The BRCT domain from the large subunit of human Replication Factor C

Kobayashi, Masakazu

Chapter 3

Amino acid determinants for DNA binding by the BRCT region of human RFC p140 subunit

Abstract

Sequence alignment of the BRCT region in the p140 subunit of Replication Factor C from various eukaryotes has identified very few conserved amino acid residues within the core BRCT domain, while none were found in sequences immediately N-terminal to the BRCT domain. However, mapping of the limited number of conserved residues that were found, onto a homology model of the BRCT domain revealed a cluster of conserved residues on one side of the molecular surface. Site-directed mutagenesis was performed on these residues as well as some lesser conserved amino acids of the N-terminal motif of the BRCT region of human p140. The structural changes due to the introduced mutations were rigorously checked using 1D ^{1}H-NMR and CD spectroscopy of the purified proteins. Mutation of weakly conserved residues in the region N-terminal to the BRCT domain generally resulted in reduced DNA binding. Mutation of any of the conserved residues within the BRCT domain negatively affected DNA binding. This latter group of residues is conserved amongst the BRCT domains from the bacterial DNA ligases, which recently have been shown to bind DNA. Furthermore, these conserved residues are positionally equivalent to those required for phosphate recognition by the tandem BRCT domains from BRCA1 suggesting that they may be involved in binding the 5' phosphate of dsDNA.

M, Kobayashi. F, Figaroa and G. Siegal (Parts of this chapter and of chapter 2 will be submitted for publication)

Introduction

The BRCT domain (**BR**CA1 **C**-**T**erminus) was first identified as a tandem repeat of roughly 90 amino acids at the C- terminus of the BRCA1 (Breast Cancer susceptibility 1) protein. Extensive amino acid sequence profiling led to the discovery of a vast number of proteins (current entry 915 open reading frames in Pfam) carrying the BRCT domain and, strikingly, most of those characterized are either directly or indirectly associated with various aspects of DNA metabolism including: DNA repair, DNA replication or cell cycle-checkpoint (1). Despite the low sequence similarity, all the BRCTs with known structures share the same overall protein fold (2-4). BRCT domains have been reported to be responsible for homotypic (BRCT-BRCT) or heterotypic (BRCT-non BRCT) protein-protein interactions. For example, in the crystal structure of the BRCT domains of XRCC1 (5), there is evidence that a homo-dimer is formed through both hydrophobic and salt-bridge interactions, the same that mediate hetero dimer formation between the BRCT domains of XRCC1 and DNA ligase III (6). Additionally, recent studies have revealed that some BRCT domains bind specifically to a phosphoserine, pointing towards a role in signal transduction events that regulate the various DNA repair pathways (7-9). The subsequent structure determination of a complex between the tandem BRCT repeat of BRCA1 and a phosphopeptide derived from BACH1 revealed the detailed molecular basis of how the phosphopeptide is specifically recognized (10-12).

Replication factor C (RFC) is a five subunit complex that plays an important role in efficient loading of PCNA onto primer-template DNA during the synthesis of the daughter strands in DNA replication. RFC consists of four subunits of 35-40 kDa and a fifth large subunit (p140) of 140 kDa. The C-terminus of p140 shares homology with the four small subunits, while the unique N-terminal sequence (roughly 50 kDa), which contains a single BRCT domain, is dispensable for PCNA loading. This N-terminal sequence has been shown to bind dsDNA at the 5'-phosphorylated terminus. The region of the sequence responsible for this DNA binding activity was mapped and it includes the single BRCT domain (13). Truncation studies indicated that the BRCT domain from human RFC p140 cannot bind to DNA by itself but requires an extra 28 amino acids at its N-terminus (Figure 2.1). Similar requirements have also been reported for *Drosophila* RFC p140 (13), therefore extending the 5'-phosphate binding activity to what here is referred to as the "BRCT region".

While the majority of BRCT domains are, or are thought to be, bona fide protein-protein interaction modules, only a limited number appear to bind DNA. The BRCT region of RFC p140 is a member of this small class of DNA binding BRCT domains. In order to further characterize its DNA binding properties, site-directed mutagenesis has been performed on the BRCT region of RFC p140 to delineate potential residues involved. Sequence alignment and homology modeling of the human BRCT region revealed an interesting cluster of conserved residues on one side of the molecular surface. Amino acid substitution of any of these residues resulted in mutants with low or undetectable DNA binding activity. Despite the lack of absolute amino acid conservation, mutations of the moderately conserved amino acids within the N-terminal sequence of the BRCT region also diminished DNA binding activity, further supporting the direct involvement of these sequences in DNA binding.

Materials and Methods

Site directed mutagenesis

The procedure for generating point mutations was adopted from the QuickChange method (Stratagene) with some modifications. 125 ng of each primer (table 2) and 25 ng of template DNA, pET20b RFCp140 375-480, were used in the reaction. The PCR program used consisted of the following: 1 cycle of 95^{o}C for 30 seconds and 18 cycles of steps consisting of 30 sec 95^{o}C, 2 min 55^{o}C, 8 min 68^{o}C. After PCR, 10 U of DpnI was added to the reactions and incubated 1h at 37^{o}C. Subsequently 10 U of Pfu polymerase was added to the reaction and incubated 30 min at 72^{o}C, which was followed by addition of 10 U of DpnI and further incubation for 1h at 37^{o}C. The reaction mix was then finally transformed into *E.coli* DH5α supercompetent cells. The plasmids carrying the desired mutations were identified by the unique silent restriction site introduced via the mutagenesis primers and were further verified by DNA sequencing. The (R388A) mutation, was generated using SOE (Splice Overlap Extension)(14).

Protein expression and purifications

E.coli strain BL21 (DE3) pLysS was transformed with pET20b RFC p140 (375-480) bearing one of the 9 site-directed mutations and grown in LB medium containing 4 μg/ml Ampicilin and 34 μg/ml Chloramphenicol. Protein expression was induced with 1mM IPTG and the cells were harvested after 3 hours. The mutant proteins were purified

using the identical protocol used for the wildtype protein described in Chapter 2. The concentration of mutants was measure using UV absorbance at 277 nm (ε_{277nm} = 7.32 $mM^{-1}cm^{-1}$). For the mutants Y382A and Y385E, extinction coefficients were calculated according to the amino acid contents using ProtParm (http://au.expasy.org/tools/protparam.html). The extinction coefficient of the native mutants was obtained by comparing the absorbance at 277 nm measured at native and unfolded mutants in 8M urea.

Measurement of 1D ^{1}H-NMR and circular dichroism spectra

Each mutated protein was exchanged into a buffer consisting of 25mM Tris-HCl pH 7.5, 50mM NaCl and 1mM DTT. Protein concentration varied from 0.05 to 0.1mM as determined by UV absorbance at 277 nm. 1D ^{1}H-NMR spectra of each protein were recorded on a 500 MHz Bruker AVANCE instrument at 298 K (25 °C) and processed with TOPSPIN 1.3 (Bruker). For the circular dichroism measurements, the purified mutants were exchanged into 10 mM Potassium Phosphate buffer (pH 7.7) using PD10 columns (Amersham Bioscience). Protein concentration was adjusted to 10 μM (K397E, T415A, R423A, K458E and wildtype), and circular dichroism (CD) spectra were measured at 21 °C with scanning 1 nm increment between 195-260 nm in 0.1 mm path-length cuvette using a Jobin Yvon CD6 instrument. For each protein, five spectra were recorded to obtain an average spectrum from which the background spectrum of the buffer was subtracted.

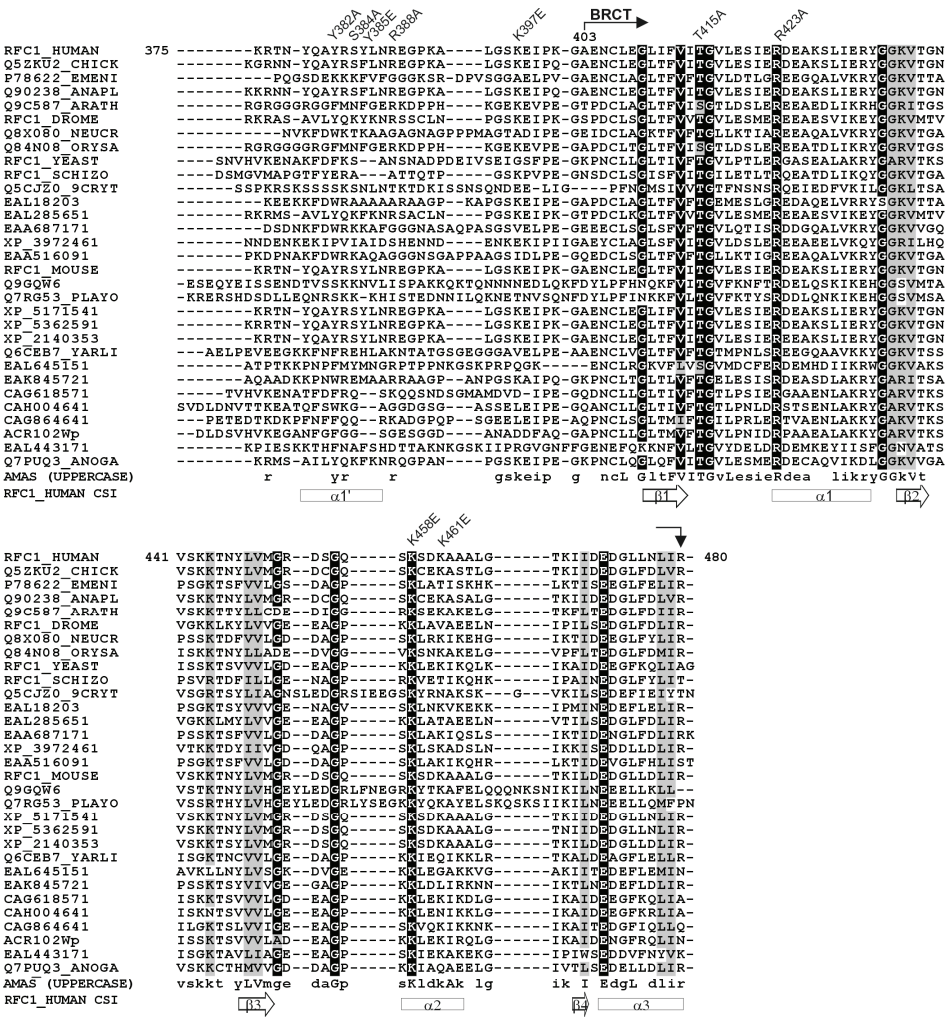
Detection of DNA-protein complex

The sequence of the DNA used in the binding assay was designed to form a hairpin with 10 base pairs (5'pCTCGAGGTCG**TCAT**CGACCTCGAGATCA 3', the letters in bold will form the loop). The 5' was radioactively labeled by exchange using T4 polynucleotide kinase and [γ-^{32}P] ATP as a substrate and purified as described in Methods Chapter 2. The mutant proteins were diluted in a buffer of 10 mM HEPES pH 7.8, 2 mM $MgCl_2$, 0.1 mM EDTA, 100 μg/ml BSA, 15 % glycerol, 0.8 μg/ml poly (dI-dC) and 2 mM DTT. Detection of DNA-protein complex was performed using the identical gel retardation assay as described in the Methods of Chapter 2. Radioactively labeled DNA was detected using film (Kodak Xomat XAR) or a phosphor-imager (BioRad) for quantification. The fraction of DNA in the DNA-protein complex was calculated as; Percentage of total shift (%) = [complex – background] / [total activity per lane - background] x 100, where total activity = complex + free DNA.

Results

Design of mutations via amino acid sequence alignment and homology modeling of RFC p140 (375-480)

The BRCT region is defined the sequences between aa 375 and 480 of the human RFC (or "p140(375-80)") whereas the region between 403 and 480 (or "p140(403-480)") is referred to as the BRCT domain. The BRCT region binds dsDNA with a 5'phosphorylated terminus (Figure 2.2). In order to find potential residues involved in the recognition of DNA, a sequence alignment of the BRCT region in the p140 subunit of 31 eukaryotic RFC complexes was analyzed, allowing the identification of a few highly conserved residues (Figure 3.1). The sequence alignment was generated using Clustal W (15;16) and the residue conservation in the alignment was further analyzed using AMAS (Analysis of Multiply Aligned Sequence)(17), which determines the degree of conservation for each position of the consensus on the basis of the amino acid properties. A limited number of conserved residues were identified within the BRCT domain while the N-terminal sequence of the BRCT region lacked any conservation among all 31 species (Figure 3.1) despite its indispensable nature in DNA binding (Figure 2.2). The amino acid backbone carbon and proton chemical shifts of p140(375-480) bound to dsDNA have been determined by NMR and with such values, one can reliably predict the secondary structure content in the protein by comparison with the chemical shift in the random coil state (Chapter 4, supplementary materials Figure 1S). This prediction suggested that the non-conserved N-terminal sequence contains a helix (α1') which is followed by a long loop that connects it to the BRCT domain (Figure 3.1 bottom). The BRCT domain itself consists of the consensus secondary structure (β1-α1-β2-β3-α2-β4-α3) commonly found in BRCT folds (2;5;18) (Figure 3.1 bottom).

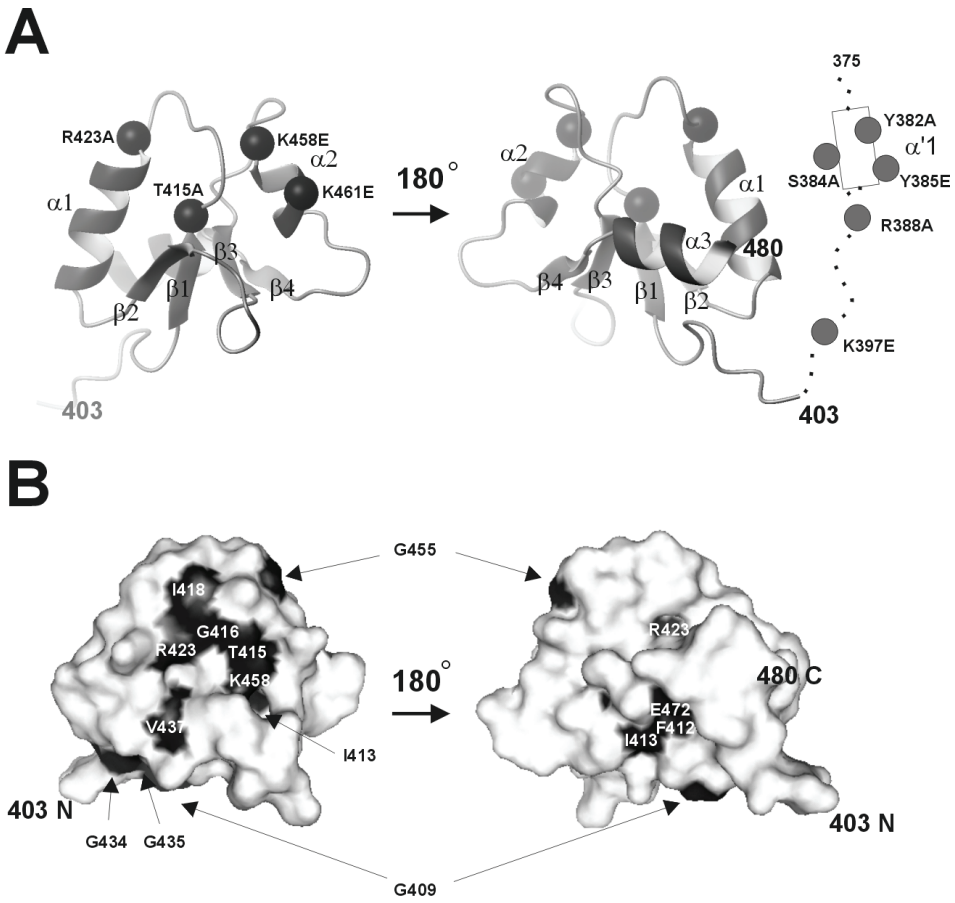


(Figure 3.1) Amino acid sequence alignment of the BRCT region of the large subunit of eukaryotic RFC. Amino acid sequences were retrieved from GenBank under the accession codes to the left of the sequence. The first raw contains the amino acid sequence of the BRCT region from human RFC p140 9RFC1_HUMAN). The alignment was generated by ClustalW (16). Residue identity above 90 % is shaded black and residue similarity above 90% is shaded grey. Analysis of the alignment by AMAS identified the conserved residues (AMAS threshold above 7), which were presented in uppercase letters at the bottom of the sequence. The lowercase letter represents an amino acid identity above 50 %. The last row indicates the secondary structures defined by the NMR data of the human p140 (375-480) (box for alpha helix and arrow for beta strands) based on secondary structure determined by the program CSI 1.0 using chemical shifts of Cα, Cβ and Hα as input (Chapter 4 supplementary materials Figure 1S).

A homology model of the RFC BRCT domain (403-480) (Figure 3.2) was generated using 3D-JIGSAW using the available structure of the BRCT domain from the bacterial NAD^+ dependent DNA ligase, which has an amino acid sequence identity of 33 % (Figure 4.3), (19). The resulting model consists of four parallel β-strands (β1-β4) forming the core, which are surrounded by three α-helices (α1–α3) (Figure 3.2). The N-

terminal sequence between residues 375 and 402 is not included in the model due to the absence of sequence homology, hence the loop (387-402) and the predicted helix (379-386) are schematically presented as a dotted line and a rectangle (α1') respectively (Figure 3.2A right). Most of the conserved hydrophobic residues are involved in stabilizing the core created by residues within the parallel β-strands and the α-helices packed against it (3-5). The structural importance of these conserved hydrophobic residues in the BRCT domain has been demonstrated by reduced stability of the folded protein when mutated (6). Mapping of the conserved residues onto the model structure reveals an interesting pattern where the majority of the conserved residues that are solvent-accessible cluster on either side of the molecular surfaces (Figure 3.2B).

Based on the sequence alignment and the structural model, amino acid residues identified as conserved or with an identity of at least 50 % were selected for mutagenesis. Conserved hydrophobic residues contributing purely to the stability of folding were excluded from mutagenesis. The surface exposed GG sequence (434, 435) (Figure 3.2B), highly conserved throughout the BRCT family, forms a tight turn between α1 and β2, which can be disrupted by substitution with bulky residues (Figure 2.2). Therefore this sequence was also excluded from mutagenesis. The assumption was made that most residues involved in DNA binding would do so via salt bridge or hydrogen bond interactions. Accordingly we selected basic and polar residues for mutagenesis. The final list includes three residues (Y382, S384, Y385) from the predicted α1' helix, two residues within the connecting loop (R388, K397), and four residues (T415, R423, K458 and K461) within the BRCT domain itself (Figure 3.1 and 2 A). These residues were preferably substituted with the negatively charged glutamic acid, if the substitution was not expected to disturb the secondary structure predicted by PSIpred (20). If the X to E substitution appeared to have deleterious structural consequences the residue X was substituted by an alanine. Therefore, four residues, Y385, K397, K458 and K461 were substituted with glutamic acid, and five residues, Y382, S384, R388, T415 and R423 were substituted with alanine. Mutated residues in this study are indicated as blue balls on the structure model in Figure 3.2A.

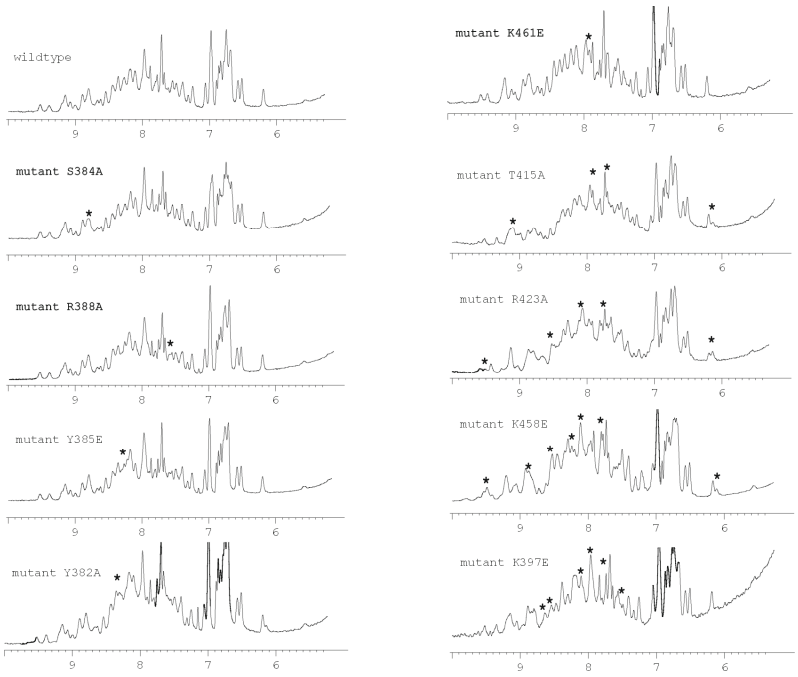


(Figure 3.2) Homology model of the BRCT domain from RFC p140 (403-480).
(A). Ribbon presentation of the model BRCT domain. The model was based on the NMR structure of the BRCT domain (PDB code:1lB7) from the NAD^+ dependent DNA ligase which has 33 % amino acids identity with RFC p140 and was generated by 3D-JIGSAW (19). The residues where structural homology was not available are schematically indicated in the figure. The predicted helix between residues 379 and 386 is presented as a box and the loop (387-402) is presented as a dotted line. The blue balls represent the Cα locations of the point mutations created in this study. **(B)**. Surface presentation of the model BRCT domain. The blue-shaded areas represent the accessible surface occupied by the conserved amino acids defined from the amino acid sequence alignment (Figure 3.1) using AMAS (17). The total accessible surface area (ASA) were calculated using MOLMOL (21).

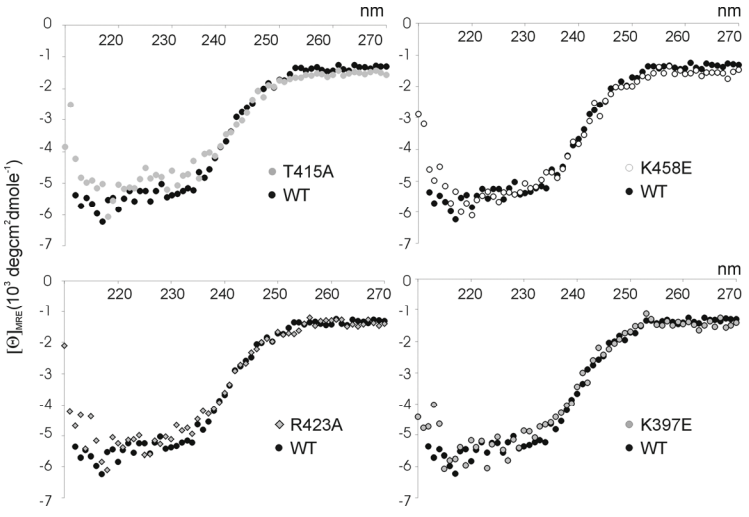
Mutagenesis and purification of proteins

Mutagenesis was performed using the whole plasmid PCR (WHOP) (see Materials and Methods) or the splicing overlap extension (SOE) protocols (14). The resulting mutants were sequenced in the expression vector and the proteins were expressed in the BL21 (DE3) strain of *E. coli*. All nine mutants contained a 6 histidine tag at the C-terminus and were purified using the identical protocol as for the wildtype protein (this thesis Chapter 2). All mutants eluted at the same position as the wt protein during gel

filtration (superpose 12) indicating that the mutants have a similar radius of gyration as the wildtype. Presumably this similarity is due to preservation of the wildtype fold and multimeric state. Proteins were analyzed using SDS-PAGE to determine purity and quantified by UV-absorbance (Materials and Methods). The existence of natively folded protein in the purified fraction was further confirmed by 1D ^{1}H NMR spectroscopy comparing the spectral region containing resonances from the backbone amide ^{1}H's (6 ~ 10 ppm) and aliphatic ^{1}H's (0 ~ 4 ppm). In Figure 3 the spectrum of each mutant protein is presented in order of decreasing similarity to the wt protein based on the number of missing and shifted peaks identified by visual inspection of each spectrum. Asterisks in the mutant spectra indicate the resonances missing or shifted from the ^{1}H amide backbone region of the wt spectrum (Figure 3.3). Mutants (T415A and R423A) exhibit increased linewidth in comparison to the wt, indicating possible aggregation in solution or structural dynamics displayed by the mutants, which are different from that of wildtype. There were no significant differences in the resonances arising from ^{1}H aliphatic carbons (results not shown). The mutants (T415, R423, K397E and K458E) that displayed spectra with some deviations from the wt protein were further checked using CD spectroscopy for secondary structure content (Figure 3.4). Due to the quality of the CD spectra, it is not possible to deduce the exact content of secondary structures in the proteins. However there are no noticeable major difference between the CD spectra of the wildtype and the mutants (Figure 3.4). The overall similarity of the ^{1}H NMR spectra and the qualitative measure indicating existence of secondary structures in the CD spectra suggested that all the mutants were folded with only minor structural changes.



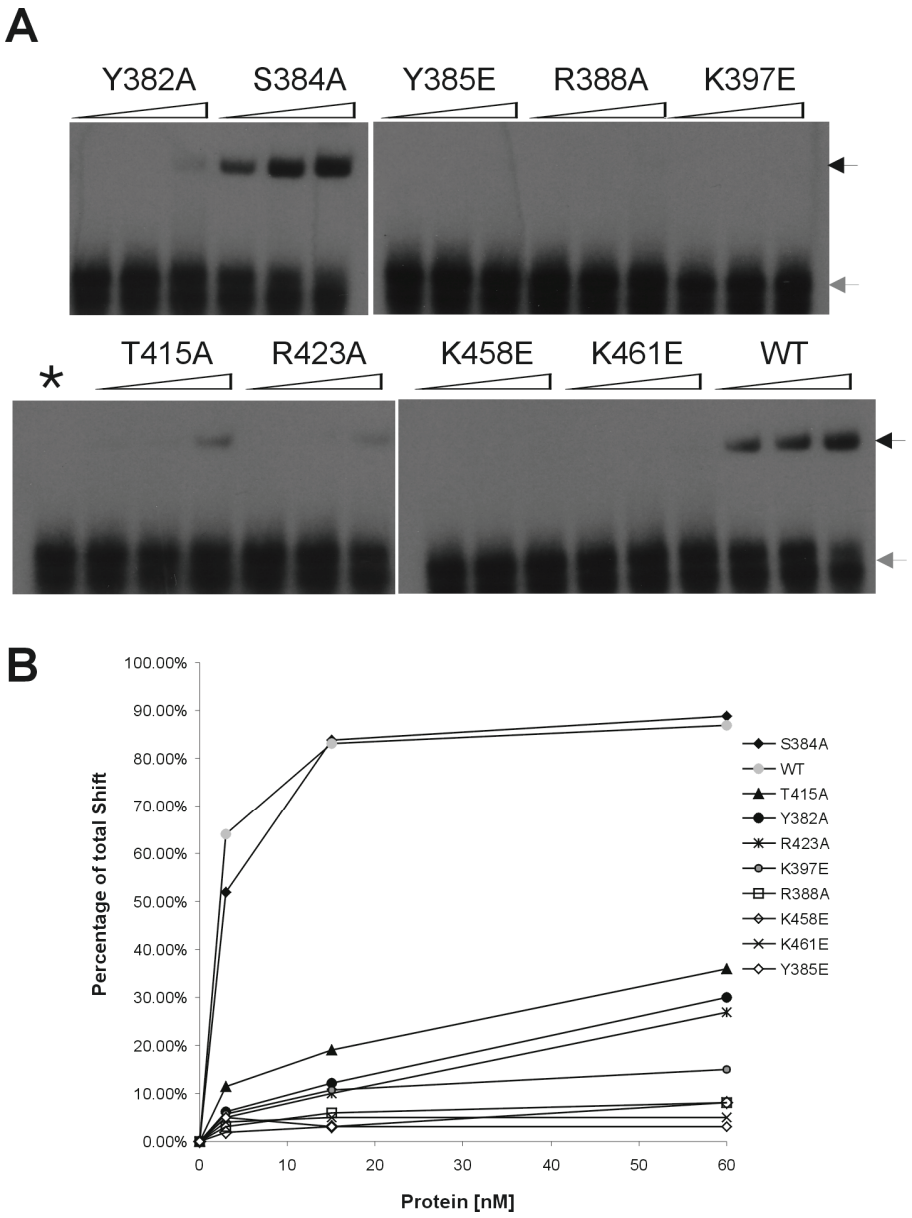
(Figure 3.3) 1D ^{1}H-NMR spectra of the purified p140(375-480) mutants**.** 1D ^{1}H-NMR spectra were recorded of each purified mutant protein described in the text. The good dispersion of resonances in the spectra region > 6.5 ppm, representing primarily amide protons, is indicative of a folded protein. The similarity of each spectrum to that of the the wt protein (the top left lane) indicates a nearly native fold. Resonances that either disappear or are different from the wt spectrum are indicated with asterisks. The spectra of the mutants (indicated on top left corner of the each spectra) are ordered in decreasing similarity to the wt protein (descending order of spectra from the top left to the bottom left and then from top right to the bottom right).



(Figure 3.4) CD spectra of the mutants T415A, R423A, K458E and K397E in comparison to that of wildtype (WT) protein. No significant differences in the secondary structure content between the mutants and wildtype can be seen.

Effect of mutations on DNA-protein complex formation

The effect of each mutation on dsDNA binding was assessed using the gel shift assay. The previously determined binding constant (K_D) of p140 (375-480) for 5' phosphorylated, dsDNA is approximately 10 nM. Under identical conditions, a constant amount of 5' [^{32}P]-labeled ds oligonucleotide (3 nM) was titrated with an increasing amount of each mutant protein (Figure 3.5 A)(3 nM, 16 nM and 66 nM). The data was analyzed with a phosphor-imager and the fraction of DNA bound protein was calculated (Materials and Methods) for each mutant (Figure 3.5B). Only the S384A mutant exhibited DNA binding equivalent to wt (Figure 3.5B). The mutations, Y385E, R388A, K458E and K461E, essentially abrogated DNA binding activity (Figure 3.5). Significant reductions in DNA binding were also observed for the mutations Y382A, T415A and R423A (Figure 3.5). As might be expected, the substitution of residues with the negatively charged glutamic acid sidechain affected DNA binding more drastically than substitution by alanine (Figure 3.5B). This observation likely reflects the greater change in surface electrostatic potential caused by incorporation of a charged residue than a neutral residue.



(Figure 3.5) **The effect of various point mutations on the DNA binding activity of RFC p140 (375-480). (A)**Titration of DNA-protein complex was performed by adding an increasing amount of each protein (3, 15, 60 nM) to a constant amount (3 nM) of the [^{32}P] labeled hairpin oligonucleotide. The resulting DNA-protein complex (black arrow) was separated from the free DNA (grey arrow) in a gel retardation assay. The amount of the complex formed by wildtype (WT) and mutants proteins (labeled on top of the lanes) were compared. The lanes with an asterisk contained only the labeled oligonucleotides as a negative control. **(B)** Quantification of DNA-protein complex formation. The intensity of the bands in the A were determined using a Phosphor-imager. The percentage of the total shift (DNA-protein complex, black arrow in **A**) was calculated as described in Materials and Methods.

Discussion

Structural consequences of mutagenesis

We used site-directed mutagenesis in order to delineate a potential DNA binding site within the BRCT region of RFC p140. This region contains 28 residues N-terminal to the consensus BRCT sequence (Figure 2.2), and the NMR data suggests that this N-terminal sequence forms an α-helix (called α1'). Residues targeted for mutagenesis were identified by sequence homology, and the resulting protein was purified to apparent homogeneity. Since structural changes caused by a particular mutation could potentially indirectly affect the activity of the protein, two qualitative measurements of a different nature, ^{1}H-NMR and CD spectroscopy, were employed to inspect structural differences between the wildtype and the mutant proteins. Slight changes in 1D NMR spectra were observed for those mutations within the BRCT domain (T415A, R423A and K458E) and within the loop (K397E), but no major differences in the secondary structure contents were seen between the wildtype and those mutants judging from their CD spectra hence indicating that the structural changes in the mutants are minor.

Effects of Mutagenesis on DNA binding

The sequence alignment of the BRCT region from 31 different species revealed a striking lack of amino acid conservation in the N-terminal sequence of the BRCT region. However, mutation of the weakly conserved aromatic groups Y382 and Y385 within the α1' helix and the positively charged residues R388 and K397 in the adjacent loop resulted in a significant reduction in DNA binding. The interaction site on the α'1 helix was further localized by the observation that the S384A mutation had no effect on DNA binding suggesting that a specific side of the helix makes contacts to the DNA. These results support our earlier observation that deletion of the N-terminal sequence of the BRCT region abrogated DNA binding. Interestingly, both aromatic residues in the α'1 helix and the arginine and lysine in the adjacent loop that were mutated in this study are conserved in Drosophila and human p140.

The absence of conserved residues in the N-terminal sequence of the BRCT region in combination with the presence of a limited number of highly conserved residues within the BRCT domain suggests that 5' phosphate recognition may be predominantly modulated by the BRCT domain. The reduced DNA binding observed for mutations at the

conserved T415, R423 and K458 and weakly conserved K461, which is located close to the first three, indicates a potential surface for DNA binding. Remarkably T415, R423, and K458 are conserved widely throughout the bacterial NAD^+ dependent DNA ligases (Figure 3.6), which together with the RFC p140, form a distinct subclass of the BRCT family (1). A site directed mutagenesis study of the BRCT domain of the NAD^+ dependent DNA ligase from *Thermus* species AK16D (DNLJ_THESK) indicated that the conserved residues including T599, R606 and G638 (T415, R423 and G455 respectively in p140) are important for DNA binding activity associated with the subsequent adenyl-moiety transfer to the 5' phosphate of the dsDNA nick (22). Deletion of the BRCT domain from the NAD^+ dependent DNA ligases of several bacterial species has indicated its crucial role in DNA binding (22-25), and recently DNA binding by the isolated BRCT domain from *E.coli* LigA has been confirmed (25). Hence this particular class of the BRCT family, with the conserved amino acids described above, may form a group of DNA binding BRCTs.

(Figure 3.6) The positional conservation of amino acids involved in phospho-moiety binding by the tandem BRCT domains from BRCA1 and the conserved residues of the distinct subclass of BRCT superfamily. The alignment was prepared using ClustalW (16) and gaps were introduced to the alignment based on the secondary structure of BRCA1 BRCT (PDB entry: 1T29) and human RFC p140. The asterisks indicate the phospho-moiety binding residues by the N-terminal domain of the tandem BRCT repeat from BRCA1. The representatives of the distinct subclass of BRCT superfamily used for the alignment are RFC p140 subunit of RFC1_HUMAN (human) and RFC1_DROME (*Drosophila melanogaster*), and NAD+ dependent ligases from DNLJ_THEFI (*Thermus filiformis*), DNLJ_THESK (*Thermus sp AK16D*), DNLJ_ECOLI (*Escherichia coli*) and DNLJ_CHLTR (*Chlamydia trachomatis*). BRCA1_HUMAN sequence is from the N-terminal BRCT of the tandem BRCTs. The secondary structure of the BRCT domain from RFC p140 as determined by NMR (this thesis) is indicated above the alignment. Arrows indicate β-sheet and square indicate α-helix.

Potential 5' phosphate binding residues

The tandem BRCT domains from BRCA1 bind a short peptide (ISRSTpSPTFNKQT) from the protein BACH1 (BRCA1 associated Carboxyl-terminal Helicase) in a phosphoserine dependent manner (7-9). The crystal structure of the complex formed between the BRCA1 tandem BRCT domains and the phosphoSerine peptide revealed that specific interaction is achieved via a network of hydrogen bonds formed by T1655, G1656 and K1702 to the phosphate oxygens. These residues belong to the first domain of the tandem BRCT repeat (10-12) and are also conserved among other tandem BRCT domains with phospho-peptide binding activity (8). Interestingly, despite the

overall lack of sequence homology between the BRCT domains of BRCA1 and RFC p140, sequence alignment based on the secondary structure elements shows a positional conservation of three key residues of BRCA1 BRCTs that are involved in phospho-moiety recognition. These residues are equivalent to the conserved T415, G416 and K458 on the surface of p140 (375-480) (Figure 3.6). The importance of T415 and K458 for DNA binding is shown here. A possible conclusion is that these residues form a patch on the surface of RFC p140 that is involved in 5' phosphate binding. Structure elucidation of p140 (375-480) will likely confirm this conclusion.

It has to be noted that the DNA binding is not restricted to the distinct class of the BRCT superfamily formed by RFC p140 and NAD^+ dependent DNA ligases. A recent report suggests that polymerase X-family members TdT, pol mu and pol lambda, all of which bear an N-terminal BRCT domain, have DNA binding activity, albeit with weak apparent affinity (26). The BRCT domains from these polymerases have little amino acid similarity to that of RFC p140 (375-480). Similarly the BRCT domain from mouse Rev1 has been shown to bind DNA stably enough to be observed on a gel retardation assay (unpublished data provided by G. Siegal). It will be interesting to see if DNA binding by the BRCTs from polX family or Rev1 has similar biochemical properties to the BRCT region of p140(375-480).

Although we could not identify which residues were directly interacting with the 5' phosphate, conserved residues which directly influence DNA binding were identified. The data presented here could potentially be useful for modeling the DNA-protein complex formed by the BRCT domain.

Reference list

1. Bork, P., Hofmann, K., Bucher, P., Neuwald, A. F., Altschul, S. F., and Koonin, E. V. (1997) *FASEB J.* **11,** 68-76

2. Gaiser, O. J., Ball, L. J., Schmieder, P., Leitner, D., Strauss, H., Wahl, M., Kuhne, R., Oschkinat, H., and Heinemann, U. (2004) *Biochemistry* **43,** 15983-15995

3. Krishnan, V. V., Thornton, K. H., Thelen, M. P., and Cosman, M. (2001) *Biochemistry* **40,** 13158-13166

4. Williams, R. S., Green, R., and Glover, J. N. (2001) *Nat.Struct.Biol.* **8,** 838-842

5. Zhang, X., Morera, S., Bates, P. A., Whitehead, P. C., Coffer, A. I., Hainbucher, K., Nash, R. A., Sternberg, M. J., Lindahl, T., and Freemont, P. S. (1998) *EMBO J.* **17,** 6404-6411

6. Dulic, A., Bates, P. A., Zhang, X., Martin, S. R., Freemont, P. S., Lindahl, T., and Barnes, D. E. (2001) *Biochemistry* **40,** 5906-5913

7. Manke, I. A., Lowery, D. M., Nguyen, A., and Yaffe, M. B. (2003) *Science* **302,** 636-639

8. Rodriguez, M., Yu, X., Chen, J., and Songyang, Z. (2003) *J.Biol.Chem.* **278,** 52914-52918

9. Yu, X., Chini, C. C., He, M., Mer, G., and Chen, J. (2003) *Science* **302,** 639-642

10. Clapperton, J. A., Manke, I. A., Lowery, D. M., Ho, T., Haire, L. F., Yaffe, M. B., and Smerdon, S. J. (2004) *Nat.Struct.Mol.Biol.* **11,** 512-518

11. Shiozaki, E. N., Gu, L., Yan, N., and Shi, Y. (2004) *Mol.Cell* **14,** 405-412

12. Williams, R. S., Lee, M. S., Hau, D. D., and Glover, J. N. (2004) *Nat.Struct.Mol.Biol.* **11,** 519-525

13. Allen, B. L., Uhlmann, F., Gaur, L. K., Mulder, B. A., Posey, K. L., Jones, L. B., and Hardin, S. H. (1998) *Nucleic Acids Res.* **26,** 3877-3882

14. Horton, R. M., Cai, Z. L., Ho, S. N., and Pease, L. R. (1990) *Biotechniques* **8,** 528-535

15. Higgins, D. G., Thompson, J. D., and Gibson, T. J. (1996) *Methods Enzymol.* **266,** 383-402

16. Combet, C., Blanchet, C., Geourjon, C., and Deleage, G. (2000) *Trends Biochem.Sci.* **25,** 147-150

17. Livingstone, C. D. and Barton, G. J. (1993) *Comput.Appl.Biosci.* **9,** 745-756

18. Thornton, K. H., Krishnan, V. V., West, M. G., Popham, J., Ramirez, M., Thelen, M. P., and Cosman, M. (2001) *Protein Expr.Purif.* **21,** 401-411

19. Bates, P. A., Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. (2001) *Proteins* **Suppl 5,** 39-46

20. Jones, D. T. (1999) *J.Mol.Biol.* **292,** 195-202

21. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J.Mol.Graph.* **14,** 51-32

22. Feng, H., Parker, J. M., Lu, J., and Cao, W. (2004) *Biochemistry* **43,** 12648-12659

23. Jeon, H. J., Shin, H. J., Choi, J. J., Hoe, H. S., Kim, H. K., Suh, S. W., and Kwon, S. T. (2004) *FEMS Microbiol.Lett.* **237,** 111-118

24. Lim, J. H., Choi, J., Kim, W., Ahn, B. Y., and Han, Y. S. (2001) *Arch.Biochem.Biophys.* **388,** 253-260

25. Wilkinson, A., Smith, A., Bullard, D., Lavesa-Curto, M., Sayer, H., Bonner, A., Hemmings, A., and Bowater, R. (2005) *Biochim.Biophys.Acta* **1749,** 113-122

26. Ma, Y., Lu, H., Tippin, B., Goodman, M. F., Shimazaki, N., Koiwai, O., Hsieh, C. L., Schwarz, K., and Lieber, M. R. (2004) *Mol.Cell* **16,** 701-713