Missense Mutations in the BRCT Domain of BRCA-1 from High-Risk Women Frequently Perturb Strongly Hydrophobic Amino Acids Conserved among Mammals

Meaghan A. Figge and Lynda Blankenship

Albany Academy for Girls, Albany, New York

Abstract

Inherited missense mutations in the tumor suppressor gene, *BRCA-1*, may predispose to breast or ovarian cancer, but the exact effects on the protein are generally unknown. The COOH-terminal region of BRCA-1 encodes two BRCT repeats, which are partially conserved in mammalian species (human, dog, rat, and mouse; 60% amino acid identity). A bioinformatic analysis was conducted to evaluate 246 BRCT missense mutations from high-risk breast and/or ovarian cancer patients (reported in the NIH Breast Cancer Information Core database). It was hypothesized that amino acids conserved in evolution would be disproportionately targeted by the mutations and that conserved amino acids with strongly hydrophobic side chains would be disproportionately perturbed. A statistical model was developed, and χ^2 tests were used to determine whether missense mutations are randomly distributed throughout the BRCT repeats or whether they disproportionately target certain amino acids. The results showed that missense mutations disproportionately target amino acids that are identical in all four mammals ($\chi^2 = 46.01$, $P < 0.001$). In addition, missense mutations disproportionately perturb conserved amino acids with strongly hydrophobic side chains ($\chi^2 = 68.57$, $P < 0.001$) and alter the strongly hydrophobic property. The two most frequently observed known cancer-predisposing missense mutations in the BRCT repeats, M1775R and A1708E, conform to this pattern. These results suggest that missense mutations affecting highly conserved amino acids with strongly hydrophobic side chains can disturb important features of the BRCA-1 protein and may play a role in breast and ovarian cancer formation. (Cancer Epidemiol Biomarkers Prev 2004;13(6):1037–41)

Introduction

It is estimated that 5% to 10% of all breast cancer cases are caused by inherited mutations in the tumor suppressor genes, *BRCA-1* and *BRCA-2* (1). *BRCA-1* mutations account for 40% to 50% of families with only hereditary breast cancer (2) and confer a 56% to 87% lifetime risk of developing breast cancer (1). *BRCA-2* mutations are linked to the formation of breast cancer in men and account for the other 50% of families with only inherited breast cancer (2). The majority of breast-ovarian cancer families are due to *BRCA-1* mutations, and most others are due to *BRCA-2* (3). Cancer-predisposing alleles of *BRCA-1* are generally recessive, and the wild-type allele is typically lost in tumor tissue (1).

The COOH-terminal BRCT domain of the BRCA-1 protein, harboring two BRCT repeats, is an evolutionarily conserved region that plays an active role in the stability of the protein's conformation (4). The BRCT domain is characterized by hydrophobic clusters of amino acids that are thought to stabilize the three-dimensional structure of the protein. These clusters are mainly found at the NH_2 and COOH termini of the domain on α-1 and α-3 and within the central β-sheet (4). The domain is essential for DNA repair (5-9) and binding (10, 11), cell cycle control (12), regulation of gene expression (13), and tumor suppressor functions (14). Evidence suggests that the BRCT domain is a protein-protein interaction macromolecule that interacts with other cellular factors involved in transcriptional control or DNA repair, including CtIP polypeptide (a substrate for the ATM protein kinase), p53, p300, and BACH1 (5, 14). Loss of these normal functions mediated by the BRCT domain by inherited mutations is thought to increase the risk of developing breast and ovarian cancer (13).

The two BRCT repeats of BRCA-1 have similar three-dimensional structures and are packed together in a head-to-tail arrangement. The secondary structure of the BRCT domain has been elucidated by X-ray crystallography. The BRCT domain is the only domain of BRCA-1 that has a known secondary structure. Each repeat is characterized by the following: β-1, α-1, β-2, β-3, α-2, β-4, and α-3 (14). The two BRCT repeats in BRCA-1 are attached by an α-helix called the linker and are joined by a salt bridge between Arg^{1699}, Glu^{1836}, and Asp^{1846} (14).

The BRCT sequences from human, dog, rat, and mouse were compared (15). The mouse and rat amino

Received 12/3/03; revised 2/7/04; accepted 2/12/04.

Grant support: Presented by M. Figge at the Intel International Science and Engineering Fair, May 2003, Cleveland, Ohio.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Meaghan A. Figge, Albany Academy for Girls, Albany, NY 12208. Phone: 518-463-2201; Fax: 518-463-5096. E-mail: figgem@nycap.rr.com

acid sequences are each 65% identical to human; the dog sequence is 92% identical to human. In the tandem BRCT repeats, 128 of the 214 amino acids (60%) are identical among the four species.

Seventy-five unique types of missense mutations (single amino acid substitutions), reported in the NIH Breast Cancer Information Core (BIC) database, are in the BRCT domain of BRCA-1 (16). All of the missense mutations that have been reported in the database have unclassified or unknown effects on the BRCA-1 protein, except for R1699W, R1699L, R1699Q, A1708E, P1749R, and M1775R. The majority of the 75 unique missense mutations have been reported more than once in the database, therefore creating a total frequency of 246 missense mutations. These 75 types of missense mutations were identified in breast and/or ovarian cancer patients from either high-risk families or patient series (17).

In this report, we tested the hypothesis that observed missense mutations in the BRCA-1 BRCT domain disproportionately target amino acids that are evolutionarily conserved among mammals. In addition, we tested the hypothesis that observed missense mutations target evolutionarily conserved amino acids with strongly hydrophobic side chains. These amino acids may be responsible for the folding of the BRCT domain through hydrophobic clustering in the interior of the protein. These hypotheses were tested statistically by an analysis of the 246 missense mutations in the NIH BIC database (16).

Methods

A total of 246 BRCT missense mutations from high-risk breast and/or ovarian cancer patients were taken from the NIH BIC database and analyzed. Only information from probands and tumors in which mutations have been identified is recorded in the BIC (17); thus, each mutation represents an independent data point.

Next, the human BRCT sequence (amino acid positions 1646 to 1859) was aligned with canine, rat, and mouse BRCT sequences, and the identical amino acids were identified. To classify the hydrophobicity of amino acid side chains, the 20 amino acids were categorized into four hydrophobicity groups (see Table 1) based on the Eisenberg et al. (18) consensus hydrophobicity scale for amino acids. This scale is based on averaged values from five individual scales (19-23). Nozaki and Tanford (19) derived their values from the free energies of transferring hydrophobic side chains and backbone peptide units from water to 100% ethanol and dioxane. The Janin (20) scale is the ratio of buried to accessible molar fractions of each amino acid in globular proteins. The Chothia (21) scale was based on the proportion of residues 95% buried in 12 proteins. Wolfenden et al. (22) measured the distribution of amino acid side chains between diluted aqueous solutions and the vapor phase at 25°C. von Heijne and Blomberg (23) estimated the free energy for transfer of a residue in a polypeptide from a random coil in an aqueous phase to a helix conformation in a nonpolar environment.

Table 1. Classification of amino acid hydrophobicity

Strongly Hydrophobic: $x \geq 0.25$	Moderately Hydrophobic: $0.25 > x > 0$	Moderately Hydrophilic: $-0.25 < x < 0$	Strongly Hydrophilic: $x \leq -0.25$
Ile 0.73	Gly 0.16	Pro −0.07	Ser −0.26
Phe 0.61	Cys 0.04	Thr −0.18	His −0.40
Val 0.54	Tyr 0.02		Glu −0.62
Leu 0.53			Asn −0.64
Trp 0.37			Gln −0.69
Met 0.26			Asp −0.72
Ala 0.25			Lys −1.10
			Arg −1.76

NOTE: The values used in this table were taken from Eisenberg et al. (18). These are consensus values compiled from five individual scales (19-23).

Arbitrary cutoff values were established for each hydrophobicity group (Table 1). Amino acids with $x \geq 0.25$ were defined as ''strongly hydrophobic,'' $0.25 > x > 0$ as ''moderately hydrophobic,'' $x \leq -0.25$ as ''strongly hydrophilic,'' and $-0.25 < x < 0$ as ''moderately hydrophilic.'' This method of grouping is similar to that employed by others for hydrophobic cluster analysis (24), except of the inclusion of alanine, which is only 0.01 point away from methionine, as strongly hydrophobic, and tyrosine as moderately hydrophobic. Furthermore, it was determined whether a strongly hydrophobic side chain was conserved among all four mammals at each codon position.

A frequency distribution analysis table was created to report the number of missense mutations that affect each codon position.

The χ^2 statistic (25) was used to determine whether the missense mutations were randomly distributed throughout the BRCT domain or whether the mutations disproportionately target certain amino acids. One degree of freedom was used for the tests. The requirement for significance was $P < 0.001$, corresponding to $\chi^2 > 10.83$. The χ^2 equation is $\chi^2 = \Sigma\ [(O - E)^2 / E]$. Expected values ($E$) were calculated using a statistical model as explained below. A separate statistical model was developed for each hypothesis.

Each statistical model was developed assuming that single base mutations would be randomly distributed throughout the codons encoding the BRCT domain. The genetic code was used to evaluate the number of ways that each specific codon in the sequence can yield a missense mutation. For testing the first hypothesis, the number of possible missense mutations in conserved amino acids was determined. Next, for testing the second hypothesis, the number of ways a missense mutation could change a conserved strongly hydrophobic property was determined. If the random mutation model were correct, then the expected distribution of missense mutations should be similar to that observed in the BIC database. However, if the mutations do not occur in a random fashion, then it can be demonstrated that mutations disproportionately target certain classes of amino acids (e.g., conserved and conserved strongly hydrophobic).

Results

The analysis of all 214 codons that encode the BRCT domain of BRCA-1 revealed that there are 1,413 ways

that a single base change can result in a missense mutation. The number of possible missense mutations that could affect a conserved amino acid was 837. The number of possible missense mutations that could alter the strongly hydrophobic property of a side chain of a conserved amino acid was 153. See Table 2 for an example of the analysis for a single codon position.

The first χ^2 test was used to determine whether the missense mutations occur at random or whether they disproportionately target evolutionarily conserved amino acids. According to the random mutation model, the probability of missense mutations affecting conserved amino acids is 0.5923 (837 / 1,413). Therefore, according to the model, in a panel of 246 missense mutations, it is expected that 146 mutations would affect conserved amino acids (0.5923 × 246). However, in the BRCA-1 database, 198 missense mutations affect the conserved amino acids. Likewise, it was expected that 100 mutations would affect nonconserved amino acids (246 − 146), but in the database only 48 were present in the nonconserved regions. Therefore, $\chi^2 = [(198 - 145.72)^2 / 145.72] + [(48 - 100.28)^2 / 100.28] = 46.01$. The χ^2 test gave a value of 46.01 with 1 degree of freedom. This value is greater than 10.83 and $P < 0.001$. Consequently, these data demonstrate that the missense mutations disproportionately target evolutionarily conserved amino acids (identical amino acids in humans, dogs, mice, and rats) in the BRCT domain. Mutations occur in the conserved amino acid positions more often than would be expected if the mutations were randomly distributed.

A second χ^2 test was used to determine whether missense mutations occur at random or whether they disproportionately perturb conserved amino acids with strongly hydrophobic side chains. The total number of possible missense mutations that perturb conserved strongly hydrophobic amino acids is 153 (Table 3). Therefore, according to the random mutation model, the probability that a missense mutation would perturb a conserved strongly hydrophobic amino acid is 0.1083 (153 / 1,413). Per the model, it was expected that 27 missense mutations would perturb conserved amino acids with strongly hydrophobic side chains (0.1083 × 246), but in the database, 67 mutations actually perturbed these amino acids. Likewise, it was expected that 219 missense mutations would affect all other amino acids (246 − 27); but 179 missense mutations were present in all other amino acid positions. Therefore, $\chi^2 = [(67 - 26.64)^2$

Table 2. Statistical model: an example of the procedure, based on the genetic code, which Is used to evaluate the number of ways that a specific codon can yield a missense mutation

Codon 1657: CTG → Leucine, strongly hydrophobic

Possible mutations:
TTG → Leucine, strongly hydrophobic
ATG → *Methionine*, strongly hydrophobic
GTG → *Valine*, strongly hydrophobic
CCG → *Proline**, moderately hydrophilic
CAG → *Glutamine**, strongly hydrophilic
CGG → *Arginine**, strongly hydrophilic
CTT → Leucine, strongly hydrophobic
CTC → Leucine, strongly hydrophobic
CTA → Leucine, strongly hydrophobic

NOTE: There are five combinations that code for missense mutations (italics). Only three (*) out of these five perturb the property, ''strongly hydrophobic.'' Refer to Table 1 for the classification of amino acid hydrophobicity. This type of analysis was performed for each of the 214 codon positions in the BRCT domain of human BRCA-1.

Table 3. Analysis of conserved hydrophobic amino acids in the BRCT domain

Amino Acid Position	Codon	Amino Acid	No. of Possible Missense That Alter Strongly Hydrophobic Property	Reported Mutations in BIC That Alter Strongly Hydrophobic Property	No. of Occurrences in BIC
1652	ATG	Met	3	Met → Thr	3
1653	GTG	Val	2		
1657	CTG	Leu	3		
1663	ATG	Met	3		
1665	GTG	Val	2		
1668	TTT	Phe	3		
1669	GCC	Ala	5	Ala → Ser	6
1676	TTA	Leu	1		
1680	ATT	Ile	3		
1687	GTT	Val	2		
1693	GCT	Ala	5		
1695	TTT	Phe	3		
1696	GTG	Val	2		
1701	CTG	Leu	3		
1704	TTT	Phe	3		
1705	CTA	Leu	3		
1707	ATT	Ile	3		
1708	GCG	Ala	5	Ala → Glu	24
1712	TGG	Trp	6		
1714	GTT	Val	2		
1718	TGG	Trp	6	Trp → Cys	1
				Trp → Ser	1
1719	GTG	Val	2		
1723	ATT	Ile	3		
1729	CTG	Leu	3		
1734	TTT	Phe	3	Phe → Ser	1
1736	GTC	Val	2		
1740	GTG	Val	2		
1741	GTC	Val	2	Val → Gly	3
1761	TTC	Phe	3		
1764	CTA	Leu	3	Leu → Pro	2
1772	TTC	Phe	3		
1775	ATG	Met	3	Met → Arg	15
1780	CTG	Leu	3	Leu → Pro	1
1783	ATG	Met	3	Met → Thr	3
1786	CTG	Leu	3		
1789	GCT	Ala	5		
1791	GTG	Val	2		
1792	GTG	Val	2		
1808	GTG	Val	2		
1810	GTG	Val	2	Val → Gly	1
1814	GCC	Ala	5		
1815	TGG	Trp	6		
1824	ATT	Ile	3		
1833	GTG	Val	2		
1837	TGG	Trp	6	Trp → Gly	2
				Trp → Arg	4
1838	GTG	Val	2		
1839	TTG	Leu	1		
1850	CTG	Leu	3		
1854	CTG	Leu	3		
1858	ATC	Ile	3		
Totals			153		67

/ 26.64] + [(179 − 219.36)^2 / 219.36] = 68.57. The second χ^2 test gave a value of 68.57 with 1 degree of freedom. The value is greater than 10.83 and $P < 0.001$. Consequently, missense mutations target and alter the hydrophobic property of conserved amino acids with strongly hydrophobic side chains. Mutations affect these amino acids more often than would be expected if the mutations were randomly distributed.

These results suggest that missense mutations affecting highly conserved amino acids with strongly hydrophobic side chains can disturb important features of the BRCA-1 protein and may play a role in breast and ovarian cancer formation.

Discussion

This report analyzes the potential role of *BRCA-1* missense mutations in patients who undergo genetic testing. Missense mutations make up a significant fraction of the *BRCA-1* gene alterations. This type of mutation, in some cases, has no effect on the protein function; in other cases, it is deleterious. The interpretation of a genetic test result is not straightforward because a missense mutation may be benign; it cannot be assumed that the mutation will predispose to cancer development.

Missense mutations in the BRCT domain from high-risk cancer patients disproportionately target amino acids that are conserved in mammalian species ($P < 0.001$). Conserved amino acids with strongly hydrophobic side chains are particularly susceptible to perturbation by missense mutations ($P < 0.001$). Such mutations may alter critical physicochemical properties that stabilize the BRCT structure or its interactions with other proteins, thereby eliminating normal function and predisposing to breast cancer. Hayes et al. (26) found that the ''BRCA-1 COOH terminus acts as a transcription activation domain, and germ line cancer-predisposing mutations in this region abolish transcription activation, whereas benign polymorphisms do not.'' These authors found that mutations of hydrophobic amino acids in the BRCT domain that are conserved in humans, dogs, mice, and rats abolished transcriptional activation. The authors concluded that ''the integrity of the BRCT domain is crucial for transcription activation and that hydrophobic residues may be important for BRCT functions'' (26).

Two cancer-predisposing mutations, M1775R and A1708E, occur frequently in the BRCT domain (see Table 3). Fifteen M1775R missense mutations and 24 A1708E missense mutations have been reported in the BIC database (16). Both mutations affect evolutionarily conserved, strongly hydrophobic residues in humans, dogs, rats, and mice.

Williams and Glover (27) assessed the structural response of the BRCT domain to M1775R using X-ray crystallography. The arginine residue was extruded from the shielded hydrophobic core, thereby perturbing the protein folding. This caused a disruption of hydrogen bonding and charge-charge repulsion at the surface of the protein and between the two BRCT repeats, resulting in conformational instability of the repeats (27).

The effects of A1708E on the three-dimensional structure of the BRCT domain are unknown. Both M1775R and A1708E can impair the DNA double-strand repair function of BRCA-1 and block the ability of the BRCT domain to interact with CtIP, histone deacetylases, and BACH1. They also interfere with transcriptional regulation (14). Based on the results of this article, further studies are warranted on the other conserved, hydrophobic residues that are affected by mutations (as shown in Table 3). Because the two examples, M1775R and A1708E, fit the pattern described in this article, it is important to determine how many other missense mutations that fit this pattern also predispose to cancer. This information would be of critical importance to genetic counselors.

Note Added in Proof. The bovine and chimpanzee BRCT amino acid sequences have recently become available. Incorporation of these into the statistical models results in a χ^2 of 49.69 for the first hypothesis and 76.99 for the second. P<0.001 in both cases.

References

1. Olopade O, Fackenthal J. Breast cancer genetics: implications for clinical practice. Hematol Oncol Clin North Am 2000;14:705-25.
2. Zheng L, Li S, Boyer T, Lee W-H. Lessons learned from BRCA-1 and BRCA-2. Oncogene 2000;19:6159-75.
3. Ford D, Easton DF, Stratton M, et al. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. Am J Hum Genet 1998;62:676-89.
4. Zhang X, Morera S, Bates P, et al. Structure of an XRCC1 BRCT domain: a new protein-protein interaction module. EMBO J 1998;17: 6404-11.
5. Wu-Baer F, Baer R. Effect of DNA damage on a BRCA-1 complex. Nature 2001;414:36.
6. Li S, Ting NS, Zheng L, et al. Functional Link of BRCA-1 and ataxia telangiectasia gene product in DNA damage response. Nature 2000; 406:210-5.
7. Lee JS, Collins KM, Brown AL, Lee CH, Chung JH. hCds1-mediated phosphorylation of BRCA-1 regulates the DNA damage response. Nature 2000;404:201-4.
8. Cortez D, Wang Y, Qin J, Elledge SJ. Requirement of ATM-dependent phosphorylation of BRCA-1 in the DNA damage response to double strand breaks. Science 1999;286:1162-6.
9. Hartman AR, Ford JM. BRCA-1 induces DNA damage recognition factors and enhances nucleotide excision repair. Nat Genet 2002;32: 180-4.
10. Paull TT, Cortez D, Bowers B, Elledge SJ, Gellert M. Direct DNA binding by BRCA-1. Proc Natl Acad Sci USA 2001;98:6086-91.
11. Yamane K, Katayama E, Tsuruo T. The BRCT regions of tumor suppressor BRCA-1 and of XRCC-1 show DNA end binding activity with a multimerizing feature. Biochem Biophys Res Commun 2000; 279:678-84.
12. Xu X, Qiao W, Linke SP, et al. Genetic interactions between tumor suppressors BRCA-1 and p53 in apoptosis, cell cycle and tumorigenesis. Nat Genet 2001;28:266-71.
13. Welcsh PL, Lee MK, Gonzalez-Hernandez RM, et al. BRCA-1 transcriptionally regulates genes involved in breast tumorigenesis. Proc Natl Acad Sci USA 2002;99:7560-5.
14. Williams RS, Green R, Glover JNM. Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. Nat Struct Biol 2001;8:838-42.
15. Orelli B, Logsdon J Jr, Bishop D. Nine novel conserved motifs in BRCA-1 identified by the chicken orthologue. Oncogene 2001;20: 4433-8.
16. Breast Cancer Information Core [accessed 2002 Aug]. Available from: http://research.nhgri.nih.gov/bic/.
17. Szabo C, Masiello A, Ryan J, The BIC Consortium, Brody L. The Breast Cancer Information Core: database design, structure, and scope. Hum Mutat 2000;16:123-31.

18. Eisenberg D, Schwarz E, Komaromy M, Wall R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. J Mol Biol 1984;179:125-42.
19. Nozaki Y, Tanford C. The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. J Biol Chem 1971; 246:2211-7.
20. Janin J. Surface and inside volumes in globular proteins. Nature 1979;277:491-2.
21. Chothia C. The nature of accessible and buried surfaces in proteins. J Mol Biol 1976;105:1-14.
22. Wolfenden R, Andersson L, Cullis PM, Southgate CCB. Affinities of amino acid side chains for solvent water. Biochemistry 1981;20: 849-55.
23. von Heijne G, Blomberg C. Trans-membrane translocation of proteins: the direct transfer model. Eur J Biochem 1979;97:175-81.
24. Gaboriaud C, Bissery V, Benchetrit T, Mornon JP. Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. FEBS Lett 1987;224:149-55.
25. Kramer M. Clinical epidemiology and biostatistics: a primer for clinical investigators and decision-makers. New York: Springer-Verlag; 1988.
26. Hayes F, Cayanan C, Barilla D, Monteiro A. Functional assay for BRCA-1: mutagenesis of the COOH-terminal region reveals critical residues for transcription activation. Cancer Res 2000;60:2411-8.
27. Williams S, Glover M. Structural consequences of a cancer-causing BRCA-1 BRCT missense mutation. J Biol Chem 2003;278:2630-5.

Missense Mutations in the BRCT Domain of BRCA-1 from High-Risk Women Frequently Perturb Strongly Hydrophobic Amino Acids Conserved among Mammals

Meaghan A. Figge and Lynda Blankenship

Cancer Epidemiol Biomarkers Prev 2004;13:1037-1041.