

Extending Structure-based Tools and Evolutionary Framework for Assessment of BRCA2 Missense Variants at the DSS1 Interface

EDIR VIDAL¹, KACPER MACIEJEWSKI¹, FLORENCIA DE LILLO¹, AND RODRIGO GALLEGOS⁴

¹ Group 5: s243564, s243548, s242869, s243563

Compiled May 11, 2025

BRCA2, a key protein involved in repairing damaged DNA, contains a DSS1-binding DNA-binding domain critical for its stable folding. We focused on the effects of different missense variants in this DBD on BRCA2 structure, stability, and binding to DSS1. Using structural models of the BRCA2 DBD, we predicted mutation-induced changes in folding free energy and in BRCA2-DSS1 binding affinity. Furthermore, to better capture the conformational flexibility changes we leveraged a series of simulations to sample the protein's dynamic ensemble. The results obtained were compared with variant effect scores from the MAVISp framework. This integrative analysis highlights areas of consensus and discrepancy between structure-based simulations and MAVISp score predictions, enriching the interpretative framework for BRCA2 variants.

<http://dx.doi.org/10.1364/ao.XX.XXXXXX>

1. INTRODUCTION

BRCA2 is a tumor suppressor that preserves genome stability by recruiting and loading RAD51 into DNA breaks through its BRC repeats and the C-terminal DNA binding domain (DBD) [1]. This function requires DSS1, a small acidic partner that stabilizes DBD, mediates proper nuclear localization, and facilitates RPA-to-RAD51 exchange on single-stranded DNA during homologous recombination [2]. Germline BRCA2 mutations confer high risks of breast, ovarian, and other types of cancer [1]. Furthermore, disruption of BRCA2-DSS1 binding abolishes RAD51 focus formation and HR efficiency, underscoring the critical role of the complex in tumor suppression [2].

BRCA2 is known to be highly intolerant to missense variation: ClinVar (April, 2024) lists almost 9,000 BRCA2 missense variants, which 95% are variations of unknown significance (VUS) or with conflicting classification [3]. Majority of known pathogenic mutations reside in the C-terminal DBD (residues 2479 - 3192), which destabilize the binding domain or perturb BRC repeats leading to a disruption in the binding of RAD51-DSS1 and the engagement of DNA [2]. Functional HDR reporter assays show that many DBD missense changes abolish BRCA2 activity, showing the challenge of classifying these VUS [4].

The MAVISp framework integrates structural and evolutionary analyses to predict missense variant effects in proteins [5]. It maps mutations onto 3D structures and interaction networks, aggregating scores from AlphaMissense [6], GEMME [7], EVE [8], DeMaSk [9], among others. By combining these with structure-based metrics, MAVISp highlights BRCA2 variants likely to

destabilize the protein or disrupt DSS1 binding. Thus, MAVISp provides multi-model pathogenicity annotations grounded in structural context.

Structure-based tools further quantify mutational energetics. MutateX automates in silico saturation mutagenesis with FoldX energy function to calculate free energy for folding and binding across all residue substitutions, handling structure preparation, parallel execution, and visualization to pinpoint BRCA2 variants that destabilize the protein or weaken DSS1 binding [10]. On the other hand, CABS-flex uses a coarse-grained Monte Carlo approach to generate near-native conformational ensembles from a static structure, capturing intrinsic dynamics faster than all-atom simulations [11]. Comparing wild-type and mutant BRCA2 flexibility at the DSS1 interface can thus reveal variants that perturb stability or inter-protein contacts.

There are variants that its pathogenicity classification on different models in the MAVISp framework differ, giving an opportunity to further enrich this analysis with structure-based methods. In this study, we combine the MAVISp platform with MutateX free energy calculations and the CABS-flex ensemble to predict structural and functional impacts of BRCA2 VUS, focusing on those at the DSS1 binding interface that have an overall prediction discrepancy between different models.

2. MATERIALS AND METHODS

To investigate the structural and functional consequences of BRCA2 missense variants at the DSS1 interface, we used an integrated pipeline combining variant annotation, structural mod-

eling, mutational scanning, and dynamics profiling. Clinically observed oncogenic mutations were first curated via CancerMut and PDBMiner. Three-dimensional models of the BRCA2-DSS1 complex were generated with AlphaFold 3, after which point substitutions were introduced and their $\Delta\Delta G$ effects on folding and binding free energies were computed using MutateX. Conformational flexibility of both wild-type and mutant complexes was then sampled with CABS-flex coarse-grained simulations.

A. Structure and Mutant Selection

A preliminary survey with PDBMiner revealed that existing BRCA2 crystal and cryo-EM structures contained unresolved regions, making them unsuitable for our pipeline. We therefore used AF3 to generate complete models of both BRCA2 and the BRCA2-DSS1 complex. In parallel, CancerMut was queried to extract oncologically relevant missense variants located at the DSS1 interface (BRCA2 2479-3192 residues). A total of 565 mutations were extracted from this query, where only 299 were filtered as a way to extract only variants of uncertain significance (VUS) according to ClinVar. Finally, we decided to use only sites that have more than 2 VUS recorded, which resulted in 12 positions on this interacting region. From this final set, each of them was annotated by the MAVISp precomputed scores to visualize the prediction among different models (AlphaMissense, GEMME, DeMaSk and EVE).

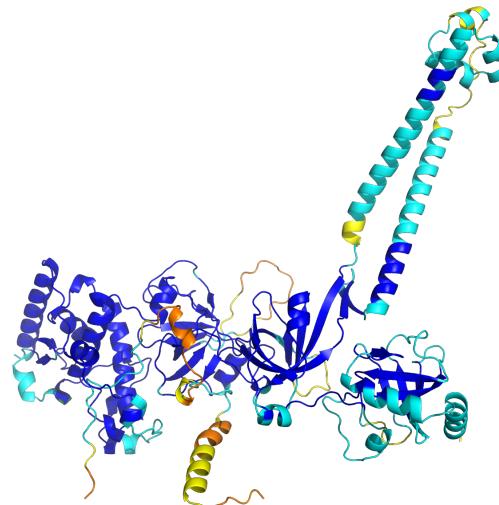
B. Stability

To assess the stability impacts of BRCA2 variants, we employed computational simulations combining MutateX [12] and CABS-flex [11] methods. Initially, structural models of the BRCA2 protein were generated using AlphaFold3. MutateX was then applied to predict the folding stability changes introduced by selected missense variants, classifying mutations into four distinct categories based on calculated $\Delta\Delta G$ values: stabilising ($\Delta\Delta G \leq 3\text{kcal/mol}$), neutral ($-2\text{kcal/mol} \leq \Delta\Delta G < 2\text{kcal/mol}$), uncertain ($2\text{kcal/mol} \leq \Delta\Delta G < 3\text{kcal/mol}$ or $-3\text{kcal/mol} \leq \Delta\Delta G < -2\text{kcal/mol}$), and destabilising ($\Delta\Delta G \geq 3\text{kcal/mol}$). To improve the prediction accuracy for borderline or uncertain cases, we employed ensemble-mode simulations using CABS-flex, which generated ten alternative conformations of the protein. This ensemble approach allowed a detailed exploration of protein flexibility, thus refining the stability predictions.

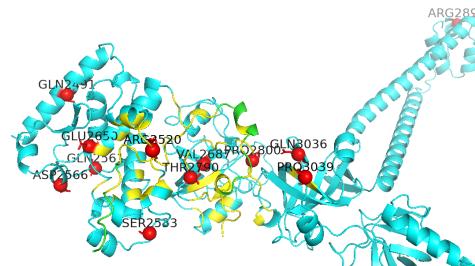
C. Binding

To assess whether the mutations of interest affected the interaction between BRCA2 and DSS1, the structure 1IYJ [13] was selected from the Protein Data Bank (PDB) as a reference. However, this structure presented missing residues in the regions relevant to our study. Therefore, to obtain a complete model of the interaction, the complex was re-modeled using AlphaFold3.

In order to achieve a high-quality prediction, only the folded portion of BRCA2 was considered, rather than the entire protein. Previous structural data and literature indicate that the only stably folded region of BRCA2 lies between residues 2479 and 3100, which is also the region covered by the experimental PDB structures. Accordingly, we provided AlphaFold3 with the FASTA sequence corresponding to this region and the entire DSS1 sequence. The AlphaFold3 prediction resulted in an ipTM score of 0.61 and a pTM score of 0.76, which are within acceptable confidence ranges, though some regions displayed locally low confidence scores, as shown in Figure 1a.



(a) BRCA-DSS1 AlphaFold3 model colored by confidence score



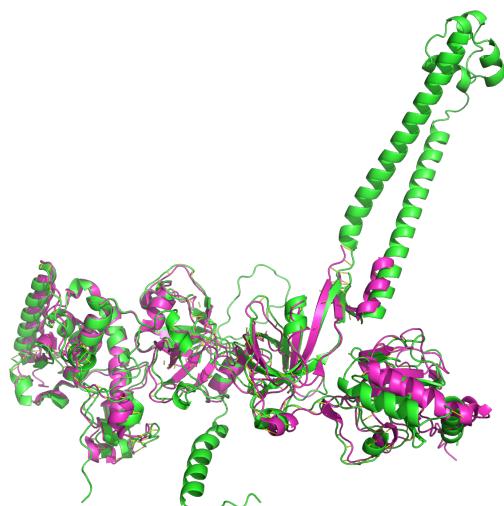
(b) Selected residues are highlighted in red, interface atoms are shown in yellow, DSS1 chain is represented in green and the BRCA full molecule is displayed in light blue.

Fig. 1. (a) Predicted 3D structure of the protein generated by AlphaFold 3. (b) Structure showing residues selected for targeted mutation analysis in red.

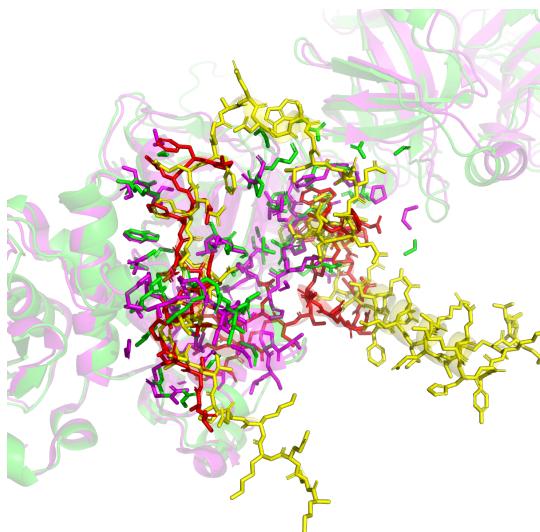
Nevertheless, when visualized and aligned with the available PDB structure with Pymol, the predicted and experimental conformations showed a high degree of similarity (Figure 2a). Special attention was given to the interaction area (Figure 2b), making sure there were no major discrepancies between the models. An unstable alpha-helix was removed from the model, and the final structure was saved as a PDB file for downstream analysis.

Binding free energy changes upon mutation were calculated using MutateX [12], (foldX5, source code: <https://github.com/ELELAB/mutatex>). Templates and configuration files used for these calculations are also available in the associated GitHub repository. The protein-protein interface was defined based on inter-chain contacts within a 4 Å cutoff (in order to include all possible interaction types and forces). Among the residues considered for mutation, only residues 2520, 2687, 2790, 2800, and 3036 were identified as part of this interface.

Since FoldX is limited to calculating local interactions, residues outside this interface are expected to have insignificant effects on the binding free energy ($\Delta\Delta G = 0$), making these five residues the primary focus for the MutateX analysis.



(a) Structural comparison between the AlphaFold model (green) and the experimental PDB structure (magenta)



(b) The red chain represents chain C from the PDB structure I1YDJ (DDS1), the yellow chain was modeled using AlphaFold, green highlights the interaction interface predicted by AlphaFold, and violet indicates interface residues identified in the experimental PDB structure.

Fig. 2. Structural Comparison of Experimental and Predicted Protein Interfaces.

Mutations were classified based on their predicted impact on the protein's free energy change. Stabilizing mutations were defined as those with $\Delta\Delta G \leq -1$ kcal/mol, and destabilizing ones as $\Delta\Delta G \geq 1$ kcal/mol. Mutations with values between -1 and 1 were considered neutral. To address borderline cases, we applied a ± 0.15 kcal/mol threshold around -1 and 1 to flag values that were very close to these cutoffs. For these non-conclusive results, the MutateX pipeline was rerun in ensemble mode, using multiple Cabflex generated models, instead of a single structure. Classification was automated using a Python script, which returned either "stabilizing," "destabilizing," "neutral" or the raw $\Delta\Delta G$ value if the result was inconclusive.

D. Ensemble generation –method

For ensemble generation, CABFlex was used to generate 10 models for each BRCA2 structure—both in isolation and in complex with DSS1. For the isolated BRCA2 structure, the models were used to assess structural stability, while for the complex, they were used to evaluate binding free energy stability. The following parameters were used for model generation.

Parameter	Value
medoids	10
montecarlo	30
annealing_cycles	10
verbosity	3
seed	10

Table 1. Model configuration parameters

3. RESULTS

A. CABFlex model assessment

To evaluate the conformational flexibility and structural robustness of key BRCA2 mutations, CAB-Flex was used to generate structural ensembles. The analysis was performed on both the BRCA2–DSS1 complex and the BRCA2 protein alone, producing 10 alternative models per case. These models were subsequently assessed using Mutatex to evaluate local structural quality through self-mutation $\Delta\Delta G$ (ddG) scans.

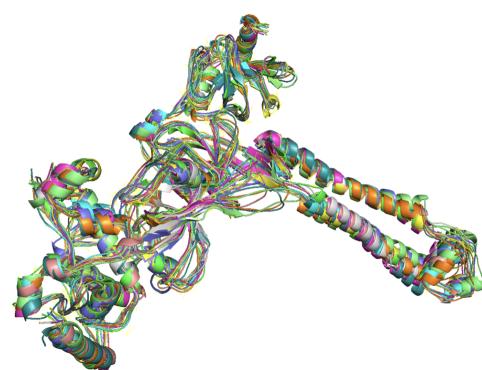
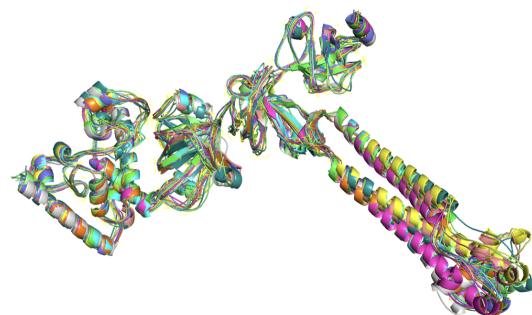


Fig. 3. Structural alignment of BRCA2 alone(up) and complex models(down) generated by CAB-Flex

As shown in Figure 3, the overall fold of the BRCA2–DSS1 complex remains highly conserved across the ensemble, with most deviations confined to flexible loop regions. This behavior aligns with CAB-Flex's goal of sampling conformational variability, particularly in regions with high inherent mobility.

Figure 3 also shows that the core topology of BRCA2 is largely retained even in the absence of DSS1. The primary deviation lies in the long α -helix, a region expected to display increased variability due to the absence of DSS1-mediated stabilization. Again, this reflects the goal of CAB-Flex to explore plausible conformational states within structurally relevant constraints.

To further assess the quality and stability of these models, self-mutation ddG values were calculated using Mutatex. High ddG values (>1 kcal/mol) may suggest unstable or poorly modeled regions, while significantly negative values could indicate overpacking or artificial stabilization—both of which should be treated cautiously when interpreting mutational effects.

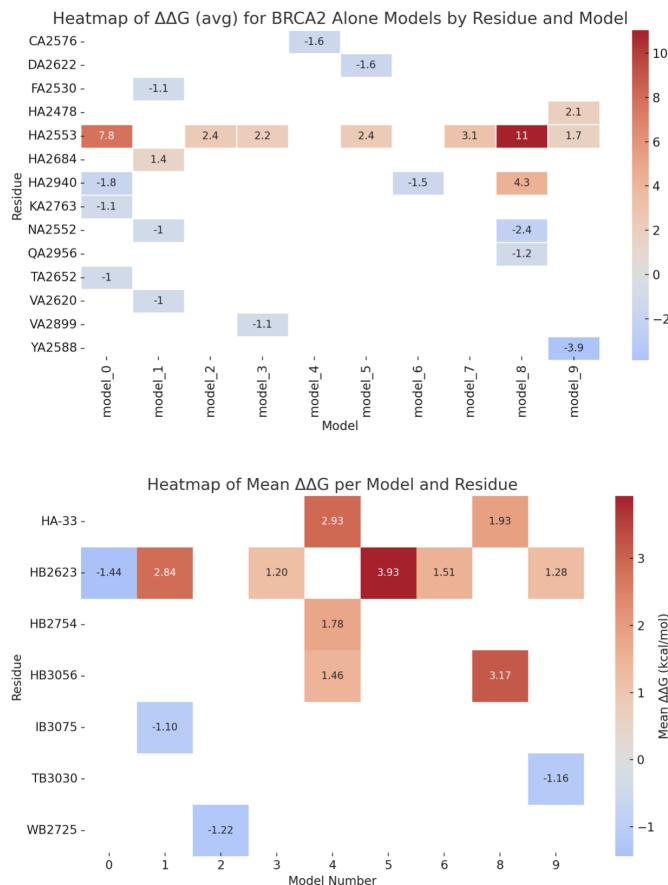


Fig. 4. Heatmap of residues with mean high and low ddG across BRCA2-alone and complex CAB-Flex models

In the BRCA2-alone ensemble, only four residues displayed consistently high ddG values across the models, suggesting that most conformations are structurally robust. Among these, HA2553 appeared in multiple models, while the others were limited to one or two occurrences, indicating localized structural sensitivity. On the other hand, a larger number of residues showed low ddG values; however, they were distributed sparsely across the models, suggesting that occasional over-stabilization exists but does not undermine the general integrity of the models.

For the complex structure, only four residues showed significant ddG variations, reinforcing the structural reliability of the ensemble. Notably, HB2623 appeared in multiple models with both high and low ddG values, indicating inconsistent local geometry or flexibility that may reduce predictive reliability in that region. Residues HA-33, HB2754, and HB3056 showed consistently high ddG, possibly reflecting unresolved strain or unfavorable contacts. In contrast, WB2725 and TB3030 showed negative ddG values, which may result from overpacking or modeling artifacts.

Taken together, the combination of structural alignment and ddG self-scans suggests that most CAB-Flex-generated models are structurally sound and suitable for downstream mutational analysis. However, certain regions—particularly those with repeated extreme ddG shifts—should be treated with caution or excluded when evaluating mutation impacts.

B. Stability

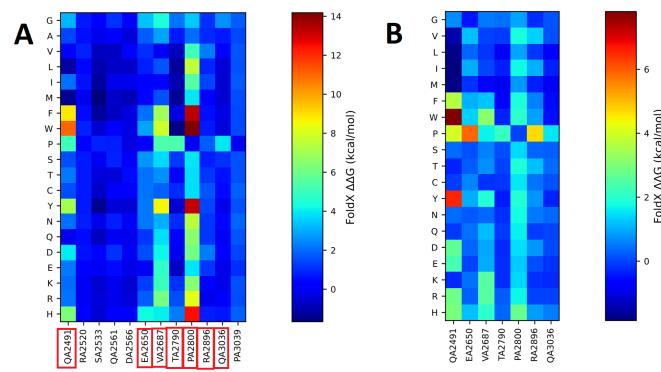


Fig. 5. Stability analysis run with MUTATEX [10] using A) single mode, and B) ensemble mode for selected (marked red) positions which were not clearly resolved by single mode.

The results of stability analysis, illustrated in Figure 5, clearly differentiated mutation impacts on BRCA2 based on single- and ensemble-mode predictions.

Residue Q2491 showed predominantly neutral stability for substitutions G, V, L, I, and M, indicating minor perturbation. However, substitutions to F (single: $\Delta\Delta G = 3.14$ kcal/mol, ensemble: $\Delta\Delta G = 3.64$ kcal/mol), W (single: $\Delta\Delta G = 8.80$ kcal/mol, ensemble: $\Delta\Delta G = 7.81$ kcal/mol), P (single: $\Delta\Delta G = 7.02$ kcal/mol, ensemble: $\Delta\Delta G = 4.01$ kcal/mol), and Y (single: $\Delta\Delta G = 7.02$ kcal/mol, ensemble: $\Delta\Delta G = 6.54$ kcal/mol) significantly destabilised the protein. The introduction of bulky aromatic side chains like phenylalanine (F), tryptophan (W), and tyrosine (Y), or the rigid cyclic structure of proline (P), likely created steric clashes or disrupted hydrophobic core interactions, severely compromising structural integrity.

Residue E2650 was generally tolerant to substitutions, maintaining neutral stability, except for substitution to P (single: $\Delta\Delta G = 5.85$ kcal/mol, ensemble: $\Delta\Delta G = 5.85$ kcal/mol). The rigidity and cyclic structure of proline likely disrupted local hydrogen bonding networks or introduced backbone rigidity, resulting in substantial structural destabilisation. Notably, G (single: $\Delta\Delta G = 3.38$ kcal/mol) was destabilising in single-mode but neutral ($\Delta\Delta G = -0.44$ kcal/mol) in ensemble-mode.

Residue V2687 showed mainly neutral stability impacts, except for substitution to W (single: $\Delta\Delta G = 3.13$ kcal/mol, ensemble: $\Delta\Delta G = 3.13$ kcal/mol). Destabilisation is likely due to

steric hindrance introduced by the bulky, aromatic side chain of tryptophan. Residue 2896 substitution to P (single: $\Delta\Delta G = 4.69$ kcal/mol, ensemble: $\Delta\Delta G = 4.69$ kcal/mol) also induced notable destabilisation, again likely due to proline-induced rigidity or disruption of local secondary structure interactions.

Residue T2790, Q3036, and P2800 showed exclusively neutral stability across all tested substitutions, demonstrating remarkable tolerance to diverse amino acid replacements without significant impact on structural integrity. Notably, no mutations across all analysed residues were identified as stabilising, highlighting a general intolerance for significant structural modifications that enhance stability.

Aromatic substitutions (F, W, Y) consistently showed a strong tendency to destabilise the BRCA2 structure. This effect is attributed to the bulky, hydrophobic, and planar nature of aromatic side chains, which can disrupt packing within the hydrophobic core, create steric clashes, or interfere with local secondary structure elements. These findings highlight a general sensitivity of BRCA2 to aromatic substitutions, especially in buried regions, underscoring the importance of side chain compatibility in maintaining protein stability.

C. Binding free energy

C.1. Single-mode

Figure 6a shows the results from running MUTATEX on single mode. As expected, those mutations present in residues that are far apart from the interface area resulted in 0.

Considering the ones close to the interface, mutations in residues 3036 and 2800 had no significant destabilizing effect, with $\Delta\Delta G$ values also close to zero, being classified as neutral according to our previously defined thresholds. This makes sense since when analyzing contacts among 4 Å in Pymol, we see no direct interaction between these residues and the DSS1 (Fig ??). In contrast, mutations at position 2520 had a severe impact on binding stability. Any substitution of Arginine with another amino acid caused substantial destabilization, which could be explained by considering the position of this residue, that has lots of interactions with DSS1 (Figure ??).

Analyzing the position and interactions of residues 2687 and 2790 (Figure ??), we might expect mutations at residue 2790 to have a greater impact than at 2687, given its higher number of interactions. However, both residues are located in densely packed regions surrounded by many atoms, which means that mutations at either site could still affect the binding interaction due to steric constraints.

For residue 2687, out of the 19 mutations, 13 resulted to have neutral effects, 5 resulted in destabilizing ($\Delta\Delta G > 1$) and 1 inconclusive result ($V \rightarrow D$). Particularly, replacing valine (V) with phenylalanine (F), tryptophan (W), tyrosine (Y), glutamate (E) or Histidine (H) weakened the interaction. Considering that valine is a small, nonpolar aminoacid, one could expect that substituting it with F, W, or Y could cause steric clashes from the introduction of large aromatic side chains. As regards E or H, they introduce charged or polar groups into a likely nonpolar environment, potentially disrupting local interactions and increasing the free binding energy.

At position 2790, out of 19 possible mutations, 4 ($T \rightarrow M, F, P, K$) resulted in inconclusive results due to predicted $|\Delta\Delta G| \approx 1$. Two mutations ($T \rightarrow D, E$) were destabilizing, while one ($T \rightarrow R$) was predicted to be slightly stabilizing. Threonine is a small, polar amino acid that can form stabilizing hydrogen bonds through its hydroxyl group. Replacing it with aspartate

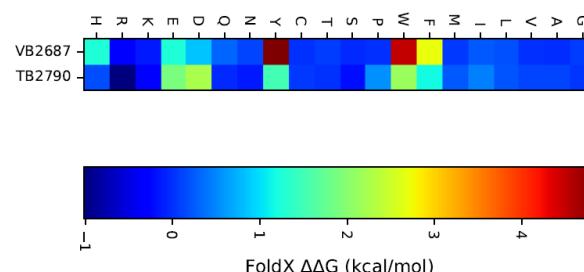
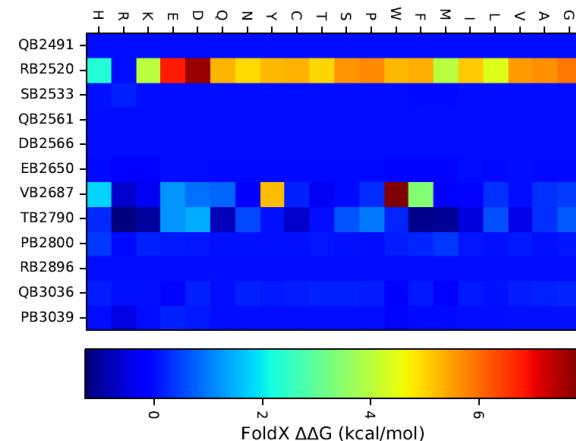


Fig. 6. Heatmaps summarizing the results from MUTATEX binding free energy: (a) single model predictions and (b) ensemble predictions.

or glutamate introduces negatively charged side chains, which may disrupt local interactions and reduce stability. On the other hand, replacing threonine for arginine, means replacing it by a larger, positively charged residue. This might create new stabilizing interactions with nearby negatively charged residues, helping to strengthen the interface and improve binding.

C.2. Ensemble-mode

Fig 6b shows a heatmap. After running ensemble mode on the residues that presented uncertain mutation effects, we compared the results with predictions obtained in single mode. As shown in the Table 2, several mutations showed logical classifications across both methods. For example, the mutation $T2790 \rightarrow M$ that was uncertainly classified as stabilizing in single mode ($\Delta\Delta G = -1.05474$) and neutral in ensemble mode, which is a borderline case.

Similarly, $T2790 \rightarrow K, D, P$ values near the classification thresholds and their classification in ensemble mode resulted in them being neutral.

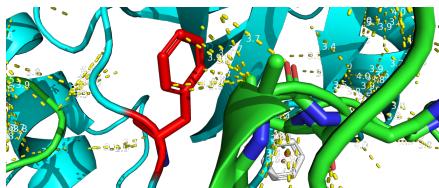
However, some results diverged. Mutations such as $T2790 \rightarrow W$ and $T2790 \rightarrow Y$ changed from neutral to destabilizing. Which may reflect that are conformation dependant effects. Finally, some mutations showed a complete discrepancy. $T2790 \rightarrow F$ shifting from a stabilizing prediction in single mode ($\Delta\Delta G = -1.14$) to a destabilizing prediction value (1.16382) in ensemble mode.

Residue	AA mut	Single_mode	Ensemble_mode
T_2790	M	-1.05474	neutral
T_2790	F	-1.14306	1.16382
T_2790	W	neutral	destabilizing
T_2790	P	0.96584	neutral
T_2790	Y	neutral	destabilizing
T_2790	K	-0.98144	neutral
T_2790	R	stabilizing	-1.01726
V_2687	D	0.88492	0.8478

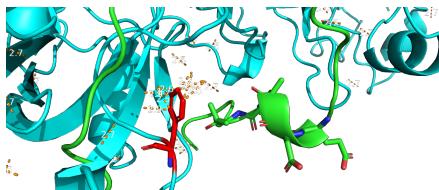
Table 2. Dicrepancy results between single and ensemble mode

This mutation was simulated in PyMOL for further analysis. Phenylalanine introduces a huge, nonpolar aromatic ring that alters existing interactions (Figure 8a) and causes steric clashes with surrounding residues (Figure 8b).

While the single-mode prediction missed this effect, the ensemble mode caught it by exploring different conformations. In some of these, the bulky phenyl ring ends up in non-favourable positions as the ones shown in Figure 8, causing clashes and interfering with the normal interactions.



(a) Interactions



(b) Clashes

Fig. 8. Interactions and Clashes simulated in Pymol when replacing T with F on residue 2790

D. Integration with MAVISp scores

To fully enrich the analysis with the proposed framework, pre-computed MAVISp scores for pathogenicity and functional consequence of mutations where integrated to the local interaction and stability scores computed by MutateX. Figure 7 completely compiles all of these scores for the single-mode.

D.1. Pathogenicity prediction

For the mutation sites, we have extracted precomputed scores and mutation predicted effect given by AlphaMissense and EVE, having mutations that are either pathogenic, benign or of uncertain effect. From these we have several mutation where the predicted effect differed, and none of the 12 sites have complete agreement of the pathogenicity effect. For example, the amino acid S2533 has every single mutation catalogued as pathogenic

by EVE, but in the case of AlphaMissense it is classified as benign, with the exception of one being classified as uncertain.

D.2. Functionality prediction

In the case of stability prediction, GEMME and DeMask precomputed predicted effect of the mutation were used. Because the precomputed scores from GEMME doesn't had a classification, it was considered everything above 0.75 of the rank-normalized score to be loss of function, while everything below 0.25 to be neutral. Anything in between is considered to be uncertain, as there is not enough evidence one can assign either both of these classes. The trend is similar than that of the pathogenicity predictors: there is a high discrepancy among predictions, and there is actually no single site where the results are conclusive.

D.3. Ensemble

To finalize the analysis, ensemble annotations where also integrated to the predictors classification. Figure 9 compiles the results from both the ensemble models generated for local interaction calculations and stability.

4. DISCUSSION

In our CABS-Flex ensembles, self-mutation $\delta\delta G$ scans revealed that the core BRCA2-DSS1 fold remains highly conserved, with only a few non-interface residues showing significant destabilization, supporting the overall structural reliability of the generated models. Notably, Q2491 and E2650 displayed pronounced sensitivity to bulky aromatic (phenylalanine, tryptophan, tyrosine) and proline substitutions, each causing $\delta\delta G$ increases of 3–8 kcal/mol in both single-structure and ensemble modes, highlighting their role in maintaining local hydrophobic and hydrogen-bond networks.

After integrating our thermodynamic stability measurements on the isolated BRCA2 domain with detailed binding-energy calculations for the BRCA2–DSS1 complex, mutational effects on the structures were compared and assessed. In the absence of DSS1, substituting V2687, T2790, or R2520 with most other amino acids tends to destabilize the BRCA2 fold. These destabilizations arise from disrupted hydrophobic cores (in the case of V2687), loss of key polar contacts (for T2790), or breakage of critical electrostatic networks (for R2520). When DSS1 is present, however, many of these destabilizing effects are at least partially rescued. At V2687 and T2790, substitutions to small or charged residues (for example, lysine or arginine at T2790) frequently recover binding-free energy to nearly wild-type levels. We attribute this rescue to the formation of new electrostatic or hydrogen-bond interactions between the substituted side chains and nearby DSS1 residues—interactions that compensate for the weakened BRCA2 core. This *molecular chaperone* effect of DSS1 illustrates how partner binding can buffer moderate instabilities in protein domains, effectively masking the phenotypic consequences of many, but not all, variants.

In contrast, substitutions to bulky aromatic side chains (phenylalanine, tryptophan, tyrosine) at either V2687 or T2790 remain severely destabilizing even in the complex. The large ring systems at these positions create steric clashes both within the BRCA2 fold and at the protein–protein interface, preventing DSS1 from fully accommodating the mutant side chain. R2520 stands apart as uniquely intolerant to any substitution. Every variant we tested at this position led to dramatic losses in binding-free energy, and none showed the neutralizing effect of DSS1 binding seen elsewhere. Structurally, R2520 participates in a dense network of salt bridges and hydrogen bonds that anchor



Fig. 7. MAVISp precomputed predicted mutation effect integrated with MutateX local interaction and stability results.

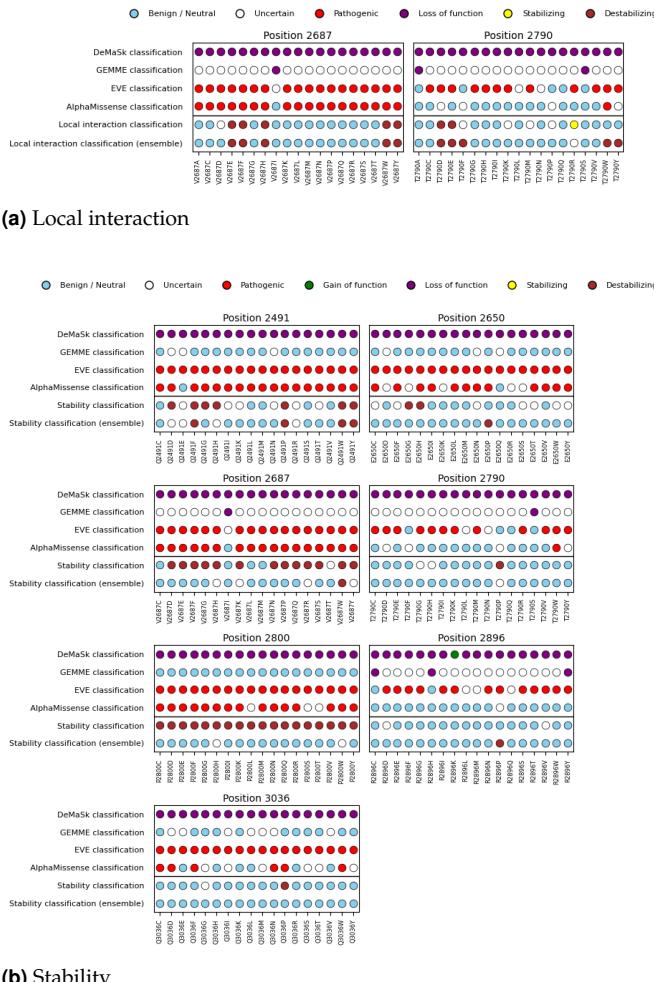
a key loop in BRCA2 to DSS1. Loss of these interactions cannot be compensated by alternative side chains or by the inherent flexibility of the interface.

Beyond these three hotspots, our analysis of additional interface-proximal residues, such as S2800 and T3036, revealed minimal changes in binding energy upon mutation. Although these side chains lie near the DSS1 contact surface, they do not participate in critical atomic contacts and their replacement has little functional consequence. Finally, by comparing single-structure predictions with ensemble-based conformational sampling (via CABS-Flex), we uncovered cases where static models misestimate mutational impact. A good example is T2790 → F where in a single mode, the bulky phenylalanine appears tolerated, but ensemble sampling reveals backbone shifts and new clashes that substantially weaken binding. This dynamic perspective argues for inclusion of flexibility in any predictive workflow and underlines the importance of ensemble-mode analyses.

In summary, our combined MAVISp–MutateX–CABS-Flex pipeline delivers a high-resolution map of mutation sensitivity across the BRCA2–DSS1 interface. It robustly identifies truly pathogenic substitutions, most notably at R2520, while also clarifying why many variants of uncertain significance exhibit moderate or context-dependent effects.

5. CONCLUSIONS

This study identified key mutations in the BRCA2 protein that significantly impact its interaction with DSS1, a partnership essential for maintaining genomic integrity. By integrating energy predictions, structural modeling, and conformational sampling, we were able to distinguish between mutations with clear destabilizing effects whose impact depends on the specific amino acid substitution and local flexibility. Our results have shown the importance of considering both structural context and dynamic behavior when interpreting the functional consequences of protein variants. Our presented integrated approach can be readily extended to other tumor suppressors or protein–protein complexes, providing a powerful strategy for mechanistic interpretation of missense variants in precision medicine.



(b) Stability

Fig. 9. Ensemble mode comparison between predictors and MutateX results.

AUTHOR CONTRIBUTIONS

- **Introduction, Discussion, Conclusion:** All authors.
- **Materials and Methods & Results:**
 - Structure and Mutant Selection: Edir Vidal
 - Stability: Kacper Maciejewski
 - Binding: Florencia De Lillo
 - CaBFlex: Rodrigo Gallegos
 - Integration with MAVISp scores: Edir Vidal

A. Disclosures

ChatGPT was used to help clarify ideas and improve the grammar and flow of the text.

B. Data Availability

All files used for this project can be found in `/home/projects/22117_proteins_2025/projects/group5_project` on `pupil2.healthtech.dtu` server or alternatively, be generated by implementing the code published in [Github](#)

REFERENCES

1. Y. Huang, W. Li, T. Foo, *et al.*, “Dss1 restrains brca2’s engagement with dsdna for homologous recombination, replication fork protection, and r-loop homeostasis,” *Nature Communications* **15**, 7081 (2024).
2. A. P. Mishra, S. A. Hartford, S. Sahu, *et al.*, “Brca2-dss1 interaction is dispensable for rad51 recruitment at replication-induced and meiotic dna double strand breaks,” *Nature Communications* **13**, 1751 (2022).
3. M. J. Landrum, J. M. Lee, G. R. Riley, *et al.*, “Clinvar: public archive of relationships among sequence variation and human phenotype,” *Nucleic Acids Research* **42**, D980–D985 (2013).
4. C. Hu, L. R. Susswein, M. E. Roberts, *et al.*, “Classification of BRCA2 variants of uncertain significance (VUS) using an ACMG/AMP model incorporating a homology-directed repair (HDR) functional assay,” *Clin. Cancer Res.* **28**, 3742–3751 (2022).
5. M. Arnaudi, M. Utichi, K. Degn, *et al.*, “Mavisp: A modular structure-based framework for protein variant effects,” *bioRxiv* (2024).
6. J. Cheng, G. Novati, J. Pan, *et al.*, “Accurate proteome-wide missense variant effect prediction with AlphaMissense,” *Science* **381**, eadg7492 (2023).
7. E. Laine, Y. Karami, and A. Carbone, “GEMME: A simple and fast global epistatic model predicting mutational effects,” *Mol. Biol. Evol.* **36**, 2604–2619 (2019).
8. J. Frazer, P. Notin, M. Dias, *et al.*, “Disease variant prediction with deep generative models of evolutionary data,” *Nature* **599**, 91–95 (2021).
9. D. Munro and M. Singh, “DeMaSk: a deep mutational scanning substitution matrix and its use for variant impact prediction,” *Bioinformatics* **36**, 5322–5329 (2021).
10. M. Tiberti, T. Terkelsen, K. Degn, *et al.*, “MutateX: an automated pipeline for in silico saturation mutagenesis of protein structures and structural ensembles,” *Brief. Bioinform.* **23** (2022).
11. M. Kucinski, T. Oleniecki, M. P. Cierny, *et al.*, “CABS-flex standalone: a simulation environment for fast modeling of protein flexibility,” *Bioinformatics* **35**, 694–695 (2019).
12. M. Tiberti, T. Terkelsen, K. Degn, *et al.*, “MutateX: an automated pipeline for in silico saturation mutagenesis of protein structures and structural ensembles,” *Briefings Bioinformatics* **23**, bbac074 (2022).
13. H. Yang, P. D. Jeffrey, J. Miller, *et al.*, “Brca2 function in dna binding and recombination from a brca2-dss1-ssdna structure,” *Science* **297**, 1837–1848 (2002).