Detection of protein folding defects caused by BRCA1-BRCT truncation and missense mutations

R. Scott Williams¹, Daniel Chasman², David Hau¹, Benjamin Hui¹, Al Lau³, and J.N. Mark Glover*‡

¹ 437 Medical Sciences Building
Department of Biochemistry
University of Alberta
Edmonton, Alberta
Canada, T6E 426
*correspondance:
mark.glover@ualberta.ca

² Division of Preventative Medicine
Brigham and Women’s Hospital
3rd Floor 900 Commonwealth Avenue
Boston MA, 02215, USA

³ Variagenics/Nuvelo, Inc.,
60 Hampshire Street,
Cambridge, MA 02139, USA

Running Title: BRCT folding mutations
SUMMARY:

Most cancer-associated BRCA1 mutations identified to date result in the premature translational termination of the protein, highlighting a crucial role for the C-terminal, BRCT repeat region in mediating BRCA1 tumour suppressor function. However, the molecular and genetic effects of missense mutations that map to the BRCT region remain largely unknown. Using a protease-based assay, we directly assessed the sensitivity of the folding of the BRCT domain to an extensive set of truncation and single amino acid substitutions derived from breast cancer screening programs. The protein can tolerate truncations of up to 8 amino acids but further deletion results in drastic BRCT folding defects. This molecular phenotype can be correlated with an increased susceptibility to disease. A cross-validated computational assessment of the BRCT mutation database suggests that as much as half of all BRCT missense mutations contribute to BRCA1 loss of function and disease through protein destabilizing effects. The coupled use of proteolytic methods and computational predictive methods to detect mutant BRCA1 conformations at the protein level will augment the efficacy of current BRCA1 screening protocols, especially in the absence of clinical data that can be used to discriminate deleterious BRCT missense mutations from benign polymorphisms.
INTRODUCTION

Germline mutations within the breast and ovarian cancer susceptibility gene *BRCA1* predispose carriers to early-onset breast and breast-ovarian cancers (1). Accumulating evidence points to a role for the *BRCA1* protein product in the regulation of multiple nuclear functions including transcription, recombination, DNA repair, and checkpoint control (2-4). Tumour-associated mutations occur throughout the *BRCA1* coding sequence, but cluster to sequences encoding the N-terminal RING finger domain and the two carboxy-terminal repeat BRCT domains (5-7).

The molecular details of how BRCA1 mutations contribute to the pathogenesis of cancer remain largely unknown. The functional significance of the BRCT region is highlighted by the high degree of sequence conservation within the BRCT regions of among mammalian, Xenopus, and avian BRCA1 homologues (8-10). Several lines of evidence reveal the BRCT is required for tumour suppressor function. A nonsense mutation, which removes 11 C-terminal residues of the second, BRCT (Tyr1853 → stop), is associated with early-onset breast cancer (11). Two cancer-linked BRCT missense mutations (12) that destabilize the BRCT fold (13-15), A1708E and M1775R, ablate the double-strand break repair and transcription function of BRCA1 (16) and inhibit BRCT interactions with histone deacetylases (17), BACH1 (18) and the transcriptional co-repressor CtIP (19, 20). Furthermore, mice with homozygous targeted mutations removing the C-terminal half of BRCA1 are viable but develop...
tumours, suggesting the missing BRCT and/or other domains are expendable for survival, but not for tumour suppression (21).

While all frameshift or nonsense mutations recorded in the Breast cancer Information Core (BIC) resulting in BRCA1 protein truncation are viewed as functionally deleterious (22, 7), the physiological significance of the majority of missense variants has not been determined due to the absence of a distinctive functional assay for BRCA1. More than 70 missense substitutions have been recorded that alter the primary sequence of the tandem BRCT-repeats but pedigree analysis clarifying the disease linkage of these alleles is available for only eight of these variants (6, 7, 12, 23-27). Many of these amino acid substitutions may be linked with disease but remain as unclassified in the BIC because the presence of the allele has not been tested in the general population, or the segregation of the allele with disease within a family is unclear (6, 7).

The recent determination of the x-ray crystal structures of the rat and human BRCA1 BRCT repeat domains were important first steps towards understanding tumourigenic BRCT mutations and provide a novel platform for the interpretation of the effects of these alterations in the absence of clinical data (15, 28). In the present study we directly evaluate the consequences missense mutation on the structure of the human BRCA1 BRCT repeats. Using a proteolysis-based assay to probe the BRCT for non-native conformations, we show that the majority of the tested missense and truncations alter the folding state of the BRCT. Cross-validated computational analyses using the BRCT structure and the sequences of proteins homologous to the human BRCT from
other organisms further suggest that many of the unclassified BRCT missense mutations are likely to be disease predisposing and perturb BRCA1 structure/function through protein destabilizing effects.
EXPERIMENTAL PROCEDURES:

Mutagenesis and vector construction

Coding sequences for BRCT C-terminal truncations of human BRCA1 were amplified from the T7 promoter based expression vector for pLM1-BRCA1 (1646-1863) (13) using the following oligonucleotides: (fragment 1646-1859) FT7-5'-gga cga gaa ttc tta acc agg gag ctg att atg gtg aac aaa aga atg tcc atg-3', CD6-5'-gat ctg gga tcc tca ggg gat ctg ggg tat cag-3'; (fragment 1646-1858) FT7, CD7-5'-gat ctg gga tcc tca gat ctg ggg tat cag gta-3'; (fragment 1646-1857) FT7, CD1-5'-gat ctg gga tcc tca ctg ggg tat cag gta ggt-3'; (fragment 1646-1855) FT7, CD5-5'-gat ctg gga tcc tca tat cag gta ggt gtc cag-3'; (fragment 1646-1853) FT7, CD4-5'-gat ctg gga tcc tca gta ggt gtc cag ctc ctg-3'; (fragment 1646-1852) FT7, 1853Ystop-5'-gat ctg gga tcc tca ggt gtc cag ctc ctg gca-3'; (fragment 1646-1851) FT7, CD3-5'-gat ctg gga tcc tca gtt gtc cag ctc ctg gca ctg-3'; (fragment 1646-1849) FT7, CD2-5'-gat ctg gga tcc tca ctc ctg gca ctg gag-3'; (fragment 1646-1829) FT7, 1829stop -5'-gat ctg gga tcc tca aca cat ctg ccc aat tgt gcc aag ggt-3'. The 5' primer FT7 incorporates a ribosome binding site and an EcoRI site for cloning. 3' oligonucleotides include the relevant stop codons and a BamHI restriction site. Gel purified PCR products were digested with EcoRI and BamHI and cloned into BamHI-EcoRI digested pLM1 (29).
All BRCT single amino acid substitutions were introduced into the BRCT fragment 1646-1859. For missense mutations A1708E, M1775R, and W1837R, mutated BRCA1 coding sequences were used as template for PCR amplification with the FT7 and CD6 primers. All other missense substitutions were engineered using PCR splicing methods (30). Primary PCR mutagenesis reactions used oligo FT7 with the appropriate reverse (R) mutagenesis oligonucleotide (see below) and CD6 with the appropriate forward (F) mutagenesis oligonucleotide. PCR products from the primary reactions were gel purified from 1.5% agarose gels using a QIAEX2 kit (Qiagen) and mixed together with oligonucleotides FT7 and CD6 in the secondary PCR splicing reactions to generate mutated PCR products that were subsequently digested with EcoRI/BamHI and ligated to pLM1. The mutagenesis oligos used were: D1692Y, F - 5’ - gtt atg aaa aca tat gct gag ttt tgt - 3’, R - 5’ - cac aaa ctc agc ata tgt ttt cat aac - 3’; F1695L, F - 5’ - aca gat gct gag ctt tgt gaa cg - 3’, R - 5’ - ccg ttc aca cac aag ctc agc atc tgt - 3’; V1696L, F - 5’ - gat gct gag ttt tgt gtg gaa cgg caa - 3’, R - 5’ - tgt ccg ttc aca caa aca ctc agc atc - 3’; C1697R, F-5’-gct gag ttt tgt gtg cgt gaa cgg aca ctg-3’, R-5’-cag tgt ccg ttc acg cac aaa ctc agc - 3’; R1699W, F - 5’ - tgt gtt tgt gaa tgt cat gga caa ctg aat - 3’, S1715R, F-5’-aaa tgg gta gtt aga tat tgt cca ttc aca cac aac - 3’; W1718C, F-5’-gtt agc tat ttc tgt gtg acc cac tct - 3’, R-5’-aga tgt ggt cac aca gaa ata gct aac - 3’; T1720A, F - 5’ - tat ttc tgt gtg gac cgc cac tct att aaa - 3’, R - 5’ - tgt aat aga ctg ggc cac cca gaa ata - 3’; G1738E, F-5’-ttt gaa
gtc aga gaa gat gtg gtc aat g – 3’, R-5’- cat tga cca cat ctt ctc tga ctt caa a-3’;
G1738R F-5’- ttt gaa gtc aga aga gat gtg gtc aat g – 3’, R-5’- cat tga cca cat ctc
ttc tga ctt caa a-3’;  P1749R, F-5’- aac cac caa ggt cgt aag cga aga g – 3’, R-
5’- ctc ttg ctc gct tac gac ctt ggt gtt t –3’;  R1751Q, F - 5’ - caa ggt cca aag caa
gca aga gaa tcc – 3’, R - 5’ - gga ttc tct tgc ttg ctt acc ttg - 3’;  A1752P, F - 5’ -
ggt cca aag cga cca aga gaa tcc cag – 3’, R - 5’ - ctg gga ttc tct tgg ctt ctt acc - 3’;  I1766S, F-5’- agg ggg cta gaa agc tgt tgc tat ggg – 3’, R-5’- ccc ata gca aca
gct ttc tag ccc cct-3’;  M1783T F-5’- caa ctt gaa tgg acc gta cag ctt tgt g - 3’, R-
5’- cac aca gct gta cgg tcc att cca ggt t;  G1788V, F - 5’ - gta cag ctt tgt gtt gct tct
gtg gtg – 3’, R - 5’ - cac cac aga agc aac aca cag ctt tac - 3’;  V1804D, F - 5’ - ctt
ggc aca ggt gac cac cca att gtg – 3’, R - 5’ - cac aat tgg gtg gtc acc tgt gcc aag - 3’;  V1809F, F - 5’ - cac cca att gtt ttt tgt cag cca gat – 3’, R - 5’ - atc tgt cag cac
aaa cac aat tgg gtg - 3’;  W1837G, F -5’- gtg acc cga gag ggg gtg ttg gac agt g –
3’, R-5’- cac tgt cca aca ccc cct ctc ggg tca c –3’.

All vectors were sequenced to confirm the success of the mutagenesis reactions.

**Proteolysis assays**

0.2 to 0.5 µg of pLM1 plasmid encoding the BRCT variants were used
directly as template for protein synthesis reactions with the TNT-Quick *in vitro*
transcription/translation system (Promega). Immediately prior to proteolytic
digestion, proteins were translated and labelled with ³⁵S-methionine at 30 °C for
2 hr. The reticulocyte lysates were then centrifuged for 2 min. at 10000 x g to
remove insoluble material and 3 µL of the lysate supernatants containing the
labelled translation products were added to 12 µL digestion buffer (150 mM NaCl, 50 mM potassium phosphate, pH 7.5) containing increasing concentrations of trypsin (Sigma) or Tlck-treated chymotrypsin (Sigma). After digestion at 20 °C for 12 minutes, the reactions were stopped with PMSF. Digestion products were electrophoresed on 15% SDS-PAGE gels and visualized with a phosphorimaging plate and a Molecular Dynamics Typhoon scanner. A local average background correction was used during quantification of the reaction products with ImageQuant (Amersham-Molecular Dynamics).

**Molecular graphics**

Structural diagrams were created with Bobscript (31, 32) and rendered using Povray (www.povray.org).

**Computational analysis of risks associated with missense mutations**

**Method 1:** Structure and sequence-based analysis. A probability of an effect on function for the missense mutations in BRCT was determined exactly as described using both feature set A and feature set B in reference (33). Briefly, the crystal structure (PDB ID: 1JNX), the multiple sequence alignment for proteins homologous to human BRCT (see Figure 3), and the chemical nature of the amino acid substitution are used to compute the values of features that are useful for predicting the effects of amino acid substitutions on protein function. For example, the quantitative estimate of solvent accessibility for a residue in a structure or its normalized phylogenetic entropy from a multiple sequence
alignment are both features that can be viewed as having a quantitative relationship to the probability of an effect on function for the introduction of a mutant amino acid (33). A probability of an effect on function for a test mutation is estimated by conditional probability as the fraction of training mutations derived from exhaustive mutagenesis of the Lac repressor (34) and T4 lysozyme (35) with an effect on function from among those with feature values are similar to the feature values of the test mutation.

**Method 2:** Refined sequence-based analysis. The sequence-based procedure (Lau and Chasman, submitted) for predicting the functional consequences of a mutation in a residue of human BRCT is an extension of the direct inspection of alternative amino acid existing at the corresponding residue in proteins homologous to the human BRCT. In essence, mutations that introduce amino acids observed at the corresponding residue of homologous proteins are judged to be tolerated by the human BRCT. The extension involves inferring, through the use of the Blocks9 mixture of Dirichlet priors (36), the more likely of two hypotheses explaining why some amino acids are not observed at corresponding residues of homologous proteins. Either the sequences of the homologous proteins represent an incomplete sampling of all twenty amino acids at a mutated residue position or some of the twenty amino acids are incompatible with the structural and functional constraints on the mutated residue position. Inclusion of the alternative amino acid from a mutation in the inferred set of acceptable amino
acids is evidence for its compatibility with biological function. Exclusion leads to a prediction of incompatibility.
RESULTS:

Structural effects of BRCT truncation mutations

We previously demonstrated that the tandem BRCT repeat region of human BRCA1 forms a proteolytically resistant globular domain and that a cancer-linked mutation, Y1853ter, which removes the 11 C-terminal residues of the protein, reduces this proteolytic stability (13). To determine to what extent the BRCT fold could tolerate truncation mutations, we subjected a series BRCT deletion mutants to a proteolytic sensitivity assay (Fig. 1, Experimental procedures). The oncogenic mutation Y1853ter and all larger C-terminal deletions of the protein were degraded by the lowest concentrations of trypsin whereas the full length BRCT (aa 1646-1863) is highly resistant to cleavage (Fig. 1). Included with these mutations are the truncation protein products of two of the most common BIC frameshift mutants, 5382insC and IVS21-36del510, that result in stop codons at positions 1829 and 1805 of the BRCA1 coding sequence (7) (Figs. 2, 3). Thus, BRCT folding defects resulting from cancer predisposing BRCA1 truncation mutations can be assayed for and detected at the protein level using a simple protease sensitivity assay.

The deletion experiment also demonstrates the protein can tolerate removal of up to 8 residues, but further deletion from the C-terminus greatly impairs the native folding of the domain, rendering it highly sensitive to proteolysis (Fig. 1, Fig. 2d). Consistent with this finding, the transcriptional activation activity of the BRCT domains was abolished by C-terminal deletions that truncate beyond a hydrophobic pair of residues, L1854 and I1855 (24).
These hydrophobes mark the C-terminal boundary for conservation of mammalian, avian and Xenopus BRCA1 homologues (Fig. 3), and make critical aliphatic contacts to the β-sheet of the C-terminal BRCT in the structures of the human and rat BRCA1-BRCT repeats (Fig. 2d) (13, 28). Hence, the transcriptional activation defects observed for BRCT deletion mutants likely result from destabilization of the protein.

**Missense substitutions destabilize the BRCT**

Similar to the truncation mutants, two cancer predisposing missense mutations, A1708E and M1775R, are destabilizing and exhibit altered BRCT protease susceptibility (13-15). To gain insights into the effects of other patient-derived mutations recorded in the BIC, we generated 23 additional missense variants and tested these proteins for proteolytic sensitivity (see Experimental Procedures, Fig. 4). 20/25 of the missense mutations tested showed varying degrees of enhanced sensitivity to tryptic digestion at 20 °C (Fig. 4a). Five of six of the mutations that substitute an arginine into the protein (C1697R, S1715R, G1738R, P1749R and W1837R) also show increased sensitivity to chymotryptic cleavage at 20 °C (Fig. 4b) suggesting that destabilizing effects, rather than the introduction of a new trypsin cleavage site, are responsible for the protease sensitivity. Mutant M1775R is also clearly destabilizing (14) and shows sensitivity to chymotrypsin at elevated temperatures (15).

The expression levels of the BRCT variants in the reticulocyte lysates typically range between 0.3-1.2 fold of wild type levels. Since the expressed variants constitute less than 5-10% of the total protein digested in the lysates and
we are using logarithmic increases in trypsin concentrations, we can quantify the percent protein remaining following digestion at each level of protease, and directly compare these values to establish a proteolysis-based hierarchy for the severity of the destabilizing effects (Fig. 5). Here we define highly destabilizing mutations as those mutants for which >60% of the protein is degraded at the lowest concentration (6 µg/mL) of trypsin. Intermediately destabilizing variants are >60% degraded at the intermediate trypsin concentration (60 µg/mL). Finally, the mutants showing wild type digestion profiles, with limited degradation until exposure to the highest trypsin concentration, are classified as having no destabilizing effect. Based on these criteria, the majority of the variants (13/25) are highly destabilizing, 7/25 are intermediately destabilizing and 5/25 have no apparent effect.

Homology models of the human BRCA1-BRCT repeats, built from the XRCC1 carboxyl terminal BRCT structure (37), have been used to describe structural environments of BRCA1-BRCT missense variants (38). Since these descriptions are inaccurate in many respects, we have reclassified the BRCT missense mutations into the four following categories based on their distribution in the human BRCT repeat structure (13) (Fig. 5, Table 1):

i. **Surface mutations:** This class of mutations includes amino acid residues found on the surface of the BRCT that appear to make little contribution to the structure of the domain. Three of five of the tested surface mutants (F1695L, T1720A, and V1804D) have no destabilizing effect. Two exposed mutations localized to the β3-α2 connecting loop (D1692Y and V1696L) confer moderate
protease susceptibility to the domain (Fig. 5a). As highlighted by Joo et al. (28), this loop forms a extended β-hairpin structure, participates in the formation of one of the two conserved BRCT surfaces targeted by missense mutations, and may be the primary site of interaction with the BRCA1-associated helicase BACH1. Hence, the BACH1 binding defect reported for mutation F1695L likely results from disruption of a contact site, whereas reduced BACH1 binding for the V1696L mutation may be due to a destabilizing effect, disruption of the contact site, or both.

**ii. BRCT-interface mutations:** The BRCA1 C-terminal domain consists of two BRCT repeats that pack together head-to-tail via a conserved triple-helical interface and several of the key residues mediating these BRCT-BRCT contacts are targeted by mutation (Fig. 5b, Table 1) (13). Four of five of the tested BRCT-interface mutants are destabilizing. Three destabilizing mutants, A1708E, M1775R and M1783T, likely disrupt the hydrophobic packing between the repeats. The crystal structure of the M1775R variant revealed that mutation-induced structural rearrangements, including flipping of the mutated arginine out of the hydrophobic core of the protein, contribute to fold destabilization (15). The intermediate protease sensitivity of mutation M1783T likely results from combined deleterious effects of protein core cavitation and the burial of a polar hydroxyl group at the BRCT interface. Residue R1699 normally participates in a salt bridge between the BRCT repeats. The loss of salt-bridging interactions and steric strain associated with accommodating the tryptophan may contribute to conformational instability of the R1699W mutant. The intermediate stability of
R1699W explains the temperature-sensitive transcription phenotype ascribed to this mutation (27). Conversely, R1699Q has little to no effect on BRCT structure and appears to have little effect on transcription activation (25).

iii. BRCA1-fold mutations: We have designated a third class of mutations as BRCA1-fold mutants. These substitutions include residues that participate in folding of the BRCT linker region, and residues that do not fall at BRCT-fold special positions, but are buried and conserved amongst BRCA1 homologues (Fig. 5c). All of the tested BRCA1-fold class mutants tested alter the folding of the domain. The majority of these mutants (C1697R, S1715R, G1738R, G1738E, and P1749R) introduce charged residues into the protein core and are highly destabilizing. The A1752P mutant likely disrupts the linker helix and is highly destabilizing. The position of R1751 in the crystal structure of human BRCA1 is unclear, but the equivalent residue in the rat structure indicates this residue is involved in salt bridging interactions and the packing of BRCT linker helix (28). This arginine is conserved amongst all known BRCA1 homologues and the R1751Q mutation may disrupt similar electrostatic stabilization in the human protein.

iv. BRCT-fold mutations: A conserved hydrophobic clustering signature for the BRCT fold superfamily of proteins was originally identified using sequence based methods (39, 40). Residues at these positions dictate the fold of an individual BRCT and participate in formation of the BRCT hydrophobic core or are found in turns. As shown in Fig 5d, 7/8 of the mutations tested (W1718C, I1766S, G1788V, V1809F, W1837G, W1837R, and Y1853C) at BRCT-fold positions are
intermediately or highly destabilizing. The highly destabilizing mutations W1718C, W1837G and W1837R mutate the invariant BRCT-fold tryptophan of both the amino- and carboxyl-terminal BRCT domains. This critical core residue appears intolerant to both cavitating (W1718C and W1837G) or charge substitution mutations (W1837R) and mediates van der Walls contacts from helix $\alpha_3$ to other secondary structure elements of the BRCT fold including $\beta$-sheet, helix $\alpha_1$ and the $3_{10}$ helix. Mutation G1788V disrupts the conserved tight turn between $\alpha_1'$ and $\beta2'$ of the C-terminal BRCT. Figures 2d and 5d highlight the role of Y1853 in positioning the C-terminal BRCT $3_{10}$ helix that packs against the $\beta$-sheet, and substitution of this residue with a cysteine is highly destabilizing. Two of the three tested BRCT-fold class mutants that target residues that contribute to intra $\beta$-sheet packing, I1766S and V1809F, are destabilizing. The third $\beta$-sheet mutant, M1652I, has been classified as a benign polymorphism (41, 42) and does not increase the protease sensitivity of the domain.

_Mutations that destabilize the BRCT predispose carriers to disease_

Pedigree analysis clarifying the disease predisposition of _BRCA1_ alleles is currently available for 8 of the 79 reported BRCA1-BRCT single amino acid substitution variants (Table 1). Seven of the mutations, D1692Y, C1697R, R1699W, A1708E, S1715R, P1749R and M1775R, are destabilizing (Figures 4,5) and are linked to cancer (Table 1). In contrast, the frequently recorded BIC polymorphism M1652I exhibits no structural defect, indicating the protease stability assay can successfully discriminate benign mutations from disease-
causing variants. The cancer-associated truncation mutants are also protease sensitive (Fig. 2).

Taken together, these results indicate that protease based detection of altered BRCT stability provides a novel and powerful predictive tool that can be used to assess disease linkage of BRCT mutations in instances where pedigree data is not available. Thus, we suggest that the 20 destabilizing missense mutants and truncations greater than eight amino acids are cancer-predisposing.

**Predicting the structural consequences of mutation on the BRCT**

The recent development of computational methods that incorporate detailed structural and sequence information to predict the effects of single amino acid substitutions on protein structure/function provides us with alternative tools to study the BRCT mutations (33, Chasman and Lau - submitted). We have applied two independent methods to predict the potential effects of the 25 missense mutations studied here (Table 1), and all known BRCT single amino acid substitutions recorded in the breast cancer information core (www.nhgri.nih.gov/Intramural_research/Lab_transfer/Bic/) (Supplementary Table 1).

In the first computational method, a set of quantitative and qualitative features are defined for each amino acid substitution in the BRCT based on its structural disposition, an assessment of the sequence conservation mapped onto the structure, and the potential consequences of introducing the alternative amino acid into the structural environment of the mutated residue. An overall probability of an effect on function is then calculated by comparing the values of
the features to the values of features of a large number of mutations with known effects on function (in this case from the Lac repressor and T4 lysozyme (33). Predictions using two sets of features (A and B in ref. 33) yielded similar results (see Supplementary Table 1). About 51% (35 out of 69) of the mutations are predicted to effect BRCT structure or function with a probability greater than 0.5. Comparison to the experimental analysis of the mutations by proteolysis indicated disagreement for six of 22 of the mutations (0.27), consistent with previous estimates of the cross-validated error described previously (Table 1) (33).

The second computational methodology estimates potential effects of mutations on function by extracting as much information as possible from the observed amino acid substitutions in the comparison of the human BRCT sequence to its homologues from other organisms. The method (Chasman and Lau, submitted) relies extensively on Bayesian prior information representing empirically observed amino acid exchangeability in a large number of sequence families (BLOCKS database, ref. 36; Blocks9 priors, ref. 43) to infer whether an alternative amino acid introduced by mutation is consistent with the biological and chemical character of amino acids found at the corresponding residues of BRCT homologues. This method predicts 14 of 25 of the mutations tested by proteolysis are incompatible with normal BRCT function, representing agreement with the experimental data for 19 of them (0.76). The first method combining structure and sequence is consistent with predictions from method 2 for 18/24
mutations (0.75), and both methods agree with the proteolysis data for 14/24 (0.58) of the mutations.
**Discussion:**

*Protein destabilization ablates BRCT mediated transcriptional activation*

When tethered to a GAL4 DNA binding domain, the BRCT domains can activate transcription in yeast and mammalian systems (44-46). Significantly, potential targets of BRCA1 transcriptional regulation include the p53-responsive genes encoding p21 as well as GADD45 (47, 48) suggesting that BRCA1 has a role in regulating DNA repair and checkpoint controls. The BRCT may modulate these functions through direct recruitment of the RNA polymerase holoenzyme (49, 50), however, the physiological significance of these effects and the precise biochemical mechanism by which the BRCT activates transcription remains unclear (reviewed in 2). Nevertheless, this intrinsic activity forms the basis for a BRCA1 functional assay that has been used to probe for defects caused by several BRCT missense mutations (24, 25, 27).

Comparison of the transcription and protease based assay data reveal a striking correlation between destabilizing phenotypes and transcriptional defects (Table 1). That is, less stable BRCT variants including C1697R, R1699W, A1708E, S1715R, G1738E, and M1775R, as well as the truncation mutants, disrupt transactivation function, whereas mutations with no effect on structure (M1652I and R1699Q) are fully active in these assays. However, it has yet to be determined whether BRCT protein misfolding causes BRCA1 tumour suppressor inactivation via BRCA1 transcription function, DNA repair function, or both.
The Protease based assay for ranking BRCT destabilizing effects

Proteolytic degradation proceeds via an unfolded state for small globular proteins (51, 52) indicating that a correlation between proteolytic resistance and the thermodynamic stability of a protein may exist. This principle forms the basis for phage-based proteolytic selection methods where the evolution of proteins with increased thermodynamic stability closely follows the selection of polypeptides with enhanced resistance to degradation by increasing concentration of protease (53, 54). Thus, the application of a protease-based assay to assess the structural consequences of missense mutations on the BRCT provides a quick, effective, complimentary method to categorize and rank the extent of destabilization of the mutant BRCT proteins.

A recent biophysical assessment of the effects of 8 missense substitutions and the truncation Y1853ter on the thermodynamic stability of the BRCT revealed that four of these missense mutations and the truncation were highly destabilizing and could not be produced as soluble protein in *E. coli* (14). All four of these missense mutants, A1708E, G1738E, G1788V and W1837R and the truncation show extreme sensitivity to tryptic digestion (Table 1, Fig. 4, Fig. 5). The remaining four (M1775R, M1783T, V1808A, and V1833M) can be produced recombinantly, but destabilize the protein by 3.5–5.5 kcal/mol. Two of these, M1775R and M1783T, show an intermediate sensitivity to proteolysis. Six of the seven other BRCT mutant proteins (V1696L, R1699W, R1751Q, M1783T, V1809F, and Y1853C) with intermediate protease sensitivity are soluble in *E. coli* when expressed at 20 °C (Table 1). Altogether, these data indicate a
three-tiered hierarchy of destabilizing effects inferred from the proteolytic data is consistent with results obtained from solubility analysis and direct thermodynamic measurements of BRCT protein stability. Highly destabilizing mutations show sensitivity to low levels of trypsin and tend to be degraded or insoluble when expressed in *E. coli*. Intermediate thermodynamically destabilizing mutations are sensitive to moderate levels of protease and can be produced in soluble form in *E. coli*.

The remaining set of mutations may not affect the folding detected by the proteolysis assay and yet still affect the functional properties of human BRCT. The computational methods we have explored represent a first attempt to identify alternative correlates within this class of disease predisposing substitutions. The computational methods were largely consistent with the proteolysis data, whether or not there was an effect on protein stability. The purely sequence-based computational methodology was more consistent with the experimental evidence than the structure- and sequence-based approach. Whether the discrepancy can be interpreted or not remains to be seen through further studies. Methods for predicting the biological consequences of amino acid substitutions is an area of active research, especially since the genome initiatives are discovering too large a number of amino acid altering genetic variants with potential effects on biological function for experimental analysis (for example, see 55, 56).
Detection of BRCT mutations

The observation that the BRCA1 BRCT domains form a proteolytically resistant domain, and that cancer predisposing BRCT variants (but not benign single amino acid substitutions) have compromised stability, indicates that a protease-based screen for mutant BRCT conformations could be incorporated into routine BRCA1 screening protocols. The protease-based assay we have used tests the stability of BRCT domains expressed in commercially available reticulocyte lysates. These lysates are employed in protein truncation test (PTT) genetic screens which have been used for detecting BRCA1 mutations (5, 57, 58). In this method, patient-isolated RNA is reverse transcribed to generate a cDNA that is then amplified using oligonucleotides that target a coding region of the protein that is to be tested for protein truncating effects. The cDNA message is then transcribed/translated and radiolabelled within the lysate, and the presence of truncated protein product is visualized by SDS-PAGE. Although effective at detecting frameshifts, nonsense and deletion mutations that lead to truncation of the expressed protein message, this technique is incapable of identifying missense substitutions.

An adaptation of the PTT, where a protease digestion step is added could be appropriate for the detection of the large majority of cancer-associated BRCT mutations (Fig. 6a). Here, oligonucleotides would be specifically designed to amplify BRCT coding sequence (aa 1646-1863) from patient samples, and the translation step would be followed by a trypsinolysis series. This test would have the distinct advantage of sensing the protein destabilizing effects of both
missense and truncation mutations. Conservative estimates indicate it could
detect as much as 80% of the cancer-associated mutations that fall within the
BRCT coding region. Alternatively, for cases where a BRCT missense
mutation has already been detected by sequencing, the mutant BRCT coding
sequence could be produced by PCR (Fig. 6b). Direct transcription/translation
from the PCR product, followed by protein digestion, would provide a quick,
relatively inexpensive test for mutant BRCT conformations.

Conclusions

Greater than 60% of clinically relevant BRCA1 mutations delete a portion
of or all of the BRCT domains, and the majority of BRCT missense alterations
tested which target the 3 key classes of BRCT folding determinants (BRCT-fold,
BRCA1-fold and Interface mutations) are destabilizing. It is apparent that BRCT
destabilization or loss of function through truncation or missense substitution is
sufficient to confer disease predisposition in carriers for these alleles. Such
mutations are comparable to the subset of β-sandwich and zinc binding
mutations that unfold the core DNA binding domain of the p53 tumour suppressor
(reviewed in 59). Conversely, the p53 core domain mutation database is largely
populated by mutations that have little effect on stability, but directly target
residues involved in sequence specific DNA binding. The identification of
analogous cancer-associated mutations that are not destabilizing, but disrupt
specific BRCT protein-protein or protein-DNA binding would provide strong
support for the role of these interactions in mediating BRCA1 tumour suppressor
function. To this end, two patient derived mutations (F1695L and V1696L) on the
surface of the BRCT that affect BACH1 binding have recently been identified (28). Further biochemical and structural characterization of these interactions will be necessary to confirm the nature of these defects.

We have established a set of probability and protease-based criteria on which we can define the structural effects of mutation on the BRCT at the protein level. The early identification of carriers of potentially deleterious BRCA1 alleles is an essential component of breast and ovarian cancer screening programs that facilitates detection, surveillance, and prevention of tumour growth. Further development of complimentary methods that test the destabilizing and biological repercussions of missense variants will provide clinicians and researchers with important tools to unravel BRCA1 function and misfunction.
References:

1. Nathanson, K. L., Wooster, R., Weber, B. L., and Nathanson, K. N. (2001) *Nat Med* 7, 552-556

2. Monteiro, A. N. (2000) *Trends Biochem Sci* 25, 469-474

3. Scully, R. and Livingston, D. M. (2000) *Nature* 408, 429-432

4. Venkitaraman, A. R. (2002) *Cell* 108, 171-82

5. Shattuck-Eidens, D., McClure, M., Simard, J., Labrie, F., Narod, S., Couch, F., Hoskins, K., Weber, B., Castilla, L., Erdos, M. et al. (1995) *JAMA* 273, 535-541

6. Couch, F. J. and Weber, B. L. (1996) *Hum Mutat* 8, 8-18

7. Shen, D. and Vadgama, J. V. (1999) *Oncol Res* 11, 63-69

8. Sharan, S. K., Wims, M., and Bradley, A. (1995) *Hum Mol Genet* 4, 2275-2278

9. Chen, K. S., Shepel, L. A., Haag, J. D., Heil, G. M., and Gould, M. N. (1996) *Carcinogenesis* 17, 1561-1566

10. Joukov, V., Chen, J., Fox, E. A., Green, J. B., and Livingston, D. M. (2001) *Proc Natl Acad Sci U S A* 98, 12078-12083

11. Friedman, L. S., Ostermeyer, E. A., Szabo, C. I., Dowd, P., Lynch, E. D., Rowell, S. E., and King, M. C. (1994) *Nat Genet* 8, 399-404.

12. Miki, Y., Swensen, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L. M., Ding, W. et al. (1994) *Science* 266, 66-71

13. Williams, R. S., Green, R., and Glover, J. N. (2001) *Nat Struct Biol* 8, 838-42

14. Ekblad, C. M., Wilkinson, H. R., Schymkowitz, J. W., Rousseau, F., Freund, S. M., and Itzhaki, L. S. (2002) *J Mol Biol* 320, 431-442

15. Williams, R. S. and Glover, J. N. (2003) *J Biol Chem* 278, 2630-2635

16. Scully, R., Ganesan, S., Vlasakova, K., Chen, J., Socolovsky, M., and Livingston, D. M. (1999) *Mol Cell* 4, 1093-1099

17. Yarden, R. I. and Brody, L. C. BRCA1 interacts with components of the histone deacetylase complex. (1999) *Proc Natl Acad Sci U S A* 96, 4983-
18. Cantor, S. B., Bell, D. W., Ganesan, S., Kass, E. M., Drapkin, R., Grossman, S., Wahrer, D. C., Sgroi, D. C., Lane, W. S., Haber, D. A., and Livingston, D. M.. (2001) *Cell* **105**, 149-160

19. Yu, X., Wu, L. C., Bowcock, A. M., Aronheim, A., and Baer, R. (1998) *J Biol Chem* **273**, 25388-25392

20. Li, S., Chen, P. L., Subramanian, T., Chinnadurai, G., Tomlinson, G., Osborne, C. K., Sharp, Z. D., and Lee, W. H. (1999) *J Biol Chem* **274**, 11334-11338

21. Ludwig, T., Fisher, P., Ganesan, S., and Efstratiadis, A. (2001) *Genes Dev* **15**, 1188-93

22. Couch, F. J. and Weber, B. L. (1996) *Hum Mutat* **8**, 8-18

23. Futreal, P. A., Liu, Q., Shattuck-Eidens, D., Cochran, C., Harshman, K., Tavtigian, S., Bennett, L. M., Haugen-Strano, A., Swensen, J., Miki, Y. et al. (1994) *Science* **266**, 120-122

24. Hayes, F., Cayanan, C., Barilla, D., and Monteiro, A. N. (2000) *Cancer Res* **60**, 2411-2418

25. Vallon-Christersson, J., Cayanan, C., Haraldsson, K., Loman, N., Bergthorsson, J. T., Brondum-Nielsen, K., Gerdes, A. M., Moller, P., Kristoffersson, U., Olsson, H., Borg, A., and Monteiro, A. N. (2001) *Hum Mol Genet* **10**, 353-60

26. Carvalho, M. A., Billack, B., Chan, E., Worley, T., Cayanan, C., and Monteiro, A. N. (2002) *Cancer Biol Ther* **1**, 502-508

27. Worley, T., Vallon-Christersson, J., Billack, B., Borg, A., and Monteiro, A. N. (2002) *Cancer Biol Ther* **1**, 497-501

28. Joo, W. S., Jeffrey, P. D., Cantor, S. B., Finnin, M. S., Livingston, D. M., and Pavletich, N. P. (2002) *Genes Dev* **16**, 583-593

29. Sodeoka, M. Larson C. J. Chen L. Leclair K. P. and Verdine G. L. A (1993) *Bioorg. Med. Chem. Lett* **3**, 1089

30. Horton, R. M., Ho, S. N., Pullen, J. K., Hunt, H. D., Cai, Z., and Pease, L. R. (1993) *Methods Enzymol* **217**, 270-9

31. Esnouf, R. M. (1997) *J Mol Graph Model* **15**, 132-134, 112-113

32. Esnouf, R. M. (1999) *Acta Crystallogr D Biol Crystallogr* **55**, 938-940
33. Chasman, D. and Adams, R. M. (2001) J Mol Biol 307, 683-706
34. Markiewicz, P., Kleina, L. G., Cruz, C., Ehret, S., and Miller, J. H. (1994) J Mol Biol 240, 421-433
35. Rennell, D., Bouvier, S. E., Hardy, L. W., and Poteete, A. R. (1991) J Mol Biol 222, 67-88
36. Henikoff, S., Henikoff, J. G., and Pietrokovski, S. (1999) Bioinformatics 15, 471-479
37. 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
38. Huyton, T., Bates, P. A., Zhang, X., Sternberg, M. J., and Freemont, P. S. (2000) Mutat Res 460, 319-332
39. Callebaut, I. and Mornon, J. P. (1997) FEBS Lett 400, 25-30
40. Bork, P., Hofmann, K., Bucher, P., Neuwald, A. F., Altschul, S. F., and Koonin, E. V. (1997) FASEB J 11, 68-76
41. Deffenbaugh, A. M., Frank, T. S., Hoffman, M., Cannon-Albright, L., and Neuhausen, S. L. (2002) Genet Test 6, 119-121
42. Monteiro, A. N., August, A., and Hanafusa, H. (1997) Am J Hum Genet 61, 761-762
43. Sjolander, K., Karplus, K., Brown, M., Hughey, R., Krogh, A., Mian, I. S., and Haussler, D. (1996) Comput Appl Biosci 12, 327-45
44. Chapman, M. S. and Verma, I. M. (1996) Nature 382, 678-679
45. Monteiro, A. N., August, A., and Hanafusa, H. (1996) Proc Natl Acad Sci U S A 93, 13595-13599
46. Haile, D. T. and Parvin, J. D. (1999) J Biol Chem 274 2113-2117
47. Somasundaram, K., Zhang, H., Zeng, Y. X., Houvras, Y., Peng, Y., Zhang, H., Wu, G. S., Licht, J. D., Weber, B. L., and El-Deiry, W. S. (1997) Nature 389, 187-190
48. Harkin, D. P., Bean, J. M., Miklos, D., Song, Y. H., Truong, V. B., Englert, C., Christians, F. C., Ellisen, L. W., Maheswaran, S., Oliner, J. D., and Haber, D. A. (1999) Cell 97, 575-586
49. Scully, R., Anderson, S. F., Chao, D. M., Wei, W., Ye, L., Young, R. A., Livingston, D. M., and Parvin, J. D. (1997) Proc Natl Acad Sci U S A 94,
50. Anderson, S. F., Schlegel, B. P., Nakajima, T., Wolpin, E. S., and Parvin, J. D. (1998) *Nat Genet* **19**, 254-256

51. Imoto, T., Yamada, H., and Ueda, T. (1986) *J Mol Biol* **190**, 647-649

52. Wang, L. and Kallenbach, N. R. (1998) *Protein Sci* **7**, 2460-2464

53. Sieber, V., Pluckthun, A., and Schmid, F. X. (1998) *Nat Biotechnol* **16**, 955-960

54. Pedersen, J. S., Otzen, D. E., and Kristensen, P. (2002) *J Mol Biol* **323**, 115-123

55. Ramensky, V., Bork, P., and Sunyaev, S. (2002) *Nucleic Acids Res* **30**, 3894-3900

56. Saunders, C. T. and Baker, D. (2002) *J Mol Biol* **322**, 891-901

57. Petrij-Bosch, A., Peelen, T., van Vliet, M., van Eijk, R., Olmer, R., Drusedau, M., Hogervorst, F. B., Hageman, S., Arts, P. J., Ligtenberg, M. J., Meijers-Heijboer, H., Klijn, J. G., Vasen, H. F., Cornelisse, C. J., van't Veer, L. J., Bakker, E., van Ommen, G. J., and Devilee, P. (1997) *Nat Genet* **17**, 341-345

58. Vossen, R and den-Dunnen J. T. (1999) *Current Protocols in Human Genetics*. John Wiley &Sons, Inc, 9.11.1-9.11.21

59. Bullock, A. N. and Fersht, A. R. (2001) *Nat Rev Cancer* **1**, 68-76
FOOTNOTES:

* To whom correspondence should be addressed. Tel.: 1-780-492-2136;
Fax: 1-780-492-0886; E-mail: mark.glover@ualberta.ca

‡ This work was supported by funding from the Canadian Breast Cancer Research Initiative, the Canadian Institutes of Health Research, the National Cancer Institute of Canada, and the Alberta Heritage Foundation for Medical Research.

ABBREVIATIONS:

The abbreviations used are: BRCT, BRCA1 carboxy-terminal domain; BIC, Breast Cancer Information Core Database;
Figure Legends

Fig. 1. Destabilization of the BRCT domains by truncation mutation. The indicated BRCT truncation mutations were *in vitro* transcribed and translated, and then digested with increasing amounts of trypsin. Lanes 1-4: 0 µg/mL, 6 µg/mL, 60 µg/mL and 600 µg/mL trypsin. Translated protein products with C-termini at positions 1804, 1828, and 1852, correspond to the deletion products of the cancer-predisposing mutations whose truncation effects are depicted in Figure 2.

Fig. 2. Structural effects of cancer-associated BRCA1 BRCT truncation-causing mutations. A. A stop codon at position 1805 results from frameshift IVS21-36del510, removing much of the C-terminal BRCT domain. B. Frameshift 5382insC creates a stop codon at position 1829 in BRCA1 and is one of the most commonly recorded BIC mutations. C. A nonsense mutation 1853-ter results in the removal of the 11 C-terminal residues of the protein, and is linked to disease. For A-C, red portions of the structures are deleted residues caused by truncation mutations D. Interaction of the C-terminal tail of BRCA1 with BRCT-C. Negative electrostatic potential is red and positive is blue. The C-terminus of BRCA1 forms a $3_{10}$ helix and an extended peptide that packs against $\alpha2'$ and the $\beta$-sheet. C-terminal deletions beyond the hydrophobic residues Leu-1854 and Ile-1855 are destabilizing.
Fig. 3. Amino acid sequence alignment of the BRCT repeat region of cloned BRCA1 homologues. Secondary structure elements are from the human BRCT repeat structure, RCSB: 1JNX. The positions of the 25 missense mutations studied here are indicated. Numbering is for Human BRCA1. Alignments were created with ClustalX (NCBI accession numbers: human - AAA73985, canine - AAD56289, rat - AAC36493, mouse - AAD00168, chicken - AAK83825, xenopus - AAL13037).

Figure 4. Destabilization of the BRCT domains by missense mutations. A. The indicated missense mutations were digested with increasing concentration amounts of trypsin. Lanes 1-4: 0 µg/mL, 6 µg/mL, 60 µg/mL and 600 µg/mL trypsin. B. Mutations harbouring a trypsin cleavage site were digested with chymotrypsin. Lanes 1-4: 0 µg/mL, 6 µg/mL, 60 µg/mL and 600 µg/mL chymotrypsin.

Figure 5. Quantification and Classification of the structural effects of BRCT missense mutations. A-D: The tested BRCT missense mutations have been divided into 4 classes based on their distribution in the BRCT structure (see text). The fraction remaining is the percent of starting protein present following digestion with the indicated concentrations of trypsin. Data points are the mean value of digestions performed in triplicate with error bars reflecting the standard deviations. Red - highly destabilizing mutation (BRCT variant is >60% degraded at 6 µg/mL trypsin, Blue - Intermediately destabilizing mutation (BRCT variant is
>60% degraded at 60 µg/mL trypsin), Black/Grey - no destabilizing effect (similar to wild type tryptic sensitivity). **E-H:** Structural distribution of BRCA1-BRCT missense mutations. **A,E:** Surface mutations. **B,F:** Interface mutations **C,G:** BRCA1-fold mutations. **D,H:** BRCT-fold mutations.

**Figure 6. Proposed application of BRCT protease susceptibility assays in BRCA1 screening protocols.** A. The commonly used protein truncation test could be modified to include a protease digestion step. This method typically generates coding sequence (and protein) from both alleles in an individual. In the case of a heterozygote carrier for missense mutation, highly destabilizing mutants would result in a 50% reduction in protein amounts of the BRCT. The quantitative nature of the protease susceptibility assay would allow one to monitor the disappearance of protein species “A”. Proteolytic fragments (species “C”) are also generated in a predictable manner for the wild type protein. Greater sensitivity for the assay could thus be achieved by monitoring the ratio of protein species C/A. For truncating mutations, truncated protein products (“B”) are produced that would be rapidly degraded by protease. B. BRCT missense mutations identified by sequencing could be generated by PCR and transcribed directly from PCR products, eliminating the need for a cloning step.
Table 1. Structure, function and disease effects of BRCT missense mutations

| Mutant | Secondary Structure* | Mutant Class | Protease Sensitivity b | Predictive Method 1 c | Predictive Method 2 d | Transcription e | Disease effects f | Solubility and Stability g |
|--------|----------------------|--------------|------------------------|--------------------|----------------------|----------------|------------------|-----------------------------|
| M1652I | β                    | Brct-fold    | (-)                    | (+)               | (+)                  | t/c(-) | (+)               | Soluble                      |
| D1692Y | c                    | Surface      | (+)                    | (-)               | (+)                  | t/c(-) | (+)               | Insoluble                     |
| F1695L | c                    | Surface      | (-)                    | (-)               | (-)                  | ?      | Soluble           |                             |
| V1696L | c                    | Surface      | (+)                    | (-)               | (-)                  | ?      | Soluble           |                             |
| C1697R | c                    | BRCA1-fold   | (++)                   | (+)               | (+)                  | t/c (+) | (+)               | Insoluble                     |
| R1699W | c                    | Interface    | (+)                    | (+)               | (+)                  | t/c (+) | (+)               | Soluble                      |
| R1699Q | c                    | Interface    | (-)                    | (-)               | (-)                  | t/c (-) | ?                 | Soluble                      |
| A1708E | a                    | Interface    | (+)                    | (+)               | (+)                  | t/c (+) | (+)               | Insoluble                     |
| S1715R | β                    | BRCA1-fold   | (++)                   | (+)               | (+)                  | t/c (+) | (+)               |                               |
| W1718C | a                    | Brct-fold    | (++)                   | (+)               | (+)                  | ?      | ?                 |                             |
| T1720A | a                    | Surface      | (-)                    | (-)               | (-)                  | ?      | ?                 |                             |
| G1738E | c                    | BRCA1-fold   | (++)                   | (+)               | (+)                  | t/c (+) | (+)               | Insoluble                     |
| G1738R | c                    | BRCA1-fold   | (++)                   | (+)               | (+)                  | ?      | Insoluble         |                             |
| P1749R | a                    | BRCA1-fold   | (++)                   | (+)               | (+)                  | ?      | Insoluble         |                             |
| R1751Q | a                    | BRCA1-fold   | (+)                    | (-)               | (-)                  | ?      | Soluble           |                             |
| A1752P | a                    | BRCA1-fold   | (++)                   | (-)               | (-)                  | ?      | ?                 |                             |
| I1766S | β                    | Brct-fold    | (++)                   | (+)               | (+)                  | ?      | Insoluble         |                             |
| M1775R | c                    | Interface    | (+)                    | (-)               | (+)                  | t/c (+) | (+)               | ∆ΔG= 5.0 kcal/mol             |
| M1783T | a                    | Interface    | (+)                    | (+)               | (-)                  | ?      | ∆ΔG= 4.28 kcal/mol |                             |
| G1788V | c                    | Brct-fold    | (++)                   | n/a               | (+)                  | ?      | Insoluble         |                             |
| V1804D | c                    | Surface      | (-)                    | (-)               | (-)                  | ?      | ?                 |                             |
| V1808F | β                    | Brct-fold    | (++)                   | (+)               | (-)                  | ?      | Soluble           |                             |
| W1837R | a                    | Brct-fold    | (++)                   | (+)               | (+)                  | ?      | Insoluble         |                             |
| W1837G | a                    | Brct-fold    | (++)                   | (+)               | (-)                  | ?      | Insoluble         |                             |
| Y1853C | a                    | Brct-fold    | (++)                   | (+)               | (+)                  | ?      | Soluble           |                             |

* Secondary structure is from the human BRCT domain structure (13)

b Protease sensitivity: (-) Wild type, no effect, (+) Intermediately destabilizing, (++) Highly destabilizing

c Predictive method 1. Predicted effect on function is as described by Chasman and Adams (33). (+): The mutation is predicted to effect structure/function, the probability of an effect on function is >0.5. (-): The mutation is predicted to be a benign substitution, the probability of an effect on structure/function <0.5 (See supplementary Table 1 for calculated probabilities). For one mutation, G1788V, there were too few data points to estimate a probability.

d Predictive method 2, sequence based. (+): The mutation is predicted to effect structure/function, (-): The mutation is predicted to be a benign substitution.

e Transcription effects are those reported by Monteiro et al 1996 (45), Monteiro et al 1997 (42), Hayes et al. 2000 (24), Vallon-Christersson et al. 2001 (25), and Worley et al. 2002 (27).

f Transcription (-) : no effect on transcription, t/c (+): affects t/c

g Disease linkage data is from recorded entries in the BIC, Hayes et al. 2000 (24) Vallon-Christersson et al. 2001 (25), and T.S. Frank, personal communication. (+): linked to disease, (-): not linked, (?): unknown

h Stabilities reported from Ekblad et al. 2002 (14). The reported solubility in E. coli is from this study or Ekblad et al. 2002 (14).
Supplementary Table 1: Structure and Sequence based predictions of the effects of missense mutations on the BRCA1 C-terminal domain

| Mutant | Secondary Structure | Mutant Class | Method 1 | Method 1 | Sequence based |
|--------|---------------------|--------------|----------|----------|---------------|
|        |                     | Feature Set A | Prob. of effect | Feature Set B | Prob. of effect | (+) Predicted to effect function |
|        |                     |              |          |          | (-) No predicted effect function |
| M1652  | β                   | Brct-fold    | 0.74     | 0.71     | -              |
| V1653M | β                   | Brct-fold    | 0.72     | 0.63     | +              |
| S1655F | β                   | BRCA1-fold   | 0.66     | 0.59     | +              |
| L1664P | a                   | BRCA1-fold   | 0.13     | 0.09     | +              |
| V1665M | a                   | BRCA1-fold   | 0.74     | 0.64     | +              |
| A1669S | c                   | Surface     | 0.03     | 0.04     | +              |
| E1682K | β                   | Brct-fold    | 0.55     | 0.50     | +              |
| T1685  | β                   | BRCA1-fold   | 0.66     | 0.59     | +              |
| T1685A | β                   | BRCA1-fold   | 0.50     | 0.59     | -              |
| M1689R | β                   | Brct-fold    | 0.80     | 0.84     | +              |
| D1692N | c                   | Surface     | 0.20     | 0.20     | +              |
| D1692Y | c                   | Surface     | 0.20     | 0.20     | -              |
| F1695L | c                   | Surface     | 0.13     | 0.18     | -              |
| V1696L | c                   | Surface     | 0.30     | 0.30     | -              |
| C1697R | c                   | BRCA1-fold   | 0.82     | 0.82     | +              |
| R1699W | c                   | Interface   | 0.59     | 0.58     | +              |
| R1699Q | c                   | Interface   | 0.59     | 0.58     | +              |
| R1699L | c                   | Interface   | 0.59     | 0.58     | -              |
| T1700A | c                   | BRCA1-fold   | 0.60     | 0.59     | +              |
| G1706A | a                   | BRCA1-fold   | 0.66     | 0.54     | +              |
| G1706E | a                   | BRCA1-fold   | 0.94     | 0.94     | +              |
| A1708E | c                   | Interface   | 0.83     | 0.83     | +              |
| V1713A | β                   | Brct-fold    | 0.63     | 0.61     | +              |
| S1715N | β                   | BRCA1-fold   | 0.94     | 0.93     | -              |
| S1715C | β                   | BRCA1-fold   | 0.74     | 0.64     | +              |
| W1718S | c                   | BRCA1-fold   | 0.74     | 0.64     | +              |
| D1739Y | c                   | BRCA1-fold   | 0.30     | 0.30     | +              |
| D1739V | c                   | BRCA1-fold   | 0.30     | 0.30     | +              |
| D1739E | c                   | BRCA1-fold   | 0.30     | 0.30     | -              |
| D1739G | c                   | BRCA1-fold   | 0.30     | 0.30     | -              |
| V1741G | c                   | BRCA1-fold   | 0.19     | 0.10     | +              |
| H1746N | c                   | Surface     | 0.37     | 0.23     | +              |
| P1771L | c                   | BRCA1-fold   | 0.82     | 0.82     | +              |
| T1773S | c                   | BRCA1-fold   | 0.05     | 0.07     | -              |
| M1775R | c                   | BRCA1-fold   | 0.05     | 0.07     | +              |
| D1778G | c                   | BRCA1-fold   | 0.36     | 0.24     | -              |
| L1764P | β                   | Brct-fold    | 0.32     | 0.24     | +              |
| I1766S | β                   | Brct-fold    | 0.67     | 0.62     | +              |
| N1776L | c                   | BRCA1-fold   | 0.00     | 0.00     | +              |
| T1777S | c                   | BRCA1-fold   | 0.14     | 0.14     | +              |
| G1788V | a                   | BRCA1-fold   | 0.42     | 0.42     | +              |
| G1788D | c                   | BRCA1-fold   | 0.47     | 0.47     | +              |
| V1804D | c                   | BRCA1-fold   | 0.00     | 0.03     | -              |
| P1806A | b                   | BRCA1-fold   | 0.15     | 0.22     | -              |
| V1809A | β                   | Brct-fold    | 0.48     | 0.43     | +              |
| V1809F | β                   | Brct-fold    | 0.48     | 0.43     | -              |
| V1810G | β                   | Brct-fold    | 0.15     | 0.15     | +              |
| Q1811G | β                   | BRCA1-fold   | 0.00     | 0.00     | +              |
| A1823T | c                   | Brct-fold    | 0.06     | 0.54     | +              |
| V1833M | β                   | BRCA1-fold   | 0.66     | 0.43     | +              |
| L1835F | c                   | BRCA1-fold   | 0.43     | 0.41     | +              |
| E1836K | a                   | Interface   | 0.48     | 0.42     | +              |
| W1837Y | c                   | BRCA1-fold   | 0.94     | 0.93     | +              |
| W1837G | a                   | BRCA1-fold   | 0.74     | 0.64     | +              |
| S1841N | a                   | BRCA1-fold   | 0.90     | 0.90     | +              |
| A1843P | a                   | Interface   | 0.65     | 0.64     | +              |
| Y1854G | a                   | Brct-fold    | 0.61     | 0.56     | +              |
| P1859R | c                   | Surface     | None     | None     | +              |
### Fig. 1

| C-terminal sequence | C-terminal Residue |
|---------------------|--------------------|
| QCQELDTYLPQIPHSHY   | 1863               |
| QCQELDTYLPQIP       | 1859               |
| QCQELDTYLPQI        | 1858               |
| QCQELDTYLPQ         | 1857               |
| QCQELDLYLI          | 1855               |
| QCQELDTY            | 1853               |
| QCQELDT             | 1852               |
| QCQELD              | 1851               |
| QCQE                | 1849               |
| ......               | 1828               |
| ....                | 1804               |

Trypsin activity profile for the indicated C-terminal sequences.
Fig. 2.

A

1805-Stop

B

1829-Stop

C

1853-ter

D

1853

1854

1855

1856

1857

1858

1859

1848

1849

1850

1851

1852

1853
Fig. 3.

N-terminal BRCT

| Mutation | 1 | 2 | 3 | 4 | Q |
|----------|---|---|---|---|---|
| brct_human | VNKMSN6GVS6LEEFMVRYPKHI5HIALNL5ITETVHFVMQTAETC5LKVFL | 1705 |
| brct_canine | VNKIR6SVAS6TTFEPMLVNP6KHI5HIALNL5ITETVHFVMQTAETC5LKVFL |
| brct_rat | AERIS6M6V65LEEFMVVFQ6R6YAL5TDV73BM6FTV6IM5TAETC5LKVFL |
| brct_mouse | AERIS6M6V65LEEFMVVFQ6R6YAL5TDV73BM6FTV6IM5TAETC5LKVFL |
| brct_chicken | CR7E5MVA6SL6Q6H5LY6V6P6AT6Q6ST6F6R6Y6M5D6ET6E6L6V6C5L6KV6L |
| brct_xenopus | SRR5L6F6V6A6L6Q6C5E5H5AI6V6Q6K6T6Q6S6515560656160656160656160 |

| R | C | A | R | Q |
|---|---|---|---|---|
| brct_human | GI6AG6N6V6V6S6Y6V6W6V6Q6S66K5M5L6N6H5D5E6F5V6R6G6V6Q6Y6N6H5Q6P6R6A6S6Q6R6Q6R6K |
| brct_canine | GI6AG6N6V6V6S6Y6V6W6V6Q6S66K5M5L6N6H5D5E6F5V6R6G6V6Q6Y6N6H5Q6P6R6A6S6Q6R6K |
| brct_rat | GI6AG6N6V6V6S6Y6V6W6V6Q6S66K5M5L6N6H5D5E6F5V6R6G6V6Q6Y6N6H5Q6P6R6A6S6Q6R6K |
| brct_mouse | GI6AG6N6V6V6S6Y6V6W6V6Q6S66K5M5L6N6H5D5E6F5V6R6G6V6Q6Y6N6H5Q6P6R6A6S6Q6R6K |
| brct_chicken | GI6AG6N6V6V6S6Y6V6W6V6Q6S66K5M5L6N6H5D5E6F5V6R6G6V6Q6Y6N6H5Q6P6R6A6S6Q6R6K |
| brct_xenopus | GI6AG6N6V6V6S6Y6V6W6V6Q6S66K5M5L6N6H5D5E6F5V6R6G6V6Q6Y6N6H5Q6P6R6A6S6Q6R6K |

C-terminal BRCT

| Mutation | R | C | T | Y | D |
|----------|---|---|---|---|---|
| brct_human | FR6GL6E6C6C6C6F6T6F6M6T6D6L6E6W6V6L6C6A6V6V6K6E6L6S6P6T6F6L6S6T6G6H6T6V6P6Q6P6D6K6E6T6D6S6 |
| brct_canine | FR6GL6E6C6C6C6F6T6F6M6T6D6L6E6W6V6L6C6A6V6V6K6E6L6S6P6T6F6L6S6T6G6H6T6V6P6Q6P6D6K6E6T6D6S6 |
| brct_rat | FR6GL6E6C6C6C6F6T6F6M6T6D6L6E6W6V6L6C6A6V6V6K6E6L6S6P6T6F6L6S6T6G6H6T6V6P6Q6P6D6K6E6T6D6S6 |
| brct_mouse | FR6GL6E6C6C6C6F6T6F6M6T6D6L6E6W6V6L6C6A6V6V6K6E6L6S6P6T6F6L6S6T6G6H6T6V6P6Q6P6D6K6E6T6D6S6 |
| brct_chicken | FR6GL6E6C6C6C6F6T6F6M6T6D6L6E6W6V6L6C6A6V6V6K6E6L6S6P6T6F6L6S6T6G6H6T6V6P6Q6P6D6K6E6T6D6S6 |
| brct_xenopus | FR6GL6E6C6C6C6F6T6F6M6T6D6L6E6W6V6L6C6A6V6V6K6E6L6S6P6T6F6L6S6T6G6H6T6V6P6Q6P6D6K6E6T6D6S6 |

G

| Mutation | C | R | C | R |
|----------|---|---|---|---|
| brct_human | F9A6I6Q6M6C6A6EF6V6R6E6V6L6D6V6S6A6L6V6Q6C6E6L6D6T6L6Y6P6I6P6R6S6H6Y6 | 1863 |
| brct_canine | F9A6I6Q6M6C6A6EF6V6R6E6V6L6D6V6S6A6L6V6Q6C6E6L6D6T6L6Y6P6I6P6R6S6H6Y6 |
| brct_rat | CP9D6I6Q6L6C6G6R6L6V6S6D6V6D6S6V6S6T6C6T6R6L6D6T6L6Y6P6I6P6R6S6H6Y6 |
| brct_mouse | CP9D6I6Q6L6C6G6R6L6V6S6D6V6D6S6V6S6T6C6T6R6L6D6T6L6Y6P6I6P6R6S6H6Y6 |
| brct_chicken | Y9A6I6Q6M6C6A6EF6V6R6E6V6L6D6V6S6A6L6V6Q6C6E6L6D6T6L6Y6P6I6P6R6S6H6Y6 |
| brct_xenopus | Y9A6I6Q6M6C6A6EF6V6R6E6V6L6D6V6S6A6L6V6Q6C6E6L6D6T6L6Y6P6I6P6R6S6H6Y6 |
Fig. 4.

A

|      | Trypsin  | Trypsin  | Trypsin  |
|------|----------|----------|----------|
|      | 1 2 3 4  | 1 2 3 4  | 1 2 3 4  |
| WT   |          |          |          |
| M1652I |          |          |          |
| D1692Y |          |          |          |
| F1695L |          |          |          |
| V1696L |          |          |          |
| C1697R |          |          |          |
| R1699W |          |          |          |
| R1699Q |          |          |          |
| A1708E |          |          |          |
| S1715R |          |          |          |
| W1718C |          |          |          |
| T1720A |          |          |          |
| G1738E |          |          |          |

B

|      | Chymotrypsin  |
|------|---------------|
|      | 1 2 3 4       |
| WT   |               |
| C1697R |               |
| S1715R |               |
| G1738R |               |
| P1749R |               |
| W1837R |               |
| M1775R |               |
| V1804D |               |
| V1809F |               |
| W1837G |               |
| Y1853C |               |
Fig. 5. A-D

A

Surface

Fraction remaining

[ Trypsin ] (µg/mL)

B

Interface

Fraction remaining

[ Trypsin ] (µg/mL)

C

BRCA1-fold

Fraction remaining

[ Trypsin ]

D

BRCT-fold

Fraction remaining

[ Trypsin ] (µg/mL)
Fig. 5. E-H
Fig. 6.

A. A combined PTT/protease assay for direct screening of patient isolated RNA

- RNA
- RT
- cDNA
- PCR
- Transcription/Translation
- Proteolysis and SDS-Page

B. PCR generation of a BRCT missense mutation identified by sequencing.

- PCR SOEing
- Transcription/Translation
- Proteolysis and SDS-Page
Detection of protein folding defects caused by BRCA1-BRCT truncation and missense mutations
R. Scott Williams, Daniel Chasman, David Hau, Benjamin Hui, Al Lau and J.N. Mark Glover

*J. Biol. Chem.* published online October 8, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M310182200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2003/10/29/M310182200.DC1