001) 1175-1185.
- [6] T.I. Ben-Omran, K. Cerosaletti, P. Concannon, S. Weitzman and M.M. Nezarati A patient with mutations in DNA Ligase IV: Clinical features and overlap with Nijmegen breakage syndrome, *Am. J. Med. Genet* 137 (2005) 283-287.

- [7] P. Robins and T. Lindahl DNA Ligase IV from HeLa cell nuclei, *J. Biol. Chem.* 271 (1996) 24257-24261.
- [8] U. Grawunder, D. Zimmer and M.R. Lieber DNA ligase IV binds to XRCC4 via a motif located between rather than within its BRCT domains, *Curr. Biol.* 8 (1998) 873-876.
- [9] B.L. Sibanda, S.E. Critchlow, J. Begun, X.Y. Pei, S.P. Jackson, T.L. Blundell and L. Pellegrini Crystal structure of an Xrcc4-DNA ligase IV complex, *Nat. Struct. Biol.* 8 (2001) 1015-1019.
- [10] H.S. Subramanya, A.J. Doherty, S.R. Ashford and D.B. Wigley Crystal structure of an ATP-dependent DNA ligase from bacteriophage T7, *Cell* 85 (1996) 607-615.
- [11] M. Odell, V. Sriskanda, S. Shuman and D.B. Nikolov Crystal structure of eukaryotic DNA ligase-adenylate illuminates the mechanism of nick sensing and strand joining, *Mol Cell* 6 (2000) 1183-1193.
- [12] A.J. Doherty and S.W. Suh Structural and mechanistic conservation in DNA ligases, *Nucleic Acids Res.* 28 (2000) 4051-4058.
- [13] J.Y. Lee, C. Chang, H.K. Song, J. Moon, J.K. Yang, H.K. Kim, S.T. Kwon and S.W. Suh Crystal structure of NAD(+)-dependent DNA ligase: modular architecture and functional implications, *EMBO J* 19 (2000) 1119-1129.
- [14] V. Sriskanda and S. Shuman *Chlorella* virus DNA ligase: nick recognition and mutational analysis, *Nucleic Acids Res.* 26 (1998) 525-531.
- [15] V. Sriskanda and S. Shuman Role of nucleotidyl transferase motif V in strand joining by *Chlorella* virus DNA ligase, *J. Biol. Chem.* 277 (2002) 9661-9667.

- [16] K. Håkansson, A.J. Doherty, S. Shuman and D.B. Wigley X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes, *Cell* 89 (1997) 545-553.
- [17] J.M. Pascal, P.J. O'Brien, A.E. Tomkinson and T. Ellenberger Human DNA ligase I completely encircles and partially unwinds nicked DNA, *Nature* 432 (2004) 473-478.
- [18] A.J. Doherty, S.R. Ashford and D.B. Wigley Characterization of proteolytic fragments of bacteriophage T7 DNA ligase, *Nucleic Acids Res.* 24 (1996) 2281-2287.
- [19] E. Riballo, S.E. Critchlow, S.H. Teo, A.J. Doherty, A. Priestley, B. Broughton, B. Kysela, H. Beamish, N. Plowman, C.F. Arlett, A.R. Lehmann, S.P. Jackson and P.A. Jeggo Identification of a defect in DNA ligase IV in a radiosensitive leukaemia patient, *Curr. Biol.* 9 (1999) 699-702.
- [20] E. Riballo, A.J. Doherty, Y. Dai, T. Stiff, M.A. Oettinger, P.A. Jeggo and B. Kysela Cellular and biochemical impact of a mutation in DNA ligase IV conferring clinical radiosensitivity, *J. Biol. Chem.* 276 (2001) 31124-31132.
- [21] Y.-F. Wei, P. Robins, K. Carter, K. Caldecott, D.J.C. Pappin, G.-L. Yu, R.-P. Wang, B.K. Shell, R.A. Nash, P. Scar, D.E. Barnes, W.A. Haseltine and T. Lindahl Molecular cloning and expression of human cDNAs encoding a novel DNA ligase IV and DNA ligase III, an enzyme active in DNA repair and recombination, *Mol. Cell. Biol.* 15 (1995) 3206-3216.
- [22] B. Kysela, A.J. Doherty, M. Chovanec, T. Stiff, S.M. Ameer-Beg, B. Vojnovic, P.M. Girard and P.A. Jeggo Ku stimulation of DNA ligase IV-dependent ligation requires inward movement along the DNA molecule, *J. Biol. Chem.* 278 (2003) 22466-22474.

- [23] V. Sriskanda and S. Shuman Specificity and fidelity of strand joining by Chlorella virus DNA ligase, *Nucleic Acids Res.* 26 (1998) 3536-3541.
- [24] P. Liu, A. Burdzy and L.C. Sowers DNA ligases ensure fidelity by interrogating minor groove contacts, *Nucleic Acids Res* 32 (2004) 4503-4511.
- [25] D.E. Barnes, A.E. Tomkinson, A.R. Lehmann, A.D.B. Webster and T. Lindahl Mutations in the DNA ligase 1 gene of an individual with immunodeficiencies and cellular hypersensitivity to DNA damaging agents, *Cell* 69 (1992) 495-503.
- [26] C. Prigent, M.S. Satoh, G. Daly, D.E. Barnes and T. Lindahl Aberrant DNA repair and DNA replication due to an inherited enzymatic defect in human DNA ligase I, *Mol. Cell. Biol.* 14 (1994) 310-317.
- [27] C. Harrison, A.M. Ketchen, N.J. Redhead, M.J. O'Sullivan and D.W. Melton Replication failure, genome instability, and increased cancer susceptibility in mice with a point mutation in the DNA ligase I gene, *Cancer Res.* 62 (2002) 4065-4074.

Figures

Figure 1. Sequence and structure of DNA ligase motif Va.

- a) The primary domain structure of DNA ligase IV highlighting the positions of the conserved catalytic motifs and the newly identified motif Va.
- b) Sequence conservation of motif Va. G469E is the mutational change identified in a LIG4 syndrome patient.
- c) The crystal structure of human DNA ligase I:DNA complex shows that motif Va forms a loop-like structure (blue) on the surface of the Oligo-Binding domain (OB; yellow) and this motif makes direct contact with the DNA (grey). The three conserved domains present in mammalian ligases are also highlighted.

Figure 2. Mutational changes studied and expression of wild type and mutant DNA ligase IV/Xrcc4 complexes in insect cells.

- a) Position of mutational changes studied.
- b) Coomassie Colloidal Blue staining of dialyzed and concentrated wild type and mutant DNA ligase IV/Xrcc4 complexes purified from baculovirus-infected insect cells. Values given are in μg of protein. Molecular mass standards (in kilodaltons) are shown on the left.

Figure 3. Adenylation activity is impaired in G468 and G469 mutant complexes.

- a) The wild type and the mutant DNA ligase IV/Xrcc4 complexes shown in Figure 2 were examined for adenylation complex formation following treatment with pyrophosphate and incubation with $[\alpha\text{-}^{32}\text{P}]$ ATP. After separation by SDS-PAGE, labelled ligase-AMP adducts migrated as a single band and were detected by autoradiography. G468E and G469E mutant complexes had no detectable adenylation

activity. Activity was significantly impaired but not ablated in G468A and G469A mutant complexes. Nearly normal adenylation activity was observed with complexes carrying mutations within the motif Va loop. Values given are ng of protein. The amount of G468E and G469E mutants DNA ligase IV/Xrcc4 complexes used in the reactions was assessed after Coomassie blue staining to give levels equivalent to 100, 200, 400 and 800 ng of wild type DNA ligase IV/Xrcc4 complexes.

b) Quantification of the adenylation activity of wild type and mutant DNA ligase IV/Xrcc4 complexes. The specific activities of wild type (WT) and mutant complexes were determined from the slopes of the titration curves in the linear range of enzyme dependence. The activity values for the mutant complexes were normalized to the wild type specific activity. For the G468A the residual adenylation activity of 1000 μ g of complex was compared to the activity of the same amount of wild type. The results represent the mean of three experiments.

Figure 4. All motif Va mutant complexes show reduced double-stranded ligation activity.

Double stranded (ds) ligation was assayed using wild type and mutant DNA ligase IV/Xrcc4 complexes in the absence of pyrophosphate treatment and ATP.

a) A DNA substrates of 442 bp length with 4 bp overhangs was incubated with the complexes and the ligated DNA multimers of different length were separated in an agarose gel. The assay has been carried out using an increasing amount of protein (values given are ng of protein). G468E and G469E mutant complexes showed dramatically reduced ligation activity. In marked contrast the G469A mutant complex

showed nearly normal activity. The remaining complexes showed significantly reduced activity.

b) Quantification of the double stranded ligation activity in wild type and mutant DNA ligase IV/Xrcc4 complexes. The percentage of converted substrate is plotted as a function of the quantity of DNA ligase IV/Xrcc4 complexes (logarithmic scale). The result represents the mean of three experiments.

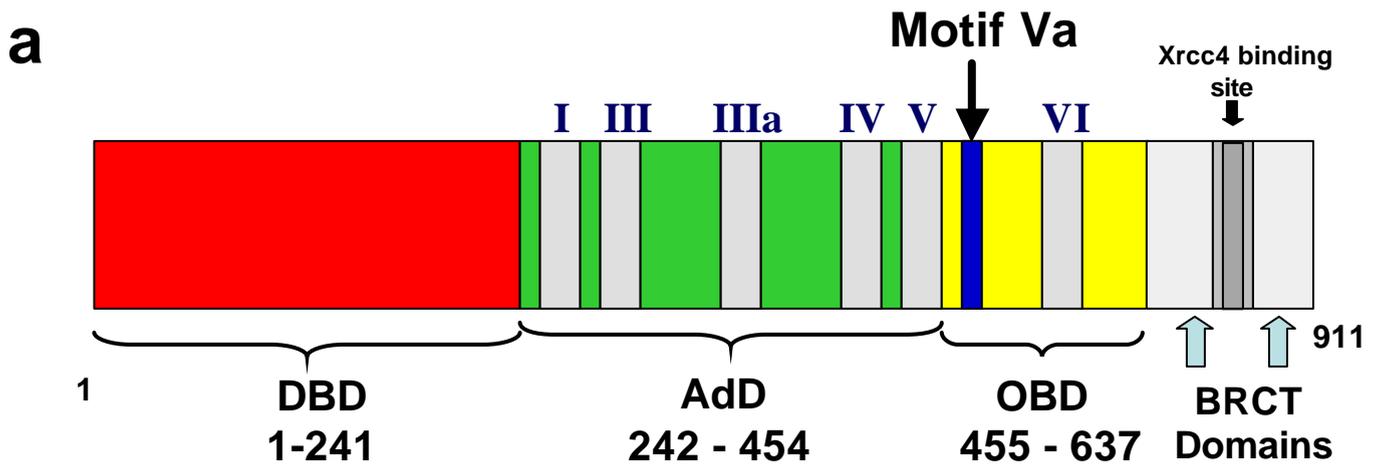
Figure 5. Motif Va mutant complexes show normal DNA binding.

DNA binding of motif Va mutant complexes was assessed using EMSA with a 54bp DNA fragment. G468E, G469E, G468A and G469A mutant complexes were not assessed for DNA binding due to their diminished purity or lower concentration, respectively. The remaining complexes showed nearly normal levels of DNA binding.

Figure 6. Structural aspects of motif Va

a) The buried position of the most conserved glycine residue (G468, black) of Motif Va is shown in the crystal structure of Lig I complex.

b) The structure of the LigI complex reveals that a number of regions are critical for preserving the B-form conformation of the DNA, including the loops L12 (blue), L45 (not shown) and α -helix S (cyan). L12 (motif Va) binds to the template strand of DNA (orange), making direct contacts, including Arg771 (R476 in Lig IV) that extends from L12 into the minor groove.



b

Hs Lig IV	468	GGYWGKGSR	476
Mus Lig IV	468	GGYWGKGSR	476
Gal Lig IV	473	GGYWGKGSR	481
Ath Lig IV	458	GGYYGSGRR	466
Hs Lig I	763	GAYLGRGKR	771
Sce Lig I	615	GAYYGRGKR	623
Spo Lig I	612	GAYYGKGR	620
Hs Lig III	609	GA FYG QGSK	617
Vac Lig I	418	GAYYGKGA K	426
T4 Lig	379	GIYPHRKTD	385
T7 Lig	258	GL...ANE	262
Consensus		GAYYGKGR	

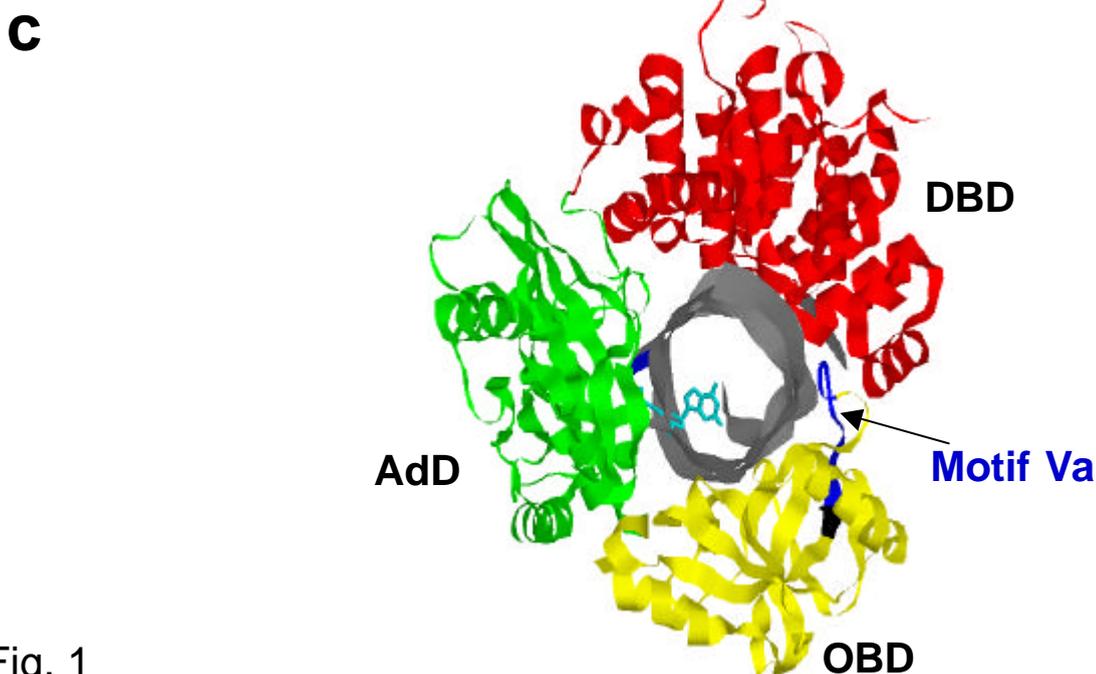


Fig. 1

a

Motif	GGYWGKGSR
G469E	G E YWGKGSR
G468E	E GYWGKGSR
G469A	G A YWGKGSR
G468A	A GYWGKGSR
Y470A	GG A WGKGSR
K473A	GGYWG A GSR
R476A	GGYWGKGS A

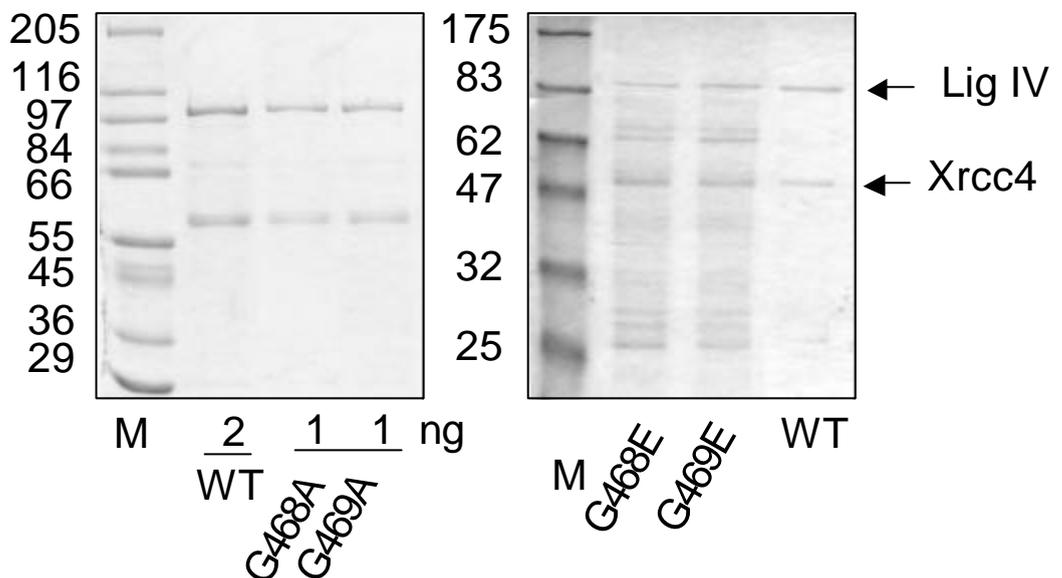
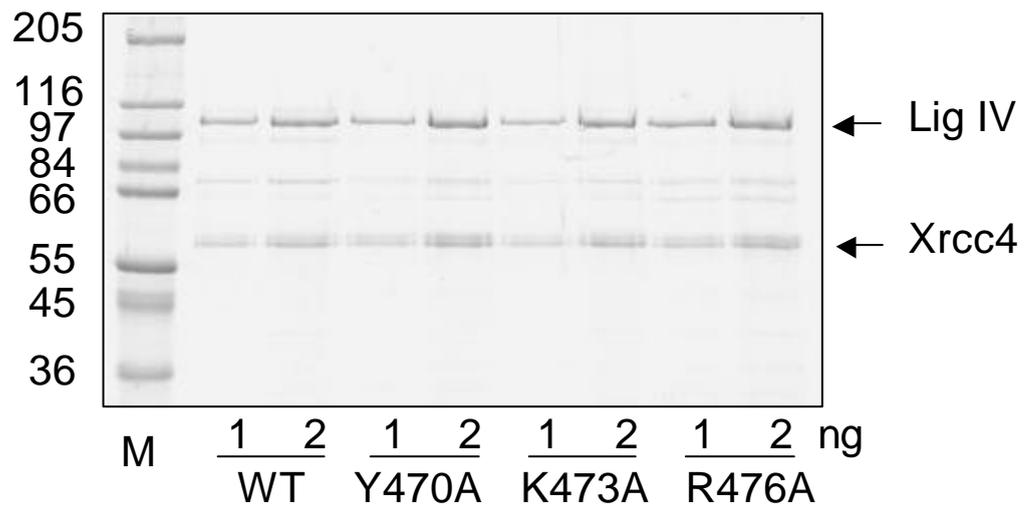
b

Fig. 2.

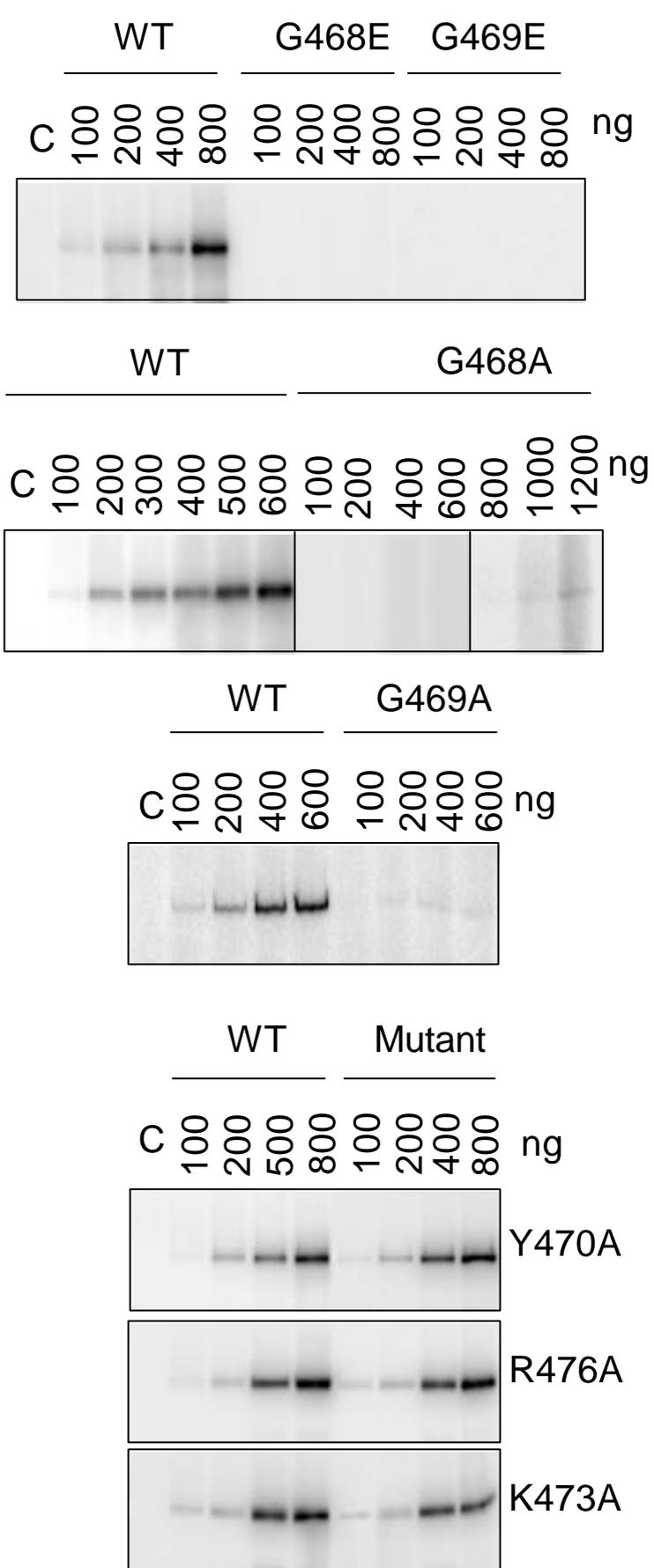


Fig 3a

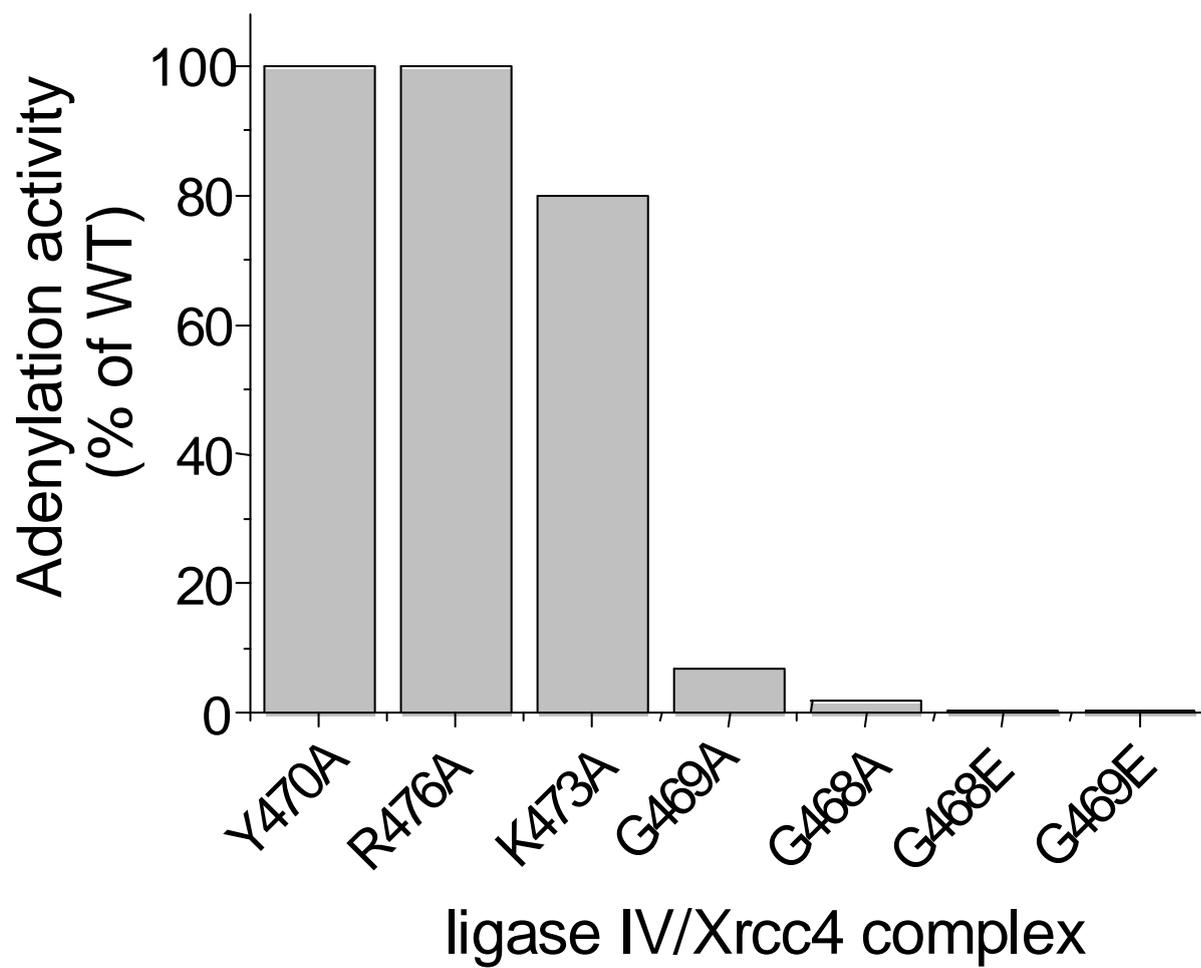


Fig 3b

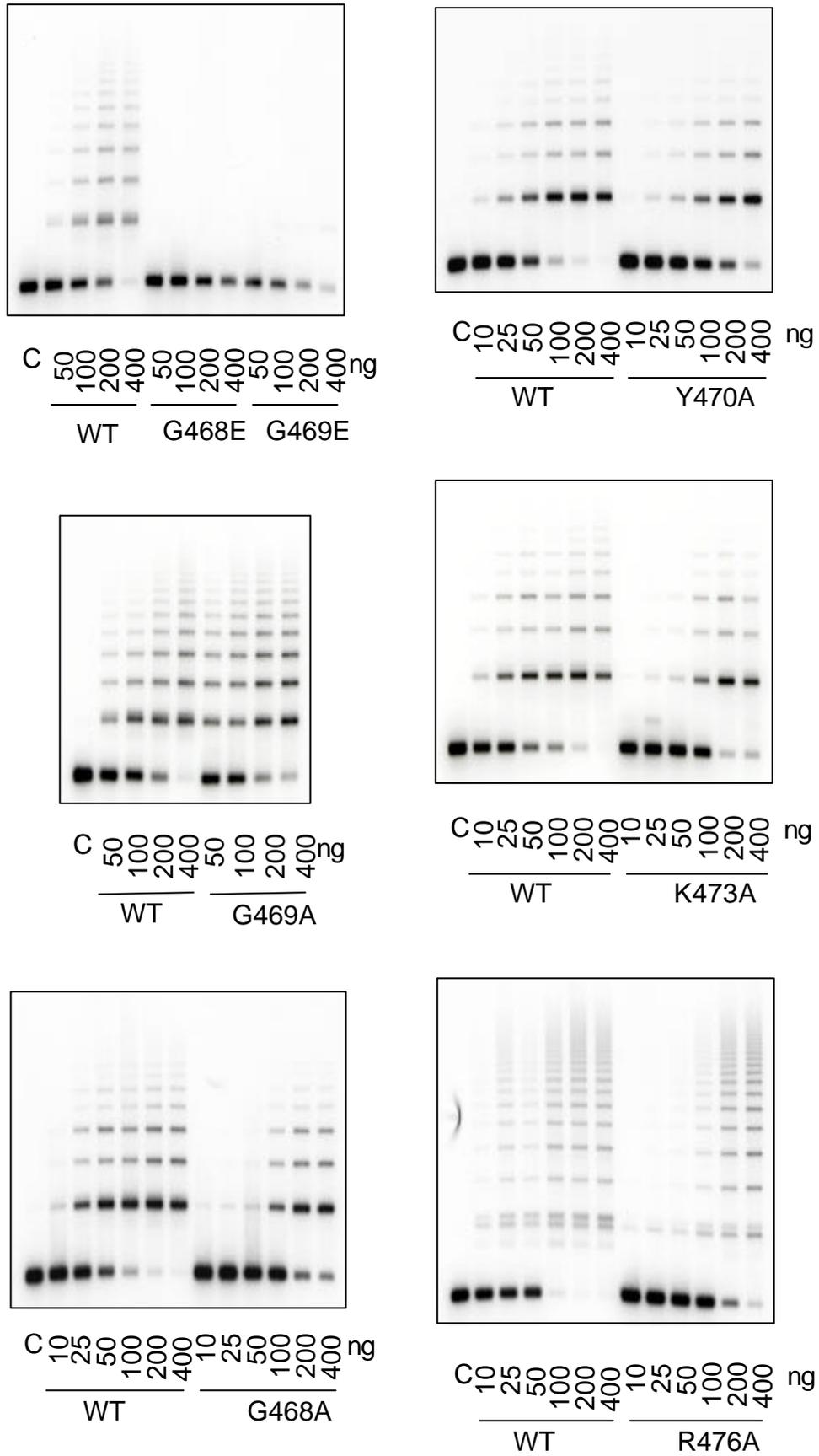


Fig 4a

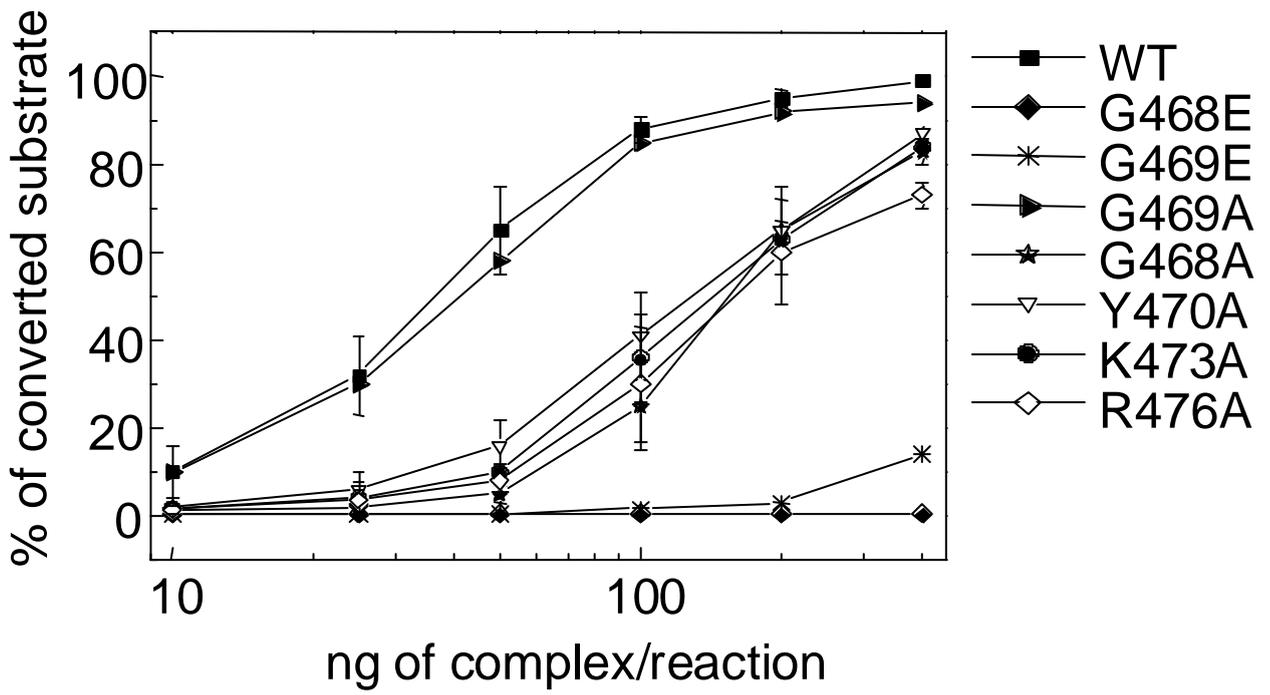


Fig 4b

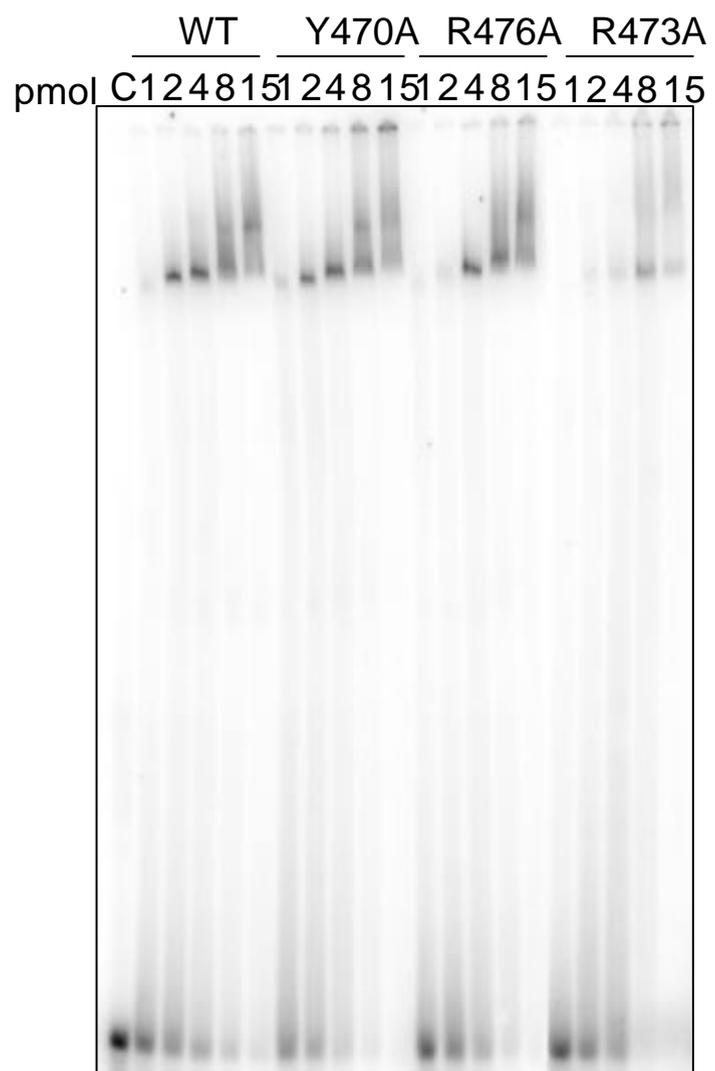
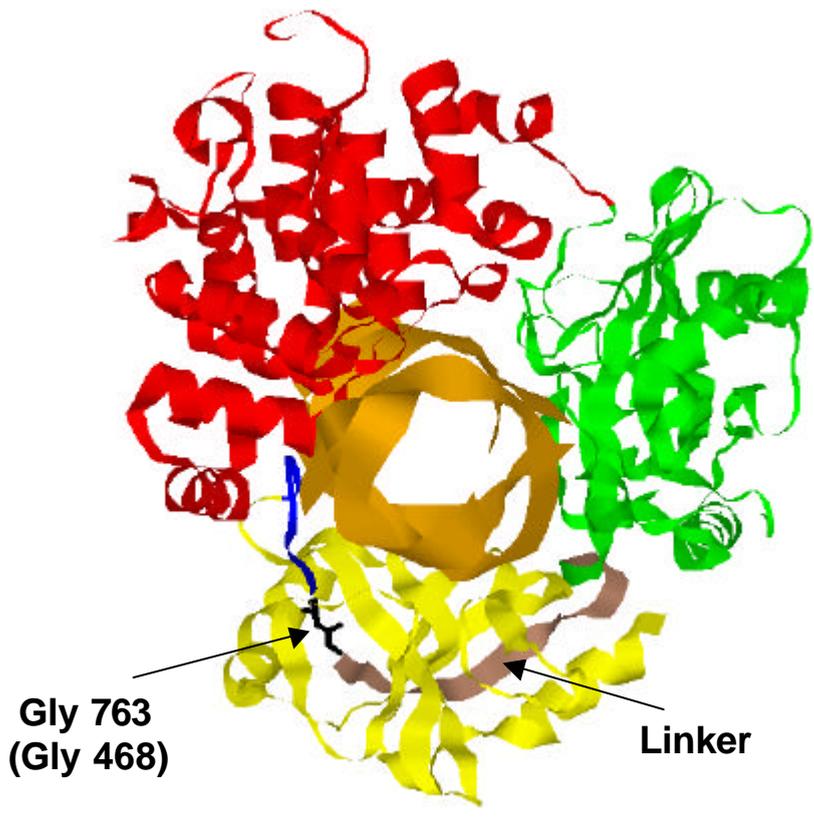


Fig 5

a



b

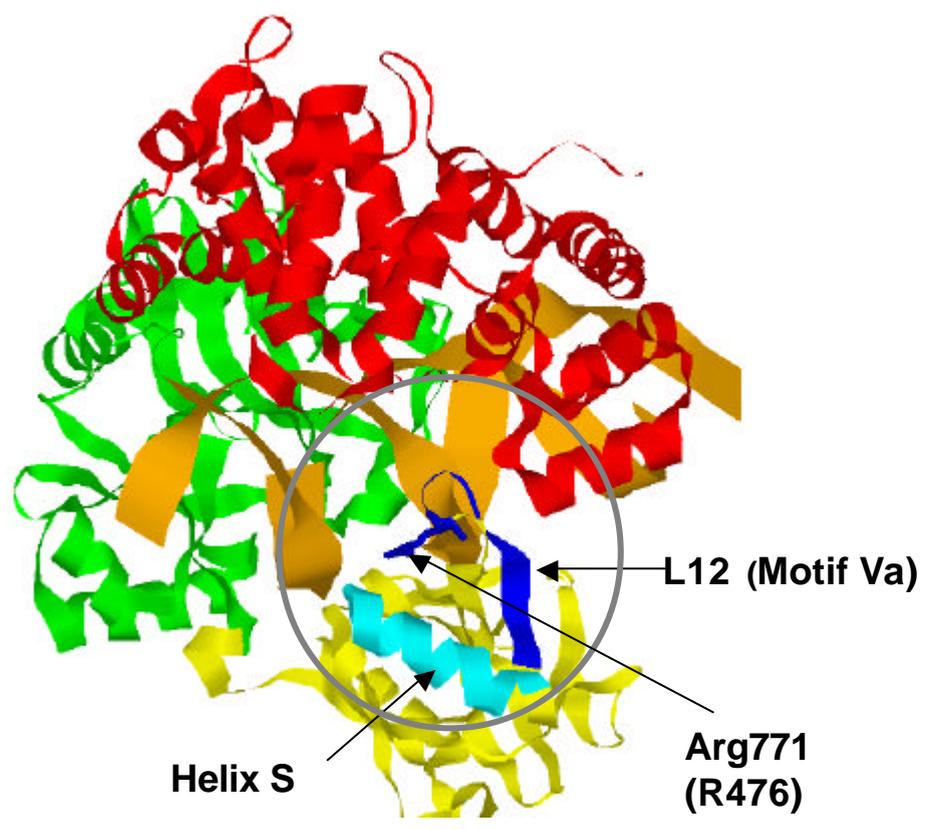


Fig. 6