ARTICLE

Pathogenicity of the *BRCA1* missense variant M1775K is determined by the disruption of the BRCT phosphopeptide-binding pocket: a multi-modal approach

Marc Tischkowitz[1,2], Nancy Hamel[1,3], Marcelo A Carvalho[4,5], Gabriel Birrane[6], Aditi Soni[6], Erik H van Beers[7], Simon A Joosse[7], Nora Wong[1,2], David Novak[1,3], Louise A Quenneville[8], Scott A Grist[9], kConFab[10], Petra M Nederlof[7], David E Goldgar[11], Sean V Tavtigian[12], Alvaro N Monteiro[4], John AA Ladias[6] and William D Foulkes*,[1,2,3]

[1]*Program in Cancer Genetics, Departments of Oncology, Human Genetics and Medicine, McGill University, Montréal, Quebec, Canada;* [2]*Department of Oncology, Segal Cancer Centre, Sir M.B. Davis-Jewish General Hospital, Montréal, Quebec, Canada;* [3]*Department of Medicine, The Research Institute, McGill University Health Centre, Montréal, Quebec, Canada;* [4]*Risk Assessment, Detection, and Intervention Program, H. Lee Moffitt Cancer Center & Research Institute, Tampa, FL, USA;* [5]*Centro Federal de Educação Tecnológica de Química, Rio de Janeiro, Brazil;* [6]*Department of Medicine, Molecular Medicine Laboratory and Macromolecular Crystallography Unit, Division of Experimental Medicine, Harvard Institutes of Medicine, Harvard Medical School, Boston, MA, USA;* [7]*Division of Experimental Therapy and Department of Pathology, Netherlands Cancer Institute, Amsterdam, The Netherlands;* [8]*Department of Pathology, Sir M.B. Davis-Jewish General Hospital, McGill University, Montréal, Quebec, Canada;* [9]*Department of Hematology & Genetic Pathology, Flinders Medical Centre, Flinders University of South Australia, Adelaide, South Australia, Australia;* [10]*Kathleen Cuningham Consortium for Research into Familial Breast Cancer, Peter MacCallum Cancer Institute, East Melbourne, Victoria, Australia;* [11]*Department of Dermatology, University of Utah School of Medicine, Salt Lake City, UT, USA;* [12]*International Agency for Research on Cancer, World Health Organization, Lyon, France*

A number of germ-line mutations in the *BRCA1* gene confer susceptibility to breast and ovarian cancer. However, it remains difficult to determine whether many single amino-acid (missense) changes in the BRCA1 protein that are frequently detected in the clinical setting are pathologic or not. Here, we used a combination of functional, crystallographic, biophysical, molecular and evolutionary techniques, and classical genetic segregation analysis to demonstrate that the *BRCA1* missense variant M1775K is pathogenic. Functional assays in yeast and mammalian cells showed that the BRCA1 BRCT domains carrying the amino-acid change M1775K displayed markedly reduced transcriptional activity, indicating that this variant represents a deleterious mutation. Importantly, the M1775K mutation disrupted the phosphopeptide-binding pocket of the BRCA1 BRCT domains, thereby inhibiting the BRCA1 interaction with the proteins BRIP1 and CtIP, which are involved in DNA damage-induced checkpoint control. These results indicate that the integrity of the BRCT phosphopeptide-binding pocket is critical for the tumor suppression function of BRCA1. Moreover, this study demonstrates that multiple lines of evidence obtained from a combination of functional, structural, molecular and evolutionary techniques, and

classical genetic segregation analysis are required to confirm the pathogenicity of rare variants of disease-susceptibility genes and obtain important insights into the underlying pathogenetic mechanisms.

Introduction

The breast cancer susceptibility gene, *BRCA1*, is one of the most studied genes in cancer research but despite the abundance of available data, new variants of unknown significance are regularly detected in clinical practice. These variants create serious management problems in the families concerned and are often frustrating to deal with. To date, there is no comprehensive functional assay available for *BRCA1* mutations and much has been written about the approaches that can be used to classify variants, summarized in Goldgar *et al*.[1] In general, it is easier to conclude that a variant is non-pathogenic than pathogenic, a point illustrated by Chenevix-Trench *et al*[2] who studied 10 *BRCA1* variants by a combination of methods and identified one as being definitely pathogenic.

In the present study, we investigated the rare variant M1775K, which occurs in the BRCT domains at the C terminus of BRCA1. The particular functional significance of the BRCA1 BRCT repeats has become increasingly recognized,[3] especially because they mediate interactions with proteins involved in cell cycle checkpoint control and double-stranded DNA repair, including BRIP1, a DNA helicase previously known as BACH1[4] that is also a breast cancer susceptibility protein,[5] and the co-repressor CtIP.[6] The M1775K variant was identified in two unrelated families of European ancestry with a history of breast cancer but its contribution to the pathogenesis of this disease has not been determined. Here, we used a combined approach encompassing a number of scientific disciplines to demonstrate that M1775K is pathogenic. Specifically, the M1775K mutation disrupts the phospho-peptide-binding pocket of the BRCA1 BRCT domains, thereby inhibiting the BRCA1 interaction with the proteins BRIP1 and CtIP. These results indicate that the integrity of the BRCT phosphopeptide-binding pocket is critical for the tumor suppression function of BRCA1. It is important to emphasize the need for a multi-disciplinary approach to determine pathogenicity, and this study, while focused on only one variant, argues for an in-depth characterization of all unclassified variants, particularly when their individual frequency is very low.

Methods
Samples
The M1775K variant was identified by sequencing in two unrelated families who had presented to cancer genetics services with a history of breast cancer and had undergone routine *BRCA1/BRCA2* mutation analysis. Both families consented to further study of the variant. Blood samples were taken from additional family members to enable segregation analysis and relevant tumor blocks were obtained. Immunohistochemical analysis for the estrogen, progesterone and HER2 receptors was performed using standard protocols.

Prior probability of pathogenicity from evolutionary conservation and substitution severity analysis
M1775K was subjected to Align-GVGD and SIFT analysis using a full-length BRCA1 protein multiple sequence alignment containing nine mammalian sequences plus sequences from chicken, frog and pufferfish; the alignment is available at http://agvgd.iarc.fr/alignments.php. The Grantham variation (GV), Grantham deviation (GD) and SIFT scores were calculated from the multiple sequence alignment and the observed missense substitution.[7,8] The combined GV–GD score was converted to a prior probability for classification as a high-risk variant based on a heterogeneity analysis of the family histories associated with 1433 variants in the Myriad Genetics Laboratory BRACAnalysis database.[9]

Incorporation of histopathology information
This was incorporated into the model using the approach of Chenevix-Trench *et al*[2] in which tumors are classified by ER status and grade and then given a likelihood ratio score based on the data of Lakhani *et al*,[10] who compared a large series of *BRCA1*-related tumors and controls for a variety of histopathological and immunological features. For example, a grade 3 ER− tumor is ~3 times more likely to be a *BRCA1*-related tumor than an age-matched sporadic tumor; conversely, a grade 1 ER+ tumor yields odds of 20:1 against being a *BRCA1*-related tumor. When one factor was missing, the marginal probabilities were used. The overall score for histopathology is the product of the likelihood ratios for each tumor evaluated carrying the M1775K variant. The Posterior probability is calculated from the prior probability (based on sequence data) and the likelihood ratio for causality (from histopathology and co-segregation) using Bayes rule, posterior = prior × LR/(prior × LR + 1 − prior).

LOH analysis

Tumor tissue from one of the affected *BRCA1* M1775K carriers was both macro- and micro-dissected (using laser capture microdissection) from formalin-fixed, paraffin-embedded (FFPE) tissue, and DNA was extracted from the collected cells using the QIAamp DNA Mini Kit (Qiagen, Mississauga, Ontario, Canada) according to the manufacturer's instructions for FFPE samples. Three microsatellite markers within *BRCA1* (D17S855, D17S1322 and D17S1323) were genotyped using radioactively labeled PCR products from DNA isolated from blood and tumor tissue from our carrier using the QIAGEN HotStar Taq PCR system (Qiagen) (primer sequences and annealing temperatures are listed in Supplementary Table 1). Products were separated by electrophoresis in a 6% denaturing acrylamide gel for approximately 2 h at 70 W and then autoradiographed. The relative intensity of the two alleles at each locus was compared and used to establish the presence of LOH at these loci. Additional primers flanking the variant were designed using the Primer3 software (Whitehead Institute for Biomedical Research, Cambridge, MA, USA; sequences and annealing temperature are listed in Supplementary Table 1). The resulting PCR products were sequenced directly using 3730XL DNA Analyzer Systems from Applied Biosystems (Foster City, CA) in both blood and tumor tissue. Chromatograms were viewed using Chromas 2.31 from Technelysium Pty Ltd. (Helensvale, Australia) and the relative intensities of the peaks in the normal and mutant traces of each sample were visually compared to determine which *BRCA1* allele was lost in the tumor sample.

Comparative genomic hybridization

DNA was isolated from 10 10-μm-thick FFPE sections, after which the DNA quality was tested by a multiplex PCR.[11] Comparative genomic hybridization (CGH) was performed on microarrays containing 3.5 K BAC/PAC-derived DNA segments covering the whole genome with an average spacing of 1 Mb, as previously described.[12] Data processing of the scanned microarray slide was performed by signal intensity measurement in ImaGene Software (BioDiscovery Inc., CA, USA) followed by median pin tip (ie subarray) normalization. Intensity ratios (Cy5/Cy3) were log2 transformed and triplicate spot measurements were averaged. Estimation of fragment copy numbers was carried out using CGH segmentation algorithm.[13]

To classify the tumor as either *BRCA1*-like or sporadic-like, a 'shrunken centroids classifier' was used.[14] The class predictor was built on 18 *BRCA1*-related and 32 sporadic tumors, and validated on an independent set of 10 *BRCA1*-related and 16 sporadic tumors.[15] On the basis of leave-one-out cross-validation, 191 features were selected to be discriminatory, with 3q and 5q the most prominent regions. Application of the shrunken centroids classifier generated a classifier score and class label (score ≥ 0 is

assessed as *BRCA1*-like, score <0 is assessed as sporadic-like). As reference interval for the classifier scores (for non-normally distributed data) the 95% interval is taken. The classifier scores for the validation set of the *BRCA1*-related tumors were in 95% of the cases above 1.6, for the validation set of the sporadic tumors, 95% of the scores were below -0.05. The reference interval for the classifier scores of *BRCA1*-related and sporadic tumors does not overlap.

Functional assays

Constructs The yeast expressing vector pLex9 carrying a wild-type *BRCA1* sequence (aa 1396–1863) fused in frame to the LexA DNA-binding domain (DBD) was used as wild-type control and as a backbone to introduce the mutations by site-directed mutagenesis using the following methods. Control constructs containing the wild-type *BRCA1*, S1613G, M1775R and Y1853X were previously described.[16] Variant M1775K was introduced by splicing using overlapping extension PCR[17] with p385-*BRCA1* (gift from Michael Erdos) as template. The first round of PCR was performed using the following primers: M1775K 3′ region (UMK, 5′-CCTTCACCAACAAGCCCACAGGATCAACTG-3′; 24ENDT, 5′-GCGGATCCTCAGTAGTGGCTGTGGGGGAT-3′); M1775K 5′ region (M1775K-R, 5′-CAGTTGATCTGTGG GCTTGTTGGTGAAGG-3′; UX13, 5′-CGGAATTCCAGAGG GATACCATGCAA-3′). Both products (5′ and 3′ regions) were combined and used as a template for a final round of PCR using the primers 24ENDT and UX13. The final PCR products were cloned into pPCR-Script Amp SK(+) vector (Stratagene). The inserts were then isolated by cutting with *Bam*HI and *Eco*RI and ligated into pLex9 or pGBT9 vectors. All mutations were confirmed by sequencing. To obtain GAL4 DBD fusions in a mammalian expression vector, pGTB9 constructs were digested with *Hin*dIII and *Bam*HI, a 1.8-kb band was isolated and ligated into equally digested pCDNA3.

Transcription assay in yeast and in mammalian cells The transcriptional assays in yeast and mammalian cells were performed essentially as previously described.[16,18,19] Briefly, *Saccharomyces cerevisiae* strain EGY48 was co-transformed with the effector plasmid pLex9 containing a fusion of LexA DBD and BRCA1 aa 1396–1863 with different variants and the plasmid reporter pRB1840, which contains a lacZ gene under the control of one LexA operator.[20,21] Competent yeast cells were obtained using the yeast transformation system based on lithium acetate (Clontech) and cells were transformed according to the manufacturers' instructions. At least three individual clones for each variant were tested for liquid β-galactosidase assays using ONPG,[22] and the assays were performed in triplicate. Activity was determined as a comparison to wild-type BRCA1 and S1613G (positive controls) or to M1775R, and Y1853X (negative controls).

For mammalian cell assays, we used pG5Luc as the reporter and transfections were normalized with an internal control phGR-TK (Promega), which contains a *Renilla* luciferase gene under a constitutive TK basal promoter. Transfections were performed with human 293T cells in triplicate using Fugene 6 (Roche), cells were harvested 24 h post-transfection, and luciferase activity was measured using a dual luciferase assay system (Promega).

Western blot analysis was performed as previously described.[3] The blots were incubated with Clontech's α-GAL4 DBD monoclonal antibody (for mammalian cells) or α-LexA DBD monoclonal antibody (for yeast cells).

Structural and biophysical analysis

Protein purification and crystallization A DNA fragment encoding the *BRCA1* BRCT domains (residues 1649–1858) harboring the variant M1775K and carrying an N-terminal 3C protease site was amplified by PCR and cloned into a modified pMAL-c2x vector (NEB). BRCA1 BRCT M1775K was expressed as a fusion with hexahistidine-tagged maltose-binding protein (6H-MBP) in *Escherichia coli* C41(DE3) cells, purified on Ni-NTA resin and eluted with 200 mM imidazole. The eluant was dialyzed against a buffer containing 25 mM Tris-HCl, pH 7.0, 300 mM NaCl, 2 mM DTT and concentrated by ultrafiltration. Cleavage of the BRCT from its 6H-MBP fusion partner was carried out with 3C protease overnight at 4°C. The protein was further purified on a Superdex 75 gel filtration column (GE Healthcare) and fractions containing the BRCT were pooled and concentrated to 23 mg/ml. Crystals were grown using a well solution containing 1.65–1.8 M ammonium sulfate, 20 mM cobalt chloride and 100 mM MES, pH 6.5, cryoprotected in mother liquor supplemented with 25% glycerol, and flash frozen in a liquid nitrogen stream. Data were collected on beamline X12B at the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory, Upton, NY, USA. The crystals belong to space group $P6_122$ with unit cell dimensions $a = b = 114.3$ Å, $c = 119.9$ Å, $\alpha = \beta = 90°$, $\gamma = 120°$. The data were processed and scaled using HKL2000[23] (Table 1).

Structure determination and refinement The structure was solved by molecular replacement with MOLREP,[24] using the BRCA1 BRCT M1775R structure (PDB entry 1N50) as the search model but omitting the side chain of Arg1775 to avoid model bias. The atomic model was built with COOT,[25] and refined to a crystallographic factor R_{cryst} of 25.1 and an R_{free} of 30.2 using TLS parameters and an overall B-factor in REFMAC5.[26] Hydrogen atoms were included at their riding positions. Because at 3.6-Å resolution many of the side chains may not be visible in the electron density, Lys1775 was removed, random shifts of 0.2 Å were introduced in the coordinates of all the atoms, and the resulting model was subjected to 10 cycles of restrained refinement. The obtained σ_A-weighted omit

Table 1 Structure determination and refinement statistics

Resolution range (Å)	50–3.6 (3.73–3.6)
Wavelength (Å)	1.000
Observed reflections	158 139
Unique reflections	5744
Completeness (%)[a]	99.9 (100)
Redundancy[a]	27.5 (28.9)
R_{sym} (%)[b]	9.3 (53.0)
Overall $\langle I/\sigma(I) \rangle$	36.1 (7.1)
R_{cryst} (%)[c]	25.1
R_{free} (%)[d]	30.2
Ramachandran plot	
Most favored (%)	72.2
Additionally allowed (%)	22.8
Generously allowed (%)	3.3
Disallowed (%)	1.7
Bond lengths[e] (Å)	0.014
Bond angles[e] (degrees)	1.635

[a]Values in parentheses are for the highest resolution shell.
[b]$R_{sym} = \sum |(I - \langle I \rangle)|/\sum(I)$, where I is the observed integrated intensity, $\langle I \rangle$ is the average integrated intensity obtained from multiple measurements and the summation is over all observed reflections.
[c]$R_{cryst} = \sum \|F_{obs}| - k|F_{calc}\|/\sum|F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively.
[d]R_{free} is calculated as R_{cryst} using 8.8% of the reflections chosen randomly and omitted from the refinement calculations.
[e]Bond lengths and angles are root-mean-square deviations from ideal values.

map was used to model the Lys1775 side chain. In addition, all possible orientations of the Lys1775 side chain were investigated employing the standard rotamer library that is implemented in the program COOT. The modeled orientation is the only one which does not project the charged Lys1775 side chain into an adjacent hydrophobic region (formed by Leu1701, Met1783, Val1838 and Leu1839) or clashes with neighboring amino-acid residues (Gln1779, Met1783 and Arg1835). The final model contains one monomer per asymmetric unit with 3340 protein atoms, two sulfate ions and one cobalt ion. The N-terminal vector-derived residues GP were not visible in the final electron density map. The residues Gln1756, Arg1762 and Thr1802 have poor stereochemistry and appear in the disallowed region of the Ramachandran plot.

Isothermal titration calorimetry Binding affinities of the BRCT M1775K protein for the BRIP1 and CtIP phosphopeptides were measured using a VP-ITC microcalorimeter (MicroCal). Briefly, 0.03 mM of BRIP1 and 0.131 mM of CtIP peptides were titrated against 0.006 and 0.008 mM BRCT M1775K protein, respectively, in phosphate-buffered saline, 300 mM NaCl, pH 7.0, at 25°C. As controls, 0.159 mM of BRIP1 and 0.191 mM of CtIP were titrated against 0.021 and 0.016 mM wild-type BRCA1 BRCT protein, respectively, under the same conditions. Titration curves were analyzed using ORIGIN 5.0 (OriginLab). Precise protein and peptide concentrations were

determined by quantitative amino-acid analysis on an ABI 420A derivatizer/analyser and an ABI 130A separation system (Applied Biosystems).

Coordinates The atomic coordinates and structure factors of BRCA1 BRCT M1775K have been deposited in the Protein Data Bank (accession code 2ING).

Results

The pedigrees of the two families in which the *BRCA1* M1775K variant was identified are shown in Figure 1 and the tumor characteristics and available genotypes are listed in Table 2. In family A, which is of Western European ancestry, the proband presented with a history of bilateral breast cancer at the ages of 43 and 44 years. The M1775K
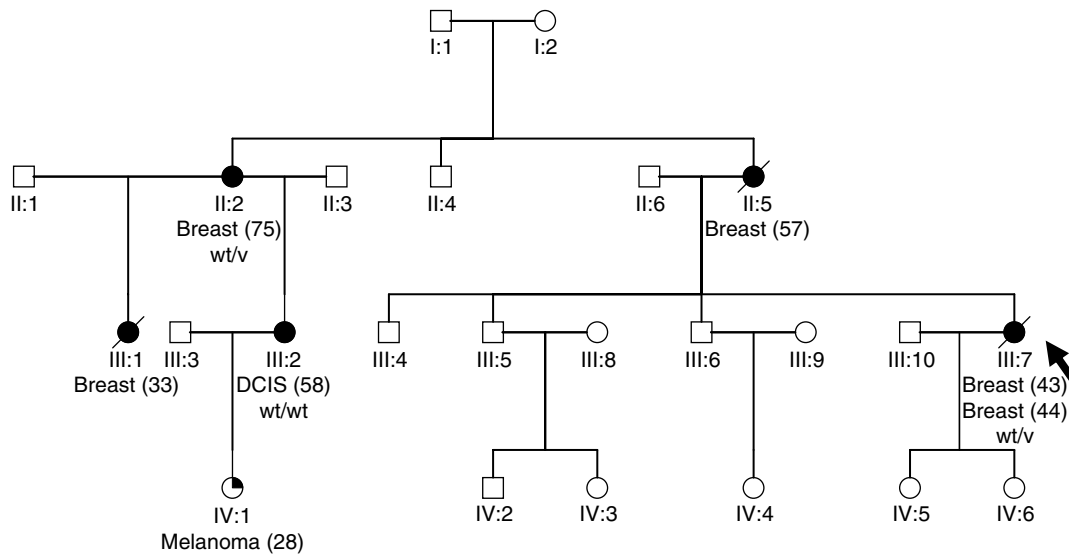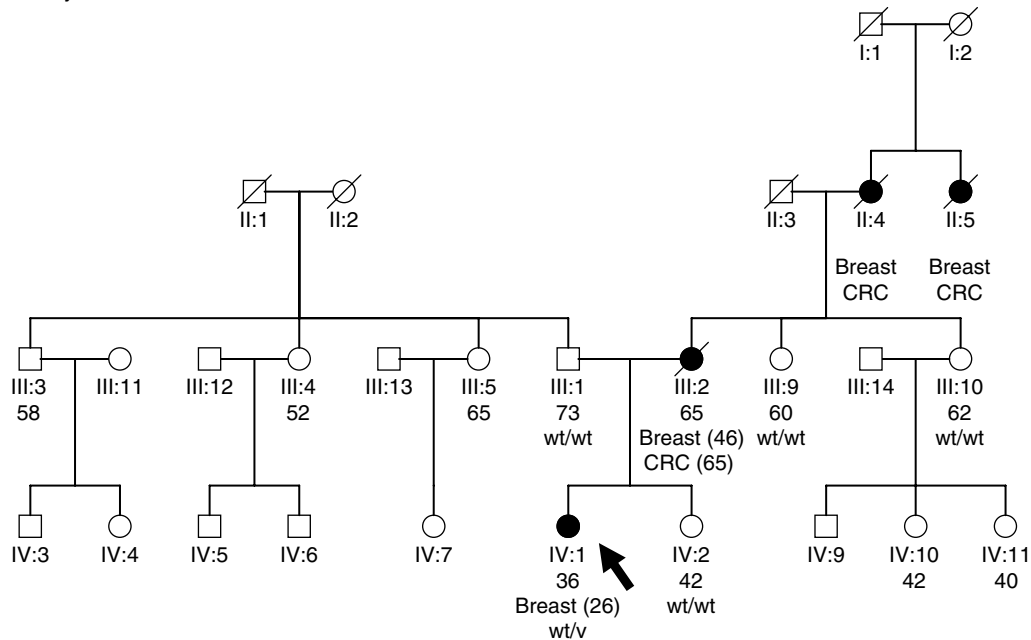

Figure 1 Pedigrees of the two families with M1775K variants. Mutation status is indicated by wt, wild-type; v, *BRCA1* M1775K variant. Blacked-in shapes indicate a cancer diagnosis.

Table 2 Histopathological characteristics of the breast cancers and genotype status of affected and unaffected members in the two families

Case	Diagnosis	Age (years)	Grade	Receptors	Genotype
A:III.7	IDC	43	3/3		wt/v
	IDC	44			
A:III.1	Medullary	31		ER?, PR+	Not done
A:II.2	IDC/medullary	75	3/3	ER−,PR−	wt/v
A:II.5	IDC	57		ER−, PR−, HER2+ (3+)	wt/v obligate
A:III.2	High-grade DCIS	58			wt/wt
B:III.2	IDC	46	2/3	ER−, PR−	wt/v[a]
B:IV.1	Medullary	26	3/3	ER−	wt/v
B:III.9	Unaffected	60			wt/wt
B:III.10	Unaffected	62			wt/wt
B:IV.2	Unaffected	42			wt/wt

v, *BRCA1* M1775K variant; wt, wild type.
[a]Assumed as BIII.1 is wt/wt.

variant was identified in her and on further testing it was also present in her affected aunt, which made her affected deceased mother an obligate carrier. One of the daughters of AII.2 had medullary-type breast cancer at 33 years (genotype unknown) and a second daughter, AIII.2, had DCIS at 58 years. This daughter did not carry the M1775K variant. In family B, which is of Greek ancestry, the proband was diagnosed at age 26 years with Stage 1 breast cancer. Pathology revealed a high-grade, 1 cm infiltrating medullary carcinoma, which had negative estrogen, progesterone and HER2 receptor status. Genetic testing detected the M1775K variant in her and in addition she had a D1778N variant 3 codons downstream, which was not present in family A, indicating that the families are unrelated. The proband's father had no personal or family history of cancer and testing showed that he did not carry the M1775K or the D1778N variant, indicating that the variants were syntenic. On the basis of this, we assumed that the variant had been inherited from her mother, an assumption strengthened by the fact that she was diagnosed with Stage 2 breast cancer at age 46 years and metastatic colon cancer at age 65 years (both cancers were confirmed on pathology). The proband reported that both her maternal grandmother and her great aunt had breast and colon cancer in their 50s (unconfirmed). Two maternal aunts, unaffected at ages 61 and 69 years were both tested and found not to carry the M1775K variant.

Prior probability based on evolutionary conservation and severity of the missense substitution

Position M1775, located in the C-terminal BRCT domain, is invariant in a BRCA1 multiple sequence alignment containing sequences from nine mammals plus chicken, frog and pufferfish (ie GV = 0.00). The GD for the substitution is 94.5, and the SIFT score is 0.0.[7,8] The prior probability that *BRCA1* or *BRCA2* missense substitutions with GV = 0 (scored with respect to the alignment through pufferfish) and GD > 60 are clinically important high-risk variants is 0.60.[9] With addition of further diverged

sequences to the alignment (sea urchin, tunicate, *Caenorhabditis elegans*), some variability is observed at this position (leucine in sea urchin and phenylalanine in tunicate and *C. elegans*), raising GV to 30.3; however, GD remains high (93.8), reinforcing evidence that lysine is outside of the range of variation tolerable at position M1775.

Likelihood from segregation analysis and histopathology

We assumed that A:III.2 who had high-grade DCIS was affected but assigned her an older age group; DCIS and breast cancer diagnosed at greater than 60 years have had similar odds ratios in previous logistic regression analyses.[27] The likelihood ratio for co-segregation in family A was 1:1 in favor of the variant being deleterious for family A and 5:1 for family B. Thus, the overall odds from segregation analysis are 5:1 as family A does not contribute. Histopathological data, using ER status and grade where available, contribute odds in favor of the variant being deleterious of 77:1. Combined together, these data result in odds in favor of M1775K being deleterious *versus* neutral of 385:1. If A:III.2 is considered as affected at the age of 58 years, the corresponding figure is 235:1 and if considered as unaffected, the odds would be 1675:1. Combining these likelihoods with the prior probability, the ranges of posterior probabilities that M1775K is a high-risk variant varies between 500:1 (A:III.2 excluded), 333:1 (A:III.2 affected) and 2500:1 (A:III.2 unaffected). The LOH data showed loss of the wild-type (wt) allele, which provided additional supportive evidence for a deleterious variant.

CGH analysis

CGH analysis of the *BRCA1* M1775K tumor from B:IV.1 using a 1 Mb BAC array demonstrated many genomic aberrations, among which were high-level amplifications (Figure 2a). The characteristic *BRCA1* aberrations, gain on chromosome 3q and loss on 5q,[28,29] are clearly present. Using the Shrunken Centroids Classifier,[28] the M1775K
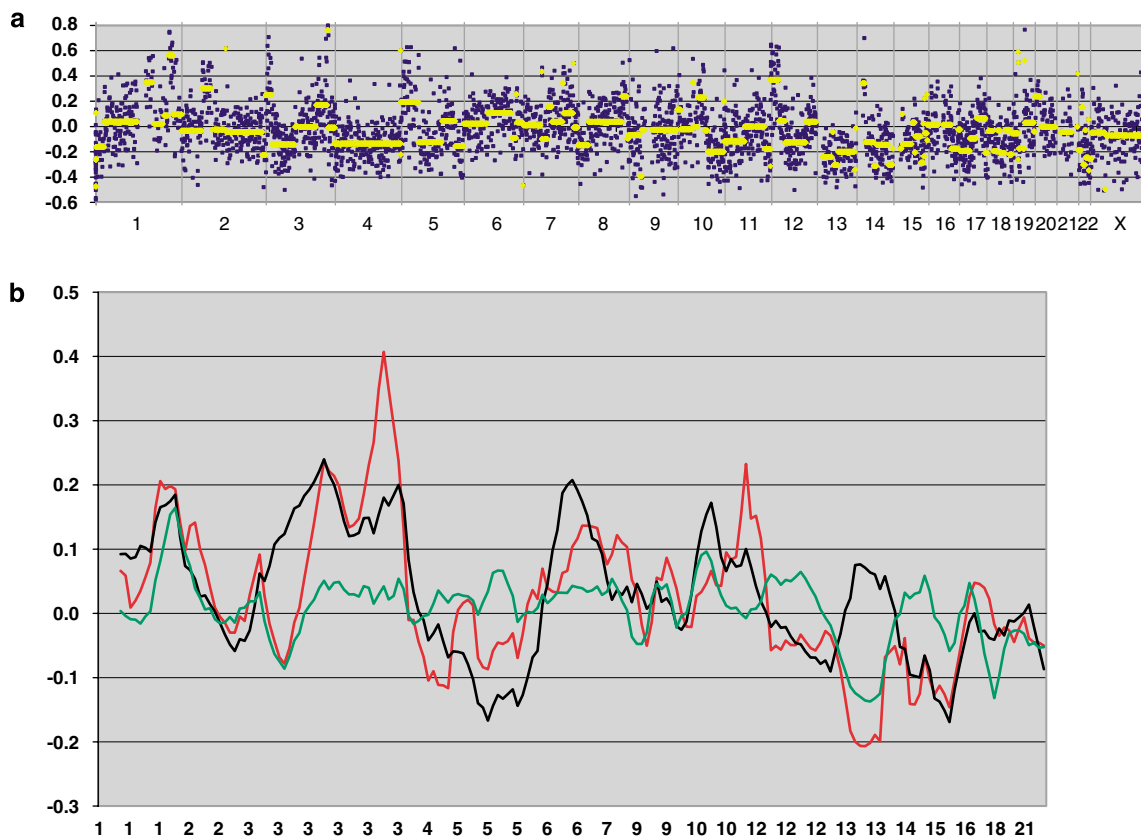
Figure 2 (**a**) Array CGH profile of M1775K from individual B:IV.1 showing the genomic position (*x* axis) and fluorescent ratio for all 3277 BACs (blue dots) and estimated copy number levels as determined by CGH segmentation software (yellow line).[13] (**b**) Average fluorescence ratio for 28 *BRCA1*-related tumors (black line) and 49 sporadic tumors (green line) and M1775K tumor from individual B:IV:1 (red line, moving average) for the 191 BAC clones in the *BRCA1* classifier.

tumor scored 1.35, which classifies it as *BRCA1*-like. In Figure 2b, the average ratios for *BRCA1*-related and sporadic tumors are depicted with the fluorescence ratios of the M1775K tumor for the 191 BACs included in the classifier. It shows that for most clones the tumor more closely resembles the aberrations present in the *BRCA1*-related tumors.

Functional analysis

To assess the functional impact of the changes in BRCA1, we determined the transcriptional activity of the C-terminal region of BRCA1 in which the variant was introduced. The M1775K missense variant was evaluated for transactivation activity in the context of stringent reporters (Figure 3) and displayed markedly reduced activity with 20 and <5% of the wild-type activity in yeast and mammalian cells, respectively (Figure 3a and b). In both cases, the activity was comparable to the deleterious variant M1775R, which has been previously described.[30,31] Although protein levels in yeast cells for the M1775K variant were lower than the wild type, they were similar to

the positive control S1613G, a neutral variant, suggesting that the reduced activity by the M1775K variant cannot be accounted for by differences in expression levels (Figure 3c). In addition, protein levels in mammalian cells were similar for all constructs tested (Figure 3d). Because an analysis of a large series of variants validated by genetic data indicated that 50% of the wt activity can be considered the threshold for classification into deleterious or neutral,[16] these functional results indicate that M1775K is likely to represent a deleterious change.

M1775K abrogates BRCA1 binding to BRIP1 and CtIP

It was previously shown that the BRCA1 BRCT domains bind to the BRIP1 phosphopeptide ISRSTpSPTFNKQ (pS denoting phosphoserine) that corresponds to residues 985–996 of human BRIP1 with a dissociation constant (K_d) of 0.7 μM, and to the CtIP phosphopeptide PTRVSpSPVFGAT (residues 322–333 of human CtIP) with a K_d of 3.7 μM.[31] To investigate the effects of the M1775K variant on the binding properties of BRCTs, we used isothermal titration calorimetry to measure the affinities
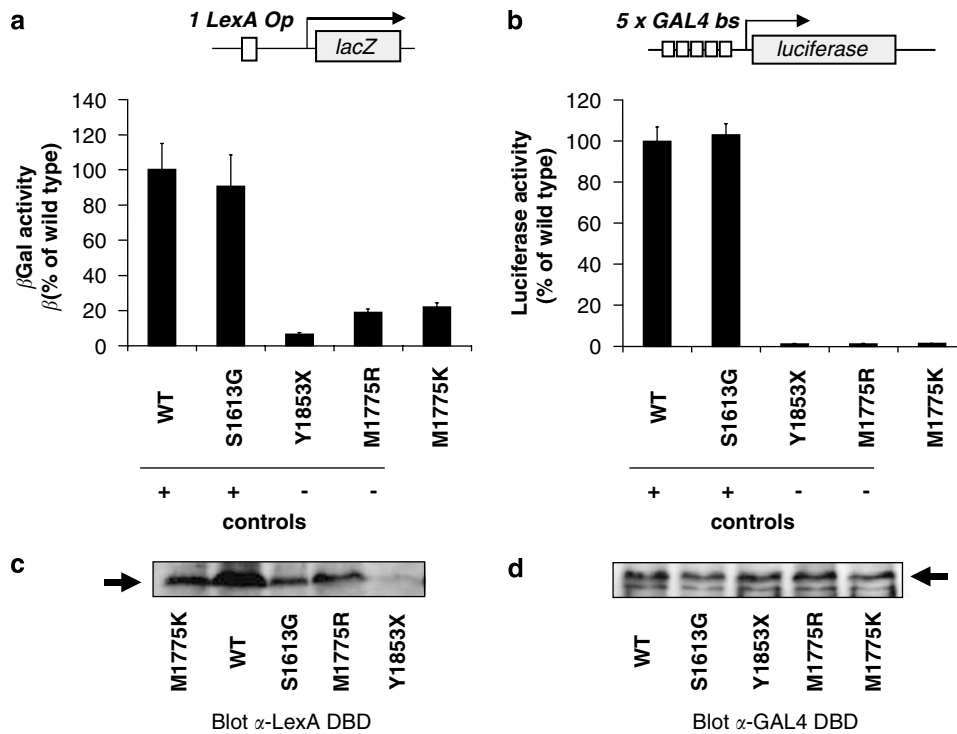
Figure 3 Functional analysis of M1775K in BRCA1. (**a**) Quantitative transcriptional assay in yeast cells. Cells were co-transformed with a LexA-responsive β-galactosidase reporter gene (diagram shown above the graph) and a LexA DBD fusion to residues 1396–1863 of wt BRCA1, or the same fragment carrying various UCVs. We used the wt and the S1613G neutral polymorphism as positive controls (+). Deleterious mutations M1775R and Y1853X were used as negative controls (−). Three independent yeast clones were tested in triplicate. The activity of the construct with wt BRCA1 was expressed as 100%, with the other results placed on this scale. (**b**) Quantitative transcriptional assay in mammalian cells. Cells were co-transfected with a GAL4-responsive firefly luciferase reporter gene (diagram shown above the graph), a *Renilla* luciferase driven by a constitutive promoter (internal control, not shown), and a GAL4 DBD fusion to residues 1396–1863 of wt BRCA1 (WT), or the same fragment carrying various UCVs. Controls are the same as described above but fused to GAL4 DBD. Measurements were carried out in triplicate and normalized against the internal transfection controls. The activity of the construct with wt *BRCA1* was expressed as 100%, with the other results placed on this scale. To control for possible variations in protein expression levels, samples were analyzed by western blot with rabbit anti-LexA DBD polyclonal antibody in yeast extracts (**c**) or mouse anti-GAL4 DBD monoclonal antibody in mammalian cell extracts (**d**).

of the BRCT M1775K protein for the aforementioned BRIP1 and CtIP phosphopeptides. Strikingly, these studies showed that BRCT M1775K had no measurable affinities for the BRIP1 and CtIP peptides (Figure 4a and b). Additional isothermal titration calorimetry experiments using increased concentrations of the protein (twofold) and peptides (fourfold) did not show any significant heat changes beyond those of dilution heat (data not shown), demonstrating that the BRCT M1775K variant does not bind to the BRIP1 and CtIP peptides. By contrast, both phosphopeptides interacted with wild-type BRCA1 BRCT protein (Figure 4c and d) with affinities identical to those reported previously.[31] Taken together, these results indicate that the Lys1775 side chain interferes with the BRCT–ligand interaction.

Structural basis for disruption of the BRCA1–BRIP1 and BRCA1–CtIP interactions by M1775K
To elucidate the mechanism by which Lys1775 interferes with ligand binding at the atomic level, we determined the

crystal structure of BRCA1 BRCT M1775K at 3.6-Å resolution. Each of the two BRCA1 BRCT domains comprises a central β-sheet formed by four parallel β-strands (β1–β4) and flanked by two α-helices (α1 and α3) on one side and a single α-helix (α2) on the other (Figure 5a). The two BRCT domains pack closely against each other in a head-to-tail manner burying a large hydrophobic interface and creating a deep surface groove. The CtIP and BRIP1 peptides bind to this groove in a two-pronged mode, with the phosphoserine at position 0 (pSer 0) binding to a shallow basic pocket (P_1) in BRCT1 and phenylalanine at position +3 (Phe +3) entering a hydrophobic pocket (P_2) made up of residues from both BRCT1 and BRCT2 repeats.[31,34,35] Superposition of the unbound mutant BRCT M1775K with the BRCA1–CtIP crystal structure[31] shows that the BRCT repeats are superimposed very well with a root-mean-square deviation of 0.85 Å for all C^α atoms (Figure 5a). Although at this resolution many of the side chains have poor electron density, an annealed σ_A-weighted omit map shows clear electron density for the side chain of Lys1775 (Figure 5b).
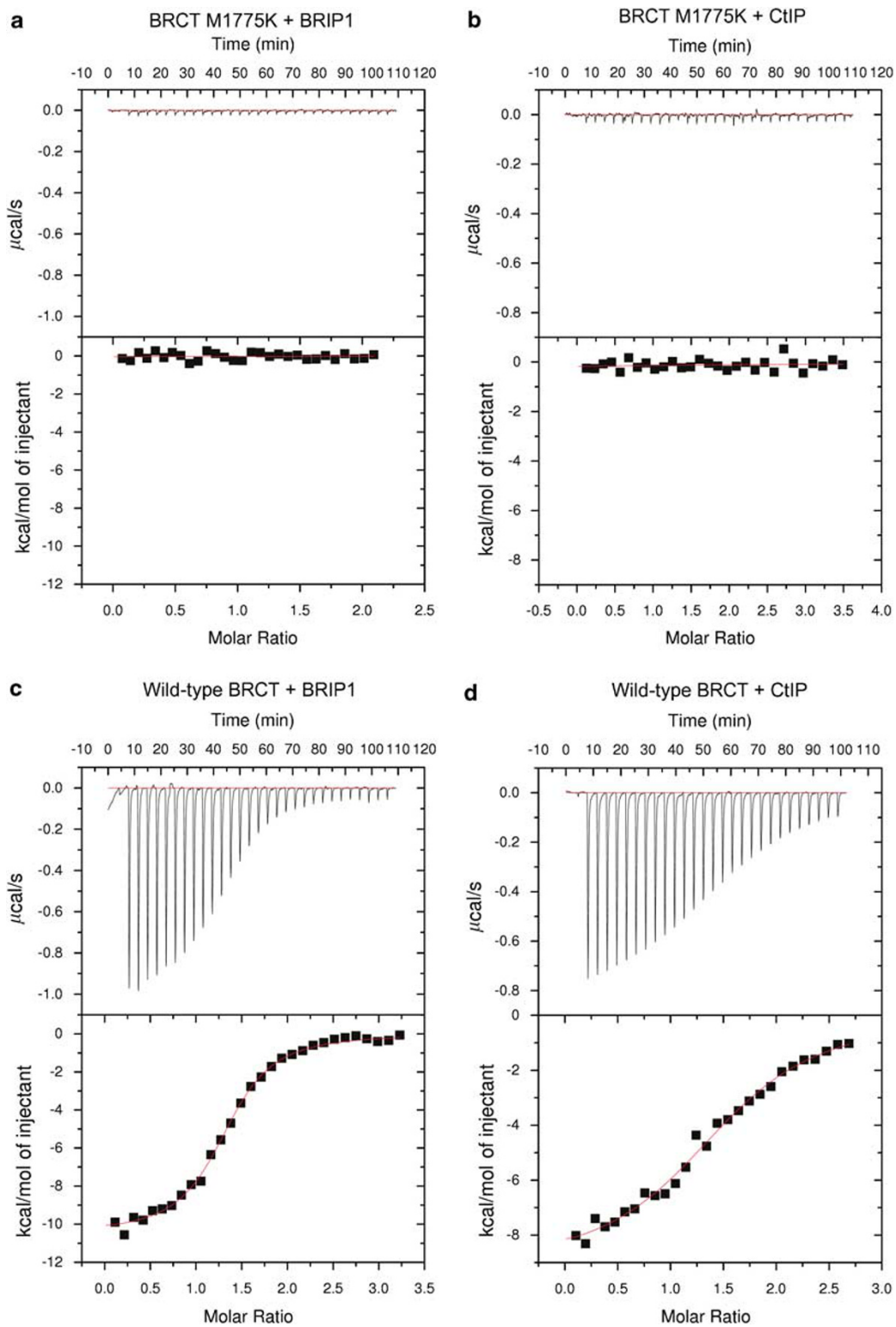
Figure 4 Representative isothermal titration calorimetry results obtained for the interaction of the BRIP1 and CtIP phosphopeptides with variant BRCA1 BRCT M1775K (**a** and **b**), respectively, and wild-type BRCA1 BRCT domains (**c** and **d**), respectively.
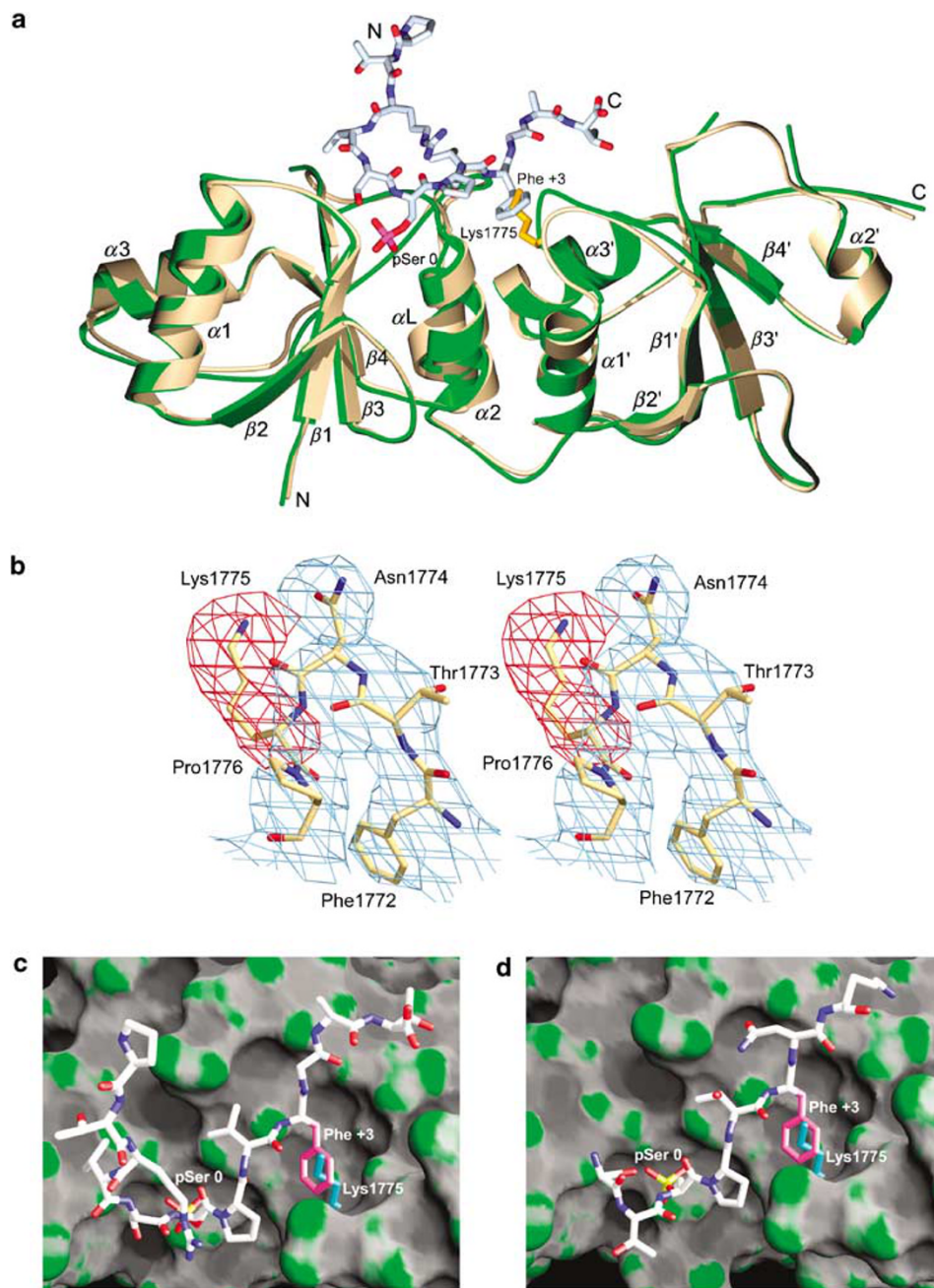
Figure 5 (**a**) Ribbons representation of the unbound BRCA1 BRCT M1775K structure (shown in green) superimposed to the BRCA1 BRCT-CtIP crystal structure (PDB 1Y98) (BRCT shown in beige and the peptide as light blue stick model). The α-helices and β-strands of BRCT1 are numbered, whereas the BRCT2 secondary structure elements are labeled with primes. The critical residues pSer 0 and Phe +3 are denoted. The side chain of Lys1775, shown as a yellow stick model, sterically clashes with the phenyl ring of Phe +3. (**b**) Stereo view of the electron density of Lys1775 and adjacent residues. An annealed σ_A-weighted omit map contoured at 3.5σ (red) is superimposed on a σ_A-weighted $2F_o-F_c$ map contoured at 1.0σ (cyan). Clear electron density is visible for the Lys1775 side chain. Carbon, nitrogen and oxygen atoms are colored beige, blue and red, respectively. (**c**) Occlusion of the CtIP Phe +3 by the side chain of Lys1775. Superposition of the BRCT–CtIP and BRCT M1775K crystal structures shows the molecular clash between the Lys1775 (cyan) and Phe +3 (pink) side chains. (**d**) Obstruction of the BRIP1 Phe +3 by variant M1775K. Superposition of the BRCT–BRIP1 (PDB 1T15) and BRCT M1775K crystal structures shows the steric hindrance between the Lys1775 (cyan) and Phe +3 (pink) side chains. (**a**) Figure was made using POVScript+ [32] and POV-Ray (www.povray.org). (**c, d**) Figure was made with GRASP. [33]

Strikingly, the side chain of the substituted Lys1775 sterically clashes with the phenyl ring of Phe +3 of CtIP (Figure 5a and c), directly obstructing the insertion of this anchoring group into the P$_2$ pocket. Likewise, superposition of the BRCT M1775K and BRCA1–BRIP1 [34,35] crystal structures reveals an identical steric hindrance between

Lys1775 and the phenyl group of the BRIP1 Phe $+3$ (Figure 5d).

Discussion

By using a combination of functional, structural, molecular and evolutionary techniques and classical genetic segregation analysis, we have demonstrated that the *BRCA1* M1775K variant is highly likely to be pathogenic. The combination of sequence analysis-based prior probability and segregation analysis-based likelihood ratio gave a strong indication that M1775K is pathogenic, but was not conclusive. Addition of the histopathology-based likelihood ratio yields a convincing posterior probability (greater than 100:1). However, histopathological characteristics have only recently been added to the integrated likelihood approach[2] and therefore we felt it important to confirm pathogenicity by other methods. In this respect, the LOH studies may add supporting evidence, although almost all studies where *BRCA1*-related tumors were shown to lose the wild-type allele[36–38] have been carried out on samples with truncating *BRCA1* mutations and the same may not be true for missense mutations.

BRCA1-related tumors in general show more genomic aberrations than sporadic tumors, and also show specific chromosomal gains and losses.[11,28,29] The M1775K tumor analyzed here shows clearly abundant chromosomal changes, many of which comprise of small regions on different chromosomes, suggesting many intra-chromosomal breaks.

The shrunken centroids classifier used for *BRCA1* class prediction shows high specificity; no false-positive and false-negative calls were observed on 10 proven *BRCA1*-related tumors and 16 sporadic control tumors that were independent of the training sets.[28] The M1775K tumor showed a relatively low score for a *BRCA1*-related tumor and is just outside the 95% reference interval for BRCA1, possibly due to the noise relating to DNA quality. However, it is classified with a positive score and predicted to be *BRCA1*-like. As can be seen in Figure 2b, the tumor does show most of the characteristic *BRCA1* aberrations.

The functional assay based on the transcriptional activity of the C-terminal region of BRCA1 functions as a monitor of the integrity of the BRCT domains and provides a simple, standardized and validated test for unclassified variants that are located in the regions encompassed by the assay.[16,18,19,39] In addition, because a large set of variants can be tested and results compared across several experiments (provided internal controls are used) it is also a powerful way to cross-validate other methods such as computation prediction models.[40–42] It is interesting to note that a recently described computation prediction model correctly predicted the structural changes and the functional impact of the M1775K change.[40] However, one

should use caution when interpreting results from functional assays in the absence of other supporting data. For example, specific surface changes that do not cause major changes in the fold but disrupt specific binding sites not required for the activity of BRCA1 in recruiting the RNA polymerase II holoenzyme might be missed by the assay. Further validation is still needed to determine whether such variants exist, and how frequent they are with clear implications for the specificity of the assay. Moreover, while the *BRCA1* M1775R variant has been described as pathogenic,[30,31] this in itself is not sufficient to class M1775K as pathogenic as it is possible for different amino-acid changes at the same position to be benign or pathogenic, for example, the BRCA1 I15T variant prevents interaction with the E2 ubiquitin conjugating enzyme UbcH5a, whereas BRCA1 I15L has no effect.[43]

The BRCA1 region spanning residues 1646–1859 folds into two tandem domains (BRCT1 and BRCT2), which interact in a phosphorylation-dependent manner with proteins involved in DNA damage-induced checkpoint control, including BRIP1[4] and CtIP.[6] The present M1775K structure represents the third crystal structure (after M1775R and V1809F[44]) of a BRCA1 BRCT mutant and reveals the mechanism underlying its pathogenicity at the atomic level. It was previously shown that the BRCA1 BRCT mutant M1775R disrupts the BRCA1 interaction with BRIP1 and CtIP through steric interference between the guanidino group of Arg1775 and the Phe $+3$ of the ligand.[31,34] It, therefore, appears that the substituted Lys1775 and Arg1775 side chains directly block access of the Phe $+3$ to pocket P_2, thereby inhibiting the BRCA1 interaction with BRIP1,[34,35,45] CtIP[46] and likely other proteins that interact with the P_1 and P_2 pockets of the BRCT repeats through a similar mechanism, leading to defects in the double-strand DNA repair[47] and transactivation functions of BRCA1.[48] Taken together, these results indicate that the integrity of the BRCT phosphopeptide-binding pocket is required for the tumor suppression function of BRCA1.

Nucleotide substitutions in disease susceptibility genes that lead to a single amino-acid change pose a real challenge in risk assessment. Unlike changes that clearly affect protein structure and where impact on function can be inferred, the consequences on protein activity resulting from missense changes are difficult to determine. The large number of missense variants detected to date in *BRCA1* (Breast Cancer Information Core database, http:// research.nhgri.nih.gov/bic/) constitutes a significant hurdle to identify individuals at high risk for breast and ovarian cancer. All of the *BRCA1* and *BRCA2* missense substitutions with frequency $>1\%$ in the general population have been classified as neutral or of little clinical importance. Therefore, all of the potentially high risk, clinically important missense substitutions are rare, making direct epidemiological studies difficult to conduct. To determine whether a

certain unclassified variant is likely to be deleterious or neutral, it is necessary to combine methods from different disciplines as shown here. Some of these methods may only be applicable to certain domains of the protein and at present there is no standardized protocol or scoring system to weigh each method in deciding whether a variant is pathogenic or not. Given the number of different complementary assays used, the work described here might be considered as the 'gold standard' in determining pathogenicity, although given the resource-intense nature of the work and the destabilizing effects on the BRCT fold by certain missense mutations, it may not be feasible to consider such a comprehensive approach for each variant identified in the clinic.

References
1 Goldgar DE, Easton DF, Deffenbaugh AM, Monteiro AN, Tavtigian SV, Couch FJ: Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. *Am J Hum Genet* 2004; **75**: 535–544.
2 Chenevix-Trench G, Healey S, Lakhani S *et al*: Genetic and histopathologic evaluation of BRCA1 and BRCA2 DNA sequence variants of unknown clinical significance. *Cancer Res* 2006; **66**: 2019–2027.
3 Hayes F, Cayanan C, Barilla D, Monteiro AN: Functional assay for BRCA1: mutagenesis of the COOH-terminal region reveals critical residues for transcription activation. *Cancer Res* 2000; **60**: 2411–2418.
4 Cantor SB, Bell DW, Ganesan S *et al*: BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell* 2001; **105**: 149–160.
5 Seal S, Thompson D, Renwick A *et al*: Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles. *Nat Genet* 2006; **38**: 1239–1241.
6 Yu X, Chen J: DNA damage-induced cell cycle checkpoint control requires CtIP, a phosphorylation-dependent binding partner of BRCA1 C-terminal domains. *Mol Cell Biol* 2004; **24**: 9478–9486.
7 Ng PC, Henikoff S: SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* 2003; **31**: 3812–3814.
8 Tavtigian SV, Deffenbaugh AM, Yin L *et al*: Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet* 2006; **43**: 295–305.
9 Easton DF, Deffenbaugh AM, Pruss D *et al*: A systematic genetic assessment of 1433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes. *Am J Hum Genet* 2007; **81**: 873–883.
10 Lakhani SR, Van De Vijver MJ, Jacquemier J. *et al*: The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. *J Clin Oncol* 2002; **20**: 2310–2318.
11 van Beers EH, Joosse SA, Ligtenberg MJ *et al*: A multiplex PCR predictor for aCGH success of FFPE samples. *Br J Cancer* 2006; **94**: 333–337.
12 Joosse SA, van Beers EH, Nederlof PM: Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material. *BMC Cancer* 2007; **7**: 43.
13 Picard F, Robin S, Lavielle M, Vaisse C, Daudin JJ: A statistical approach for array CGH data analysis. *BMC Bioinformatics* 2005; **6**: 27.
14 Tibshirani R, Hastie T, Narasimhan B, Chu G: Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci USA* 2002; **99**: 6567–6572.
15 Joosse SA, Van Beers EH, Tielen IHG *et al*: Submitted.
16 Carvalho MA, Marsillac SM, Karchin R *et al*: Determination of cancer risk associated with germ line BRCA1 missense variants by functional analysis. *Cancer Res* 2007; **67**: 1494–1501.
17 Higuchi R: Recombinant PCR; in Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds): *PCR Protocols: a Guide to Methods and Applications*. San Diego, CA: Academic Press Inc., 1989, pp 177–183.
18 Phelan CM, Dapic V, Tice B *et al*: Classification of BRCA1 missense variants of unknown clinical significance. *J Med Genet* 2005; **42**: 138–146.
19 Carvalho MA, Couch FJ, Monteiro AN: Functional assays for BRCA1 and BRCA2. *Int J Biochem Cell Biol* 2007; **39**: 298–310.
20 Estojak J, Brent R, Golemis EA: Correlation of two-hybrid affinity data with *in vitro* measurements. *Mol Cell Biol* 1995; **15**: 5820–5829.
21 Golemis EA, Gyuris J, Brent R: Two-hybrid system/interaction traps; in Ausubel FM, Brent R, Kingston R *et al* (eds): *Current Protocols in Molecular Biology*. New York: John Wiley and Sons, 1994, pp 13.14.11–13.14.17.
22 Brent R, Ptashne M: A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* 1985; **43**: 729–736.
23 Otwinowski Z, Minor W: Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 1997; **276**: 307–326.
24 Vagin A, Teplyakov A: MOLREP: an automated program for molecular replacement. *J Appl Crystallogr* 1997; **30**: 1022–1025.
25 Emsley P, Cowtan K: Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 2004; **60**: 2126–2132.
26 Murshudov GN, Vagin AA, Dodson EJ: Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 1997; **53**: 240–255.
27 Thompson D, Easton DF, Goldgar DE: A full-likelihood method for the evaluation of causality of sequence variants from family data. *Am J Hum Genet* 2003; **73**: 652–655.
28 Joosse SA, Van Beers EH, Tielen IHG *et al*: Submitted.
29 Wessels LF, van Welsem T, Hart AA, van't Veer LJ, Reinders MJ, Nederlof PM: Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. *Cancer Res* 2002; **62**: 7110–7117.

30 Williams RS, Glover JN: Structural consequences of a cancer-causing BRCA1-BRCT missense mutation. _J Biol Chem_ 2003; **278**: 2630–2635.

31 Varma AK, Brown RS, Birrane G, Ladias JA: Structural basis for cell cycle checkpoint control by the BRCA1-CtIP complex. _Biochemistry_ 2005; **44**: 10941–10946.

32 Fenn TD, Ringe D, Petsko GA: POVScript+: a program for model and data visualization using persistence of vision ray-tracing. _J Appl Crystallogr_ 2003; **36**: 944–947.

33 Nicholls A, Sharp KA, Honig B: Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. _Proteins_ 1991; **11**: 281–296.

34 Clapperton JA, Manke IA, Lowery DM _et al_: Structure and mechanism of BRCA1 BRCT domain recognition of phosphory-lated BACH1 with implications for cancer. _Nat Struct Mol Biol_ 2004; **11**: 512–518.

35 Shiozaki EN, Gu L, Yan N, Shi Y: Structure of the BRCT repeats of BRCA1 bound to a BACH1 phosphopeptide: implications for signaling. _Mol Cell_ 2004; **14**: 405–412.

36 Cornelis RS, Neuhausen SL, Johansson O _et al_: High allele loss rates at 17q12–q21 in breast and ovarian tumors from BRCAl-linked families. The Breast Cancer Linkage Consortium. _Genes Chromosomes Cancer_ 1995; **13**: 203–210.

37 Osorio A, de la Hoya M, Rodriguez-Lopez R _et al_: Loss of heterozygosity analysis at the BRCA loci in tumor samples from patients with familial breast cancer. _Int J Cancer_ 2002; **99**: 305–309.

38 Smith SA, Easton DF, Evans DG, Ponder BA: Allele losses in the region 17q12–21 in familial breast and ovarian cancer involve the wild-type chromosome. _Nat Genet_ 1992; **2**: 128–131.

39 Vallon-Christersson J, Cayanan C, Haraldsson K _et al_: Functional analysis of BRCA1 C-terminal missense mutations identified in breast and ovarian cancer families. _Hum Mol Genet_ 2001; **10**: 353–360.

40 Karchin R, Monteiro AN, Tavtigian SV, Carvalho MA, Sali A: Functional impact of missense variants in BRCA1 predicted by supervised learning. _PLoS Comput Biol_ 2007; **3**: e26.

41 Mirkovic N, Marti-Renom MA, Weber BL, Sali A, Monteiro AN: Structure-based assessment of missense mutations in human BRCA1: implications for breast and ovarian cancer predisposi-tion. _Cancer Res_ 2004; **64**: 3790–3797.

42 Monteiro AN, Couch FJ: Cancer risk assessment at the atomic level. _Cancer Res_ 2006; **66**: 1897–1899.

43 Morris JR, Pangon L, Boutell C, Katagiri T, Keep NH, Solomon E: Genetic analysis of BRCA1 ubiquitin ligase activity and its relationship to breast cancer susceptibility. _Hum Mol Genet_ 2006; **15**: 599–606.

44 Williams RS, Lee MS, Hau DD, Glover JN: Structural basis of phosphopeptide recognition by the BRCT domain of BRCA1. _Nat Struct Mol Biol_ 2004; **11**: 519–525.

45 Manke IA, Lowery DM, Nguyen A, Yaffe MB: BRCT repeats as phosphopeptide-binding modules involved in protein targeting. _Science_ 2003; **302**: 636–639.

46 Yu X, Wu LC, Bowcock AM, Aronheim A, Baer R: The C-terminal (BRCT) domains of BRCA1 interact _in vivo_ with CtIP, a protein implicated in the CtBP pathway of transcriptional repression. _J Biol Chem_ 1998; **273**: 25388–25392.

47 Scully R, Ganesan S, Vlasakova K, Chen J, Socolovsky M, Livingston DM: Genetic analysis of BRCA1 function in a defined tumor cell line. _Mol Cell_ 1999; **4**: 1093–1099.

48 Monteiro AN, August A, Hanafusa H: Evidence for a transcrip-tional activation function of BRCA1 C-terminal region. _Proc Natl Acad Sci USA_ 1996; **93**: 13595–13599.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (http://www.nature.com/ejhg)