

# Conformational Effects of a Common Codon 399 Polymorphism on the BRCT1 Domain of the XRCC1 Protein

Regina Monaco · Ramon Rosal · Michael A. Dolan ·  
Matthew R. Pincus · Paul W. Brandt-Rauf

Published online: 25 September 2007  
© Springer Science+Business Media, LLC 2007

**Abstract** The X-ray cross-complementing-1 (XRCC1) protein functions as a scaffold that coordinates the activity of the cellular machinery involved in base excision repair (BER) of DNA damage. The BRCT1 domain of XRCC1 is responsible for interacting with several of the key components of the BER machinery, and it is also the site of a common genetic polymorphism in XRCC1 at amino acid residue 399 (Arg → Gln). Experimental and epidemiologic evidence suggest that this polymorphism may alter BER capacity and increase cancer risk. The aim of this study was to investigate whether these effects could be attributable to conformational changes in XRCC1 induced by the polymorphism. Molecular dynamics techniques were used to predict the structure of the wild-type and polymorphic forms of the BRCT1 domain of XRCC1, and differences in structure produced by the polymorphic substitution were determined. The results indicate that, although the general configuration of both proteins is similar and there is little actual deviation at the site of the polymorphism itself, the substitution produces significant conformational changes at

several other sites in the BRCT1 domain, including the loss of secondary structural features such as  $\alpha$  helices that may be critical for protein–protein interactions. These results provide support for the hypothesis that this polymorphism in XRCC1 could affect DNA repair capability by altering the structure of the BRCT1 domain and thus the ability of XRCC1 to coordinate BER.

**Keywords** Genetic polymorphism · Structure–function correlation · Base excision repair · Molecular dynamics

## Abbreviations

XRCC1	X-ray cross-complementing-1
BER	Base excision repair
BRCT1	BRCA1 carboxy terminus
PARP	Poly(ADP-ribose) polymerase
hOGG1	human 8-oxoguanine glycosylase
MPG	Methyl purine glycosylase
APE1	Apurinic/aprimidinic endonuclease 1
Pol $\beta$	DNA polymerase $\beta$
Lig III	DNA ligase III
NTD	N-terminal domain

## 1 Introduction

The XRCC1 (X-ray cross-complementing-1) protein is a key component of the cellular machinery responsible for base excision repair (BER) of DNA damage [1]. XRCC1 acts as a scaffold protein in this repair pathway coordinating the activity of the other components of the BER machinery [1, 2]. Thus, XRCC1 appears to regulate the function of the poly (ADP-ribose) polymerases that catalyze the ribosylation of a

---

R. Monaco · R. Rosal · P. W. Brandt-Rauf (✉)  
Department of Environmental Health Sciences, Mailman School  
of Public Health, Columbia University, 60 Haven Avenue, B-1,  
New York, NY 10032, USA  
e-mail: pwb1@columbia.edu

R. Rosal  
Chemical Terrorism Laboratory, New York City Department  
of Health and Mental Hygiene, New York, NY 10016, USA

M. A. Dolan  
Tripos, Inc, St. Louis, MO 63144, USA

M. R. Pincus  
Department of Pathology and Laboratory Medicine, New York  
Harbor VA Medical Center, Brooklyn, NY 12209, USA

number of DNA-bound proteins to decrease their affinity for DNA allowing the repair machinery access to the damaged site (PARP-1 and PARP-2), the specific glycosylases that recognize and remove the damaged base (including hOGG1, human 8-oxoguanine glycosylase, and MPG, methyl purine glycosylase), an apurinic/apyrimidinic endonuclease responsible for cleaving the phosphodiester bond at the abasic site created by the glycosylase (APE1), a DNA polymerase that has deoxyribosephosphodiesterase activity to release the 5' sugar phosphate group and gap-filling synthesis activity to add one nucleotide to the 3'-OH (Pol  $\beta$ ), and a DNA ligase that seals the nick in an ATP-dependent fashion (Lig III) [1, 2]. XRCC1 coordinates the activity of these other proteins through three protein interaction modules that constitute independent and distinct globular domains of the protein, including an N-terminal domain (NTD), a central BRCA1 carboxy terminal domain (BRCT1), and a C-terminal BRCT domain (BRCT2) [3]. The two intervening linker regions between these globular domains have been denominated hinges and have been predicted to be unstructured [3].

The NTD is the site of interaction with Pol  $\beta$  [4], and BRCT2 is the site of interaction with Lig III [5]. The BRCT1 domain (approximately residues 315–407) is believed to be responsible for XRCC1 interaction with PARPs and the BER glycosylase, decreasing the activity of the former and increasing the activity of the latter, and thus is a critical domain for proper functioning of the repair process [6, 7]. Similar BRCT domains have been found in a number of other important cancer-related proteins, such as BRCA1, BRCA2, 53BP1, REV1, RAP1, RAD9, RAD4, Ect2, and Crb2, where they are believed to likewise be involved in critical protein–protein interactions [8]. The BRCT1 domain of XRCC1 is also the site of a very common single nucleotide polymorphism that results in the substitution of a glutamine (Gln) for the normally occurring arginine (Arg) at amino acid residue 399 [9]. Although the data to date have been somewhat inconsistent, some evidence from both experimental and epidemiologic studies suggests that this polymorphism in XRCC1 may alter DNA repair capabilities and cancer risk [10]. Thus, it is theoretically possible that this polymorphism is responsible for altering the structure of the BRCT1 domain and hence the function of the XRCC1 protein by disrupting critical protein–protein interactions and its coordination of BER. The purpose of the present study was to investigate this possibility by determining if there are differences in the structure of the BRCT1 domain with Arg and Gln at position 399 using molecular dynamics techniques.

## 2 Materials and Methods

The BRCT2 domain of XRCC1 (amino acid residues 538–633) is highly homologous to the BRCT1 domain (68%

homologous; 20% identical), and its X-ray crystallographic structure has been determined [11]. Therefore, we used this known BRCT2 structure as the starting point for determining the structure of the wild-type BRCT1 domain with Arg at amino acid residue 399 via an adaptation of a molecular dynamics approach as previously described for other protein structural determinations, which has been shown to yield results that are consistent with experimental data [12, 13].

First, the sequence of the BRCT1 domain from amino acid residues 315–407 with Arg at 399 was threaded onto the coordinates of the BRCT2 domain using commercially available software (Sybyl; Tripos, St. Louis, MO), with the ends of the peptide neutrally blocked. This alignment was done by direct substitution at homologous amino acid residues that left two gaps due to the fact that the BRCT2 domain contains three additional amino acid residues (at 553 and 598–599) that have no analogous residues in the BRCT1 sequence; however, these additional residues occur in random coil loop regions that connect  $\beta$  sheets and  $\alpha$  helices and thus can be easily deleted without expecting to produce any significant conformational effects. The resulting gaps were closed by making direct peptide bonds between the adjacent amino acid residues in the BRCT1 sequence. Next each BRCT1 substituted residue in the BRCT2 structure was relaxed and minimized using Sybyl's Structure Preparation Tool, which allowed steric clashes to be removed. The BRCT1 domain has an overall charge of +9 that was neutralized by the random placement of Cl<sup>-</sup> counterions. These counterions were allowed to relax individually to correspond to physico-chemically reasonable placement. The molecule was then immersed in a water box of 7,507 water molecules, and the water molecules were also allowed to relax while the BRCT1 domain was held restrained.

A series of nested energy minimizations were then performed on this complex of the BRCT1 molecule, counterions and water, resulting in an overall minimized structure for the wild-type BRCT1 domain. After the overall structure was minimized to a root-mean-square (RMS) gradient of 0.001 Å, the dynamics runs were carried out. The BRCT1 complex was heated to 300°K over 2 picos and then allowed to equilibrate over the balance of the dynamics run. The energies, volume and density of the completed dynamics run were examined to insure that the results were physically reasonable. The final average wild-type BRCT1 structure was determined from the last 25–50 picos of the run and was calculated from the frames along the part of the dynamics stimulation that had equilibrated.

Next, the final average wild-type BRCT1 structure was used as the starting point for determining the structure of the polymorphic BRCT1 domain with Gln substituted for the normally occurring Arg at amino acid residue 399. Then the same procedures as above were repeated to obtain

the final average polymorphic BRCT1 structure. Finally, the average structure for the polymorphic protein was superimposed on that for the wild-type protein such that the RMS deviation of the coordinates of the backbone atoms of one structure from the other was a minimum. The average RMS deviation between the polymorphic and wild-type proteins was determined for the structures as a whole, as well as for each amino acid residue individually to identify isolated regions with the most significant conformational changes.

### 3 Results

As shown in Figs. 1 and 2, coordinate fluctuations for all residues in both the wild-type and polymorphic forms of the protein were less than 1 Å (in most cases much less), confirming that the proteins converge on a specific structure in each case.

The overall RMS deviation for the average structures of the wild-type and polymorphic proteins was moderately large at 4.95 Å, even though the general configuration of the wild-type BRCT1 domain is retained in the polymorphic form. The individual residue RMS deviations for the average structures are shown in Fig. 3, demonstrating even larger deviations (>5 Å at multiple residues) at discrete regions throughout the BRCT1 domain. The most prominent of these deviations occur at Leu320-Gly322, Ser328-Gln331, Arg335-Ser336, Arg339-Ala342, Asp352-Thr354, Trp385-Val386, Arg391-Leu395, and Met403-Gly407. Interestingly, at the site of the polymorphism at residue 399, although there is some loss of secondary structure in the polymorphic form (see below), there is minimal deviation (1.94 Å) between the two structures; however, there

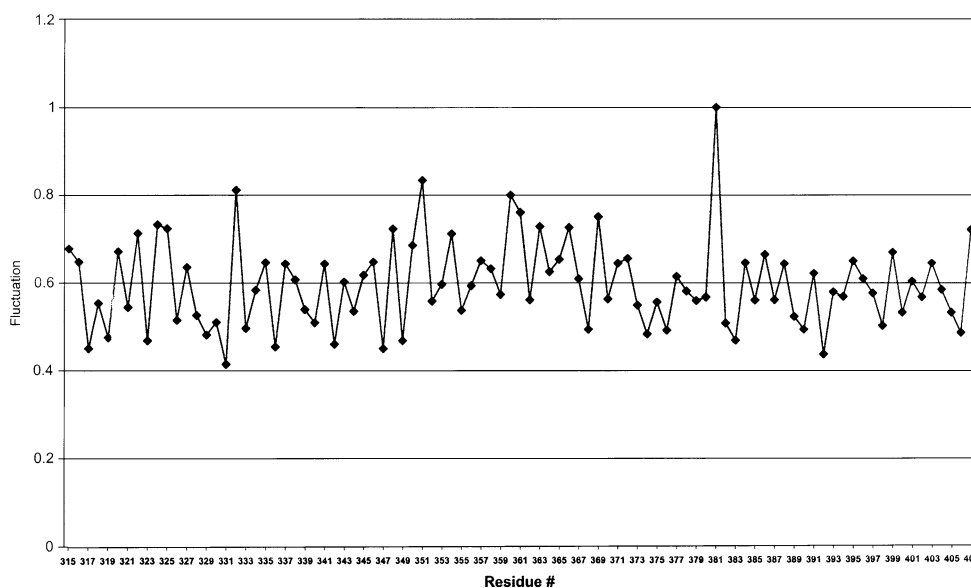
are regions of considerable deviation N-terminal and C-terminal to this site.

Perhaps even more significant are the differences in secondary structure between the wild-type and polymorphic proteins, which can be seen in Figs. 4–6. Figure 4 shows the average backbone structure for the wild-type protein, Fig. 5 shows the average backbone structure for the polymorphic protein, and Fig. 6 shows the best-fit superposition of the two average structures. The wild-type protein can be seen to have four  $\alpha$  helices and three  $\beta$  sheets (Fig. 4). However, in the polymorphic protein two extended  $\alpha$  helices from Arg335-Leu345 and from Trp385-Arg391 and a shorter  $\alpha$  helix from Ser398-Arg400 are no longer present; similarly, the  $\beta$  sheets from Val323-Leu327 and from Ala347-Tyr349 are also missing in the polymorphic protein (Fig. 5). In particular, the  $\alpha$  helices, which in many cases are known to be involved in protein interactions, occur on the surface of the BRCT1 domain where they could easily participate in such interactions, so these changes could obviously have significant functional implications.

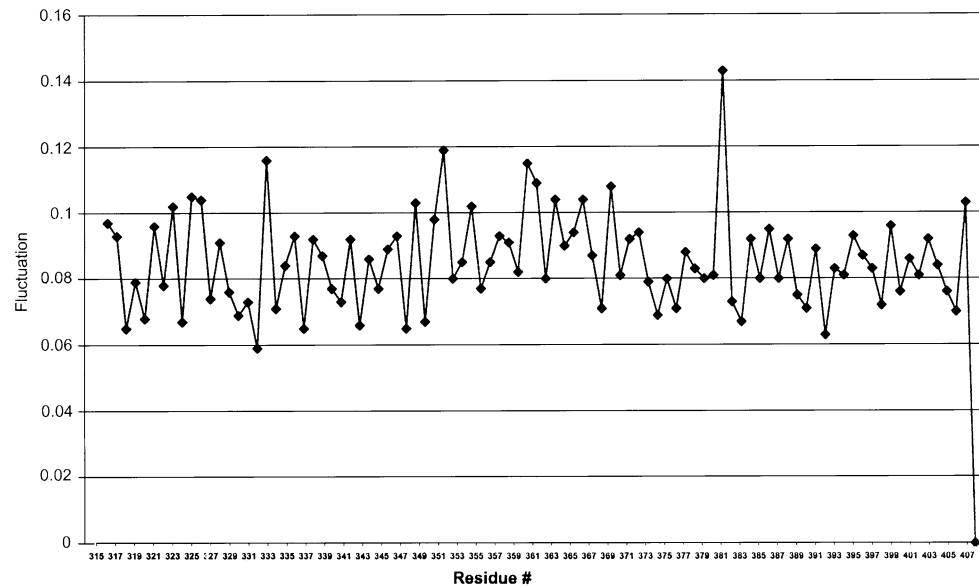
### 4 Discussion

These results suggest that the polymorphic substitution of Gln for Arg at amino acid residue 399 in the XRCC1 protein could produce significant conformational changes in the BRCT1 domain. Although the site of the substitution itself does not show much deviation in the two structures, the substitution appears to be responsible for major deviations and secondary structural changes at several other sites in the BRCT1 domain, both adjacent to the site of the substitution, as well as at a distance from it.

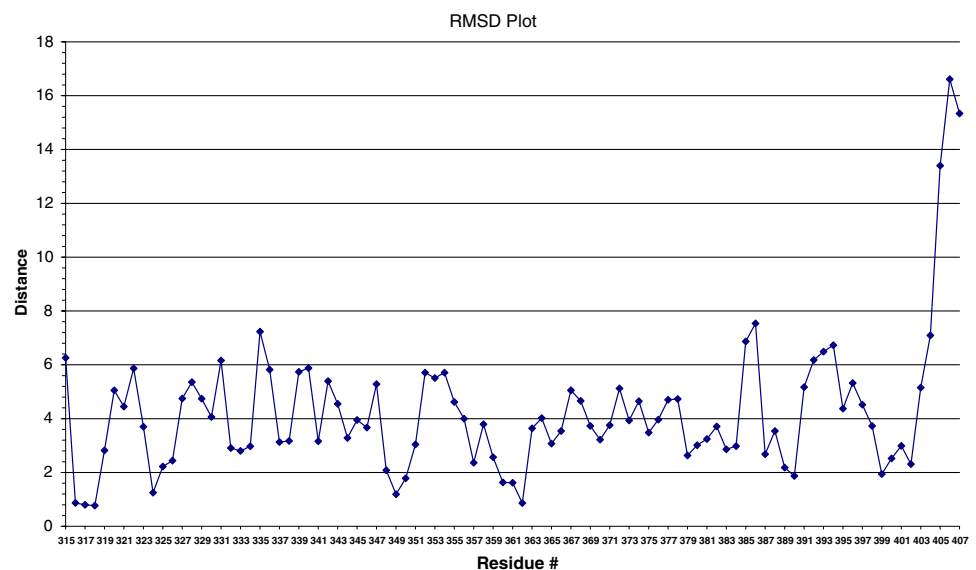
**Fig. 1** Coordinate fluctuations for the final low-energy structures on the dynamics trajectory of the wild-type BRCT1 domain of XRCC1. Scale is in Angstroms



**Fig. 2** Coordinate fluctuations for the final low-energy structures on the dynamics trajectory of the polymorphic BRCT1 domain of XRCC1. Scale is in Angstroms



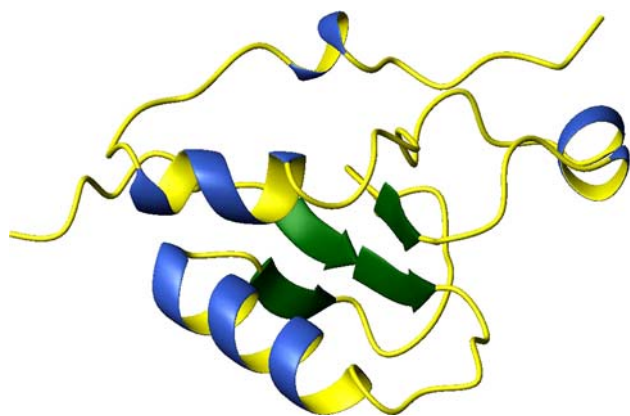
**Fig. 3** Individual residue backbone deviations for the coordinates of corresponding amino acid residues 315–407 of the average structure of the polymorphic BRCT1 domain from those of the wild-type domain average structure. Scale is in Angstroms



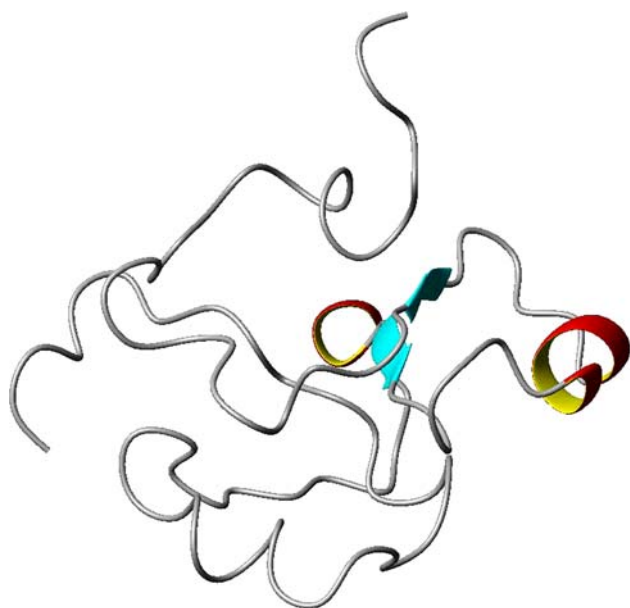
The conformational effects noted in this study seem to be consistent with other lines of evidence. For example, it is known that DNA-PK, a member of the phosphatidylinositol 3-kinase-related kinase superfamily, phosphorylates the BRCT1 domain of XRCC1 at Ser371 following exposure to ionizing radiation, causing XRCC1 dimer dissociation [14]. The presence of the polymorphic Gln399 variant does not affect this phosphorylation [14], suggesting that this site is not significantly changed in conformation from that of the wild-type protein. Ser371 is in the center of the one  $\alpha$  helix that occurs in both the wild-type and polymorphic forms of the protein from Lys369-Gln372, retaining its relative orientation with respect to the rest of the BRCT1 domain. Furthermore, our results show that the RMS deviation at Ser371 between the wild-type

and polymorphic proteins is only a moderate 3.76 Å and that the site remains exposed on the surface of the BRCT1 domain easily accessible for phosphorylation in both forms (Fig. 6).

Data from our own prior studies suggest that the Gln399 polymorphism leads to a reduction in BER capability both in vitro and in vivo [15, 16]. For example, we have studied a model population of workers exposed to the known mutagen/carcinogen vinyl chloride and the effect of this XRCC1 polymorphism on the occurrence of biomarkers of mutations in this cohort. Vinyl chloride is known to be metabolized to the reactive intermediates chloroethylene oxide and chloroacetaldehyde which form promutagenic etheno-DNA adducts. The resultant etheno-adenine DNA adduct is believed to be responsible for the production of

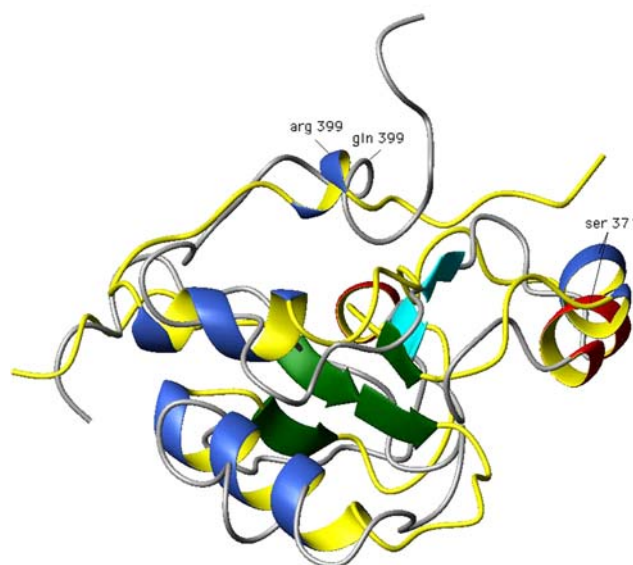


**Fig. 4** C $\alpha$  tracing of the average structure of the wild-type BRCT1 domain of XRCC1 from molecular dynamics simulations with  $\alpha$  helices indicated by broad blue bands and  $\beta$  sheets by broad green arrows. The N-terminus is to the lower left and the C-terminus to the upper right



**Fig. 5** C $\alpha$  tracing of the average structure of the polymorphic BRCT1 domain of XRCC1 from molecular dynamics simulations with  $\alpha$  helices indicated by broad red bands and  $\beta$  sheets by broad blue arrows. The N-terminus is to the lower left and the C-terminus to the upper right

A  $\rightarrow$  T transversions in the *TP53* tumor suppressor gene that occur in workers with the sentinel neoplasm for vinyl chloride exposure, angiosarcoma of the liver, and the biomarkers for these mutations that occur in exposed workers without tumors [17]. These mutant p53 biomarkers occur in a statistically significant dose-response relationship with regard to cumulative vinyl chloride exposure, but at any given exposure level there are seemingly otherwise similar individuals who differ in the occurrence of the biomarkers, suggesting that some genetically determined susceptibility



**Fig. 6** Superposition of the C $\alpha$  tracings of the average structures of the wild-type (yellow) and polymorphic (gray) forms of the BRCT1 domain of XRCC1. The site of phosphorylation at Ser371 and of the polymorphic substitution at Arg/Gln399 are indicated. The N-termini are to the lower left and the C-termini to the upper right

could account for different outcomes despite similar exposures. The presence of the Gln399 polymorphism in XRCC1 appears to explain much of this differential susceptibility. For instance, vinyl chloride-exposed workers who were homozygous for the Gln399 polymorphism were found to have a 4-fold increased risk for the occurrence of the mutant p53 biomarkers even after controlling for potential confounders including cumulative vinyl chloride exposure [15, 16]. In addition, we studied the effect in cell culture of chloroacetaldehyde exposure on the formation of etheno-adenine adducts in lymphoblast lines from individuals who were homozygous wild-type or homozygous variant for the XRCC1 399 polymorphism. The efficiency of repair of the adducts in the homozygous wild-type cells was four times greater than the efficiency of repair in the homozygous variant cells, directly analogous to the epidemiologic results in the exposed workers [16]. These findings are consistent with the fact that the repair of the etheno-adenine adducts produced by vinyl chloride occurs by BER via MPG and the other XRCC1-coordinated repair machinery. As noted, an important step in this process requires XRCC1 regulation of PARP-1, which depends on catalyzing its automodification with multiple poly(ADP-ribose) molecules [18]. The BRCT1 domain of XRCC1 contains the PAR-binding consensus sequence from Arg379 to Arg400 [18]. As found in the present study, this is a region of significant RMS deviation and secondary structural change in the polymorphic protein, including the loss of a major  $\alpha$  helical segment in the middle of the

sequence from Trp385 to Arg391 on the surface of the BRCT1 domain. Therefore, it is entirely plausible that the effects on BER of etheno-adenine adducts that we have observed in vitro and in vivo are at least partially attributable to the conformational alterations in the PAR-binding site, disrupting PARP-1 regulation and the access of the repair machinery to the site of the DNA damage. Similarly, it is possible that the conformational changes noted in other regions of the BRCT1 domain affect the interaction of XRCC1 with different BER proteins, including the glycosylases. For example, some evidence suggests that the XRCC1-hOGG1 interaction involves the BRCT1 domain and the hinge region just N-terminal to it [19], so the conformational changes that we have found in the Gln399 polymorphic protein at the N-terminal end of BRCT1, including the loss of the  $\alpha$  helix from Arg335 to Leu345, could conceivably affect the ability of XRCC1 to interact with and regulate hOGG1 activity.

In summary, the results of the present study provide support for the hypothesis that the position 399 polymorphism in XRCC1 could influence BER capability by altering the structure of the BRCT1 domain and disrupting its interaction with other components of the repair machinery. Since this is a relatively common polymorphism in many populations [10], these conformational alterations and their functional effects could account for a significant amount of the observed variability in DNA repair and hence cancer risk in these populations.

**Acknowledgments** This work was supported in part by grants from NIEHS (P30-ES09089) and NIOSH (R01-OH04192).

## References

- Caldecott KW (2003) *DNA Repair* 2:955–969
- Vidal AE, Boiteux S, Hickson ID, Radicella JP (2001) *EMBO J* 20:6530–6539
- Callebaut I, Labesse G, Durand P, Poupon A, Canard L, Chomilier J, Henrissat B, Moron JP (1997) *Cell Mol Life Sci* 53:621–645
- Marintchev A, Mullen MA, Maciejewski MW, Pan B, Gryk MR, Mullen GP (1999) *Nat Struct Biol* 6:884–893
- Nash RA, Caldecott KW, Barnes DE, Lindahl T (1997) *Biochemistry* 36:5207–5211
- Schreiber V, Ame JC, Dolle P, Schultz I, Rinaldi B, Fraulob V, Menissier-de Murcia J, de Murcia G (2002) *J Biol Chem* 277:23028–23036
- Campalans A, Marsin S, Nakabeppu Y, O'Connor TR, Boiteux S, Radicella JP (2005) *DNA Repair* 4:826–835
- Glover JNM, Williams RS, Lee MS (2004) *Trends Biochem Sci* 29:579–585
- Shen MR, Jones IM, Mohrenweiser H (1998) *Cancer Res* 58:604–608
- Goode EL, Ulrich CM, Potter JD (2002) *Cancer Epidemiol Biomarkers Prev* 11:1513–1530
- Zhang X, Morera S, Bates PA, Whitehead PC, Coffey AI, Hainbucher K, Nash RA, Sternberg MJE, Lindahl T, Freemont PS (1998) *EMBO J* 17:6404–6411
- Brandt-Rauf PW, Chen JM, Marion MJ, Smith SJ, Luo JC, Carney WP, Pincus MR (1996) *J Protein Chem* 15:367–375
- Chen JM, Rosal R, Smith S, Pincus MR, Brandt-Rauf PW (2001) *J Protein Chem* 20:101–105
- Levy N, Martz A, Bresson A, Spenlehauer C, de Murcia G, Menissier-de Murcia J (2006) *Nucleic Acids Res* 34:32–41
- Li Y, Marion MJ, Rundle A, Brandt-Rauf PW (2003) *Biomarkers* 8:408–414
- Li Y, Marion MJ, Zipprich J, Freyer G, Santella RM, Kanki C, Brandt-Rauf PW (2006) *Int J Occup Med Environ Health* 19:45–52
- Smith SJ, Li Y, Whitley R, Marion MJ, Partilo S, Carney WP, Brandt-Rauf PW (1998) *Am J Epidemiol* 147:302–308
- Pleschke JM, Kleczkowska HE, Strohm M, Althaus FR (2000) *J Biol Chem* 275:40974–40980
- Marsin S, Vidal AE, Sossou M, Menissier-de Murcia J, LePage F, Boiteux S, de Murcia G, Radicella JP (2003) *J Biol Chem* 278:44068–44074