RESEARCH PAPER

Principles of Protein Stability and their Application in Computational Design: A Comprehensive Review

A. Jasmine Sugil[1], K. Merriliance[2], Antony Vasantha Kumar C[3], S. Serina Hingis[4]

[1]Research Scholar, Department of Computer Applications and Research Centre, Sarah Tucker College (Autonomous), Affiliated to Manonmaniam Sundaranar University, Tirunelveli-627012, Tamilnadu, India

[2]Associate Professor, Department of Computer Applications and Research Centre, Sarah Tucker College (Autonomous), Tirunelveli-07, Tamilnadu, India

[3]Associate Professor, KIT-Kalaignarkarunanidhi Institute of Technology, Coimbatore, Tamil Nadu, India

[4]Regulatory Affairs, Axxelent Pharma Sciences Pvt Ltd, Chennai, India

Corresponding author: A. Jasmine Sugil | Email: findsugil@gmail.com

ABSTRACT

Recent advances in protein stability prediction have substantially influenced medicine, particularly in the development of novel therapeutic agents and immunomodulatory drugs. Despite this progress, several critical challenges remain. First, existing databases are inadequate for comprehensive thermodynamic measurements of proteins. Second, ΔΔG values exhibit considerable variability due to inconsistent experimental conditions. Third, many predictive methods neglect the anti-symmetry property of ΔΔG values between mutant and native proteins. Fourth, prediction performance is frequently overestimated as a result of sequence similarity between proteins in training and test datasets. This review examines these challenges, identifies emerging issues, and proposes strategies to enhance the accuracy and reliability of computational prediction tools. Furthermore, it discusses the potential of these methods to inform the design of precision medicine strategies for genetic disorders, such as Parkinson's disease, cancer, cystic fibrosis, and other neurodegenerative diseases associated with protein misfolding.

Keywords: Protein folding, Protein stability, Mutation, Gibbs free energy, Anti-symmetry.

How to cite this article: Sugil AJ, Merriliance K, Antony Vasantha Kumar C, Serina Hingis S. Principles of Protein Stability and their Application in Computational Design: A Comprehensive Review. Int J Drug Deliv Technol. 2026;16(55s): 18-34. DOI: 10.25258/ijddt.16.55s.3

Source of support: Nil.

Conflict of interest: None

.1. Introduction

Protein structural stability is closely linked to its biological function. Single-point residue substitutions can significantly influence protein stability, thereby elucidating molecular mechanisms underlying human diseases and facilitating drug discovery in [1]. Numerous methods have been developed to predict the stability of structural protein variants by leveraging sequence or structural characteristics [2]. These predictive approaches frequently estimate changes in Gibbs free energy (ΔΔG) of folding or binding for domains and interfaces impacted by mutations. For example, ELASPIC [3] utilises homology modelling to construct protein structures and demonstrates that such mutations can diminish protein stability and binding affinity. STRUM [4] predicts stability changes resulting from single-point mutations by employing I-TASSER simulations to generate three-dimensional models from wild-type sequences. Accurate annotation of protein function and disease diagnosis depends on understanding single-nucleotide polymorphism (SNP)-induced stability changes (ΔΔG). Among sequence-based features, the volume difference between wild-type and mutant amino acids is particularly significant. Features of mutant amino acids have an average importance score of 0.013, which is 4.33% higher than the 0.003 average for wild-type features, underscoring the importance of using mutant sequences to generate specific profiles and predictive models [4]. Non-synonymous DNA mutations can modulate protein function by altering stability, potentially disrupting essential conformational transitions [5]. Alterations in protein stability represent a principal molecular mechanism in many mutation-induced disorders [6] and may constitute a more prevalent cause of disease and functional loss than previously appreciated. Elucidating the effects of specific mutations on protein stability or molecular interactions can inform therapeutic development and help identify potential drug resistance or sensitivity. This understanding is also critical for optimising protein engineering through mutagenesis, thereby advancing precision medicine. The Gibbs free energy of unfolding (ΔG) quantifies the effects of non-synonymous variations on protein stability and is influenced by factors such as temperature, pH, salt concentration, organic

solvents, urea, and other agents. Experimentally, the difference in unfolding free energy between mutant and wild-type proteins (ΔΔGu) serves as the primary metric. The value ΔΔG = ΔG3 - ΔG2 denotes the difference in folding free energy between wild-type and mutant proteins. The sign of ΔΔG indicates whether a mutation increases (ΔΔGu>0) or decreases (ΔΔGu<0) protein stability is used to quantify nonsynonymous variants.

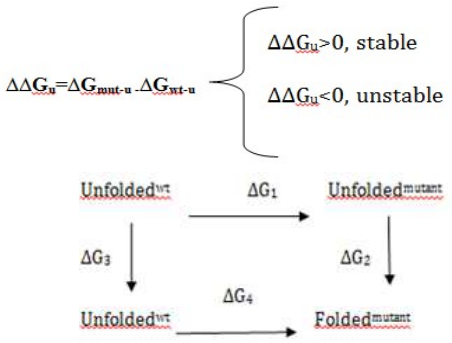


Fig.1 Representation of changes in Gibbs free energy between wild and mutant type Protein

Circular dichroism, differential scanning calorimetry, fluorescence spectroscopy employing heat, and denaturant unfolding methods have all been used in the past to test protein stability [7]. Although accurate, experimental methods to compare protein changes between mutant and wild-type proteins are time and financially intensive, various computational techniques have been created for evaluating and forecasting how single-point mutations affect protein stability [8]. The majority of them focus on either quantifying the ΔΔG values or identifying the effect (stabilizing/destabilizing/neutral). Despite the wide variety of computational methods for estimating protein stability after mutation, the following significant issues remain unresolved:

1) The dearth of thermodynamic measurements for proteins offered by the available databases;
2) Biases in the creation of predictive techniques brought on by failing to take into account the anti-symmetry of G values between mutant and native protein forms;
3) The best practices for providing a fair assessment of prediction performance.

A further goal of the current review is to offer a viewpoint on employing these prediction algorithms for specific biological applications in addition to highlighting these important topics. With the aid of these techniques, medical professionals could choose the best treatments and develop novel immunotherapeutic agents based on variations of mutant related diseases.

2. Protein Stability Data and Standard Datasets

Experimental and computational approaches currently offer only a limited understanding of protein folding and adaptation, as noted by Pucci and Rooman [9]. This constraint complicates the comparison of prediction methodologies. While VariBench and VariSNP provide numerous benchmark datasets for variation interpretation [10], few resources focus specifically on protein stability research. Historically, ProTherm [11] has served as the primary repository for protein stability data, compiling thermodynamic measurements such as Gibbs free energy change, enthalpy change, heat capacity change, and transition temperature for both wild-type and mutant proteins.

The previous version (5.0) of ProTherm contained approximately 17,000 entries from 771 proteins, with thermodynamic data reported for 7014 wild-type proteins, 8202 mutant proteins, 77 double mutants, and one triple mutant. This database is currently inaccessible and contains several discrepancies [10]. Reliable datasets require extensive manual screening and curation. Consequently, ProTherm has been used to generate numerous curated benchmark datasets, with the most frequently used listed in Table 1. Despite these efforts, there remains a substantial need for additional experimental data, particularly for stabilising mutations, which represent only about 30% of the datasets. The Protein Data Bank (PDB) [34] contained approximately 167,000 protein structures as of July 2020; however, available datasets for stability studies remain limited. The inherent variability in ΔΔG measurements present a significant limitation. When experiments are not conducted under identical conditions and methods, ΔΔG values for the same mutation may differ. Therefore, prediction methods must address uncertainties in ΔΔG measurements and account for multiple experimental values for the same mutation.

Table 1: Datasets and Subsets used from ProTherm Database [78]

Dataset	Total Variants (Proteins)	Stabilizing Variants (Pro	Destabilizing Variants (Prot	Supplementary Details

		teins)	eins)	
Q3421 [4]	3,421 (148)	763 (114)	2,658 (131)	Unique Variants/ Averaged ΔΔG
VariBench[14]	1,564 (89)	436 (70)	1,128 (78)	Unique Variants
VariBench3D[14]	1,423(79)	382 (60)	1,041 (68)	Variants with available structures from[14]
S1859 [15]	1,859 (64)	583 (48)	1,276 (55)	Replicated Variants/ Averaged ΔΔG
Fold-X [19]	964 (38)	110 (25)	854 (36)	Unique Variants
S238 [20]	238 (25)	45 (16)	193 (20)	Unique Variants/ Subset of S1948
S2648 [20]	2,648 (131)	602 (96)	2,046 (118)	Unique Variants/ Averaged ΔΔG
S350 [20]	350 (67)	90 (35)	260 (57)	Unique Variants/ Subset of S2648
S1615 [24]	1,615 (41)	449 (35)	1166 (35)	Unique Variants
S388 [24]	388 (17)	48 (12)	340 (15)	Unique Variants/ Physiological Conditions
S1948 [25]	1,948 (58)	592 (45)	1,356 (50)	Replicated Variants
p53 [27]	42 (1)	11 (1)	31 (1)	One Protein
S1925 [27]	1925 (55)	582 (42)	1,343 (48)	Replicated Variants
S2156 [27]	2,156 (84)	472 (61)	1,684 (68)	Unique Variants/ Averaged ΔΔG
Broom2017[32]	605 (58)	147 (37)	458 (54)	Unique Variants/ Background Variants
PTmul[34]	914 (90)	310 (57)	604 (68)	Unique Variants/ Multiple Variants
Cao Test [37]	276 (37)	79 (21)	197 (35)	Replicated Variants
Cao Train [37]	5,444 (204)	1,233 (150)	4,211 (185)	Replicated Variants
S3366[41]	3,366 (130)	836 (103)	2,530 (110)	Unique/Single and Multiple Variants
S3568 [45]	3,568 (154)	947 (110)	2,621 (138)	Replicated Variants
S630 [45]	630 (39)	467 (26)	163 (32)	Replicated Variants
Ssym[51]	684 (15)	342 (10)	342 (13)	Unique/Symmetric Variants
S1676 [75]	1676 (95)	453 (53)	1,223 (62)	Unique Variants/ Averaged ΔΔG

Some studies analysed selected subsets of values collected under specific experimental conditions (such as S388) or used averages or weighted averages (such as S2156 and S2648), while others used all ΔΔG values for training and testing. The resulting variability is significantly greater than the average ΔΔG distributions in major databases [15]. To ensure meaningful comparisons, the same datasets should be used across different methods.

A possible approach to generating new ΔΔG data for simulations exists. MD is a powerful tool for studying protein conformational changes (16–18). However, this requires many simulations under various conditions, such as different force fields, due to many experimental variations. The ΔΔG values are lower than the energy of a single hydrogen bond. Achieving this demands highly accurate force fields and careful MD parameter tuning. As computational power and force field accuracy improve, MD simulations may become valuable for augmenting data in this field.

3. Computational Techniques to Predict Protein Stability Variations

Structure-based features include secondary structure, solvent accessibility, hydrogen bonding patterns, and residue interactions.

Sequence-based features encompass amino acid composition, secondary structure arrangement, protein domains and motifs, surface properties, and evolutionary conservation. The presence of hydrophobic amino acids within the protein core typically enhances stability.

Energy-based features represent the balance of forces governing protein unfolding, including hydrophobic and electrostatic interactions, hydrogen bonds, van der Waals forces, solvent effects, and solvation.

Molecular-based features characterize factors that influence protein stability and folding, such as hydrophobic interactions, hydrogen bonds, disulfide bonds, salt bridges, van der Waals forces, solvent-exposed surfaces, and the protein folding energy landscape.ding energy landscape.

While some methods focus on specific feature subsets, most current approaches integrate all the criteria described above. Early protein stability prediction techniques employed force fields based on physical free energy functions from molecular mechanics [16], often in combination with molecular dynamics or Monte Carlo simulations [17-18]. Force fields estimate the potential energy of molecular systems using bonded and non-bonded interaction terms. These methods are computationally intensive, which restricts their application to a limited number of protein mutations. The introduction of problem-specific, physics-based approaches, such as those utilizing empirical energy functions [13, 19-20], has enabled the calculation of interaction propensities between atoms from collections of known protein structures using Maxwell-Boltzmann statistics. These functions are also referred to as statistical potentials, empirical potentials, knowledge-based potentials, and scoring functions. Empirical potentials use geometrical descriptors and experimental data from known protein structures, distinguishing them from force-field techniques. Grounded in statistical physics, these models balance computational accuracy with the free energy function [19], a widely used method that integrates electrostatic and van der Waals energies with hydrogen bond and solvation contributions. Model parameters are fitted to experimental data to predict mutation-induced changes in folding free energy. FoldX represents a mutation-rotamer-based approach, which permits side-chain conformational changes while maintaining a stable backbone.

The Molecular Mechanics/Poisson Boltzmann Surface Area (MM/PBSA) and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) approaches constitute another class of traditional linear $\Delta\Delta G$ predictors [21, 22]. These methods combine molecular mechanics and solvation energies by utilizing ensembles of conformations generated from molecular dynamics simulations. Recent approaches, such as SAAFEC [23], employ multiple linear regression to integrate MM/PBSA components with knowledge-based biophysical features. The objective is to characterize structural changes resulting from mutations and to elucidate the physical basis of folding free energy changes, while accurately predicting $\Delta\Delta G$ values. MM/PBSA-based algorithms generally achieve good prediction performance. To predict $\Delta\Delta G$, data-driven computational tools employing regression or machine learning techniques, such as Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Network (ANN), have been explored. In addition to the methods listed above, the most widely used and recent techniques are summarized in Table 2. These approaches are trained on datasets of proteins and their mutations, for which $\Delta\Delta G$ values have been empirically measured. The underlying relationships are modeled during the learning process; therefore, machine learning approaches do not require explicit understanding of the target function. By uncovering previously unknown patterns, linkages, and dependencies that knowledge-based models may overlook, these methods offer greater flexibility for incorporating new features. It's more flexible for adding new features.

3.1 Single point mutation vs Multisite Variants

A single point mutation alters the DNA sequence, substituting one amino acid for another in a protein. This can affect protein stability, causing misfolding, aggregation, or degradation. Multisite mutations involve changes at multiple positions within a protein sequence, and their effects on stability depend on the specific mutations and their locations.

Multipoint mutations can affect protein stability by disrupting structural elements,

altering interactions, changing hydrophobicity or charge distribution, or introducing new interactions.

3.2. Multi-point mutations

Anticipating how proteins will hold up under a flurry of mutations is a formidable challenge. Often, a single mutation is not enough to achieve the desired stability, so protein engineers turn to combinations of mutations. However, predicting the effects of these multiple changes is far more complex than dealing with just one, since it requires unravelling a web of intricate combinatorial interactions.

Moreover, experimental data on DDG for multiple-point mutations are scarce compared to those for single-point mutations. This scarcity, combined with the tangled nature of combinatorial effects, makes accurate predictions even more elusive. As a result, most current methods focus solely on single-point changes.

Generally, mutations that are close together—whether in sequence or structure—tend to interact more strongly than those spaced farther apart. When mutations are distant, one common approach is to estimate the DDG for each change separately, sum the results, and use the sum as an approximation of the overall effect. Linear algorithms such as FoldX and DDGun naturally adopt this additive method. Extending single-point prediction techniques to handle multiple mutations is straightforward, but another strategy is to average the input features across all substitutions, as seen in PROTS-RF and Prethermut. However, when mutations work together in unexpected ways, simple averaging may fall short.

Table 2: Prediction of ΔΔG values using the most employed methods[78]

Method	**Features required& Ref**	**Algorithm used**	**Anti-Symmetry**	**Multiple Mutations**
FoldX	3D[19]	linear regression	NO	YES
MUpro	3D and 1D [76]	ANN and SVMs	NO	NO
CUPSAT	3D[13]	combined statistical potential	NO	NO
I-Mutant (3.0/ 2.0)	3D or 1D[24, 2526]	SVM	NO	NO
iPTREE – STAB	1D[15]	Decision trees and Adaptive boosting	NO	NO
AUTO-MUTE (2.0)	3D[56]	Random Forest	NO	NO
Prethermut	3D[41]	SVM and Random Forest	NO	YES
POPMUSIC(3.1/ 2.1)	3D[20]	ANN and Statistical potential	NO	NO
Pro-Maya	3Dand 1D[77]	Random Forest regression	NO	NO
PROTS-RF	3D and 1D[36]	Random Forest	YES	YES
iStable (2.0)	3D or 1D [45,46]	*meta*-predictor	NO	NO
NeEMO	3D[79]	ANN regression	NO	NO
DUET	3D[27]	*meta*-predictor (SVM regression)	NO	NO
mCSM	3D[80]	Gaussian process	NO	NO

		regression and Random forest		
EASE-MM	1D[75]	SVM	NO	NO
INPS (3D)	3D or 1D [35,42]	SVM regression	YES	NO
STRUM	3Dand 1D[4]	Gradient boosting regression trees	NO	NO
ELASPIC	3Dand 1D[3]	Stochastic Gradient Boosting of Decision Trees	NO	YES
SAAFEC	3D[23]	Molecular Mechanics Poisson-Boltzmann	NO	NO
MAESTRO (web)	3D[5]	Multi-agent Prediction (ANN+ SVM+ linear regression)	NO	YES
SDM	3Dand 1D[33]	environment-specific substitution tables (ESSTs)	YES	NO
TML-MP	3Dand 1D[1]	Gradient boosting	NO	NO
ThreeFoil	3Dand 1D[32]	*meta*-predictor using11tools	YES	NO
DynaMut	3D[47]	meta-predictor based on Normal Mode Analysis	YES	NO
DDGun	3D or 1D[34]	Linear combination of features	YES	YES
DeepDDG	3D and 1D[37]	ANN	YES	NO
ProTstab	1D	Gradient boosting of regression trees	NO	NO

These approaches offer substantial speed improvements over earlier methods because, once a model is constructed from the data, predictions are generated instantaneously. Most techniques integrate non-energy variables into the scoring function by employing statistical potentials derived from environmental propensities, substitution frequencies, and observed correlations between adjacent residues in protein structures. Access to extensive and diverse experimental training data is crucial for accurate estimation; however, it also heightens the risk of overfitting. Additionally, interpreting the results in physical terms can be challenging.

The limitations of these methods are determined by the availability of data. Structure-based methods, such as those listed in Table 2, are inapplicable when the three-dimensional structures of proteins are unavailable. Sequence-based prediction tools (designated as 1D in Table 2) address this limitation by predicting changes in folding free energy directly from the amino acid sequence using machine-learning techniques. These tools are advantageous because they do not require a three-dimensional structure. In general, methods that utilize the wild-type protein's three-dimensional structure, such as I-Mutant [24-26], PoPMuSiC [20], or DUET [27], demonstrate greater effectiveness than sequence-based approaches [28]. Optimising protein geometry constitutes a combinatorial NP-hard problem, and no polynomial-time or polynomial-space algorithms are currently available to address it [30]. Consequently, all available tools depend on approximations and heuristics. Predicted structures may serve as models when experimental structures are unavailable [31]. Existing methods have undergone

comprehensive evaluation using both experimental structures and the modeling techniques employed to generate predicted structures [32].d based on both experimental structures and the modeling techniques used to generate predicted structures [32].

3.3. Anti-symmetry Analysis (ΔΔG)

The ΔΔG anti-symmetry, which serves as a physical principle at the core of each thermodynamic transformation, is a significant trait that only a small number of predictors take into account.

The relative concentrations of A and B (i.e. [A] and [B]) at equilibrium (and at constant temperature and pressure) are characterized by the following Gibbs free energy difference given a chemical (or conformational) transformation from A to B:

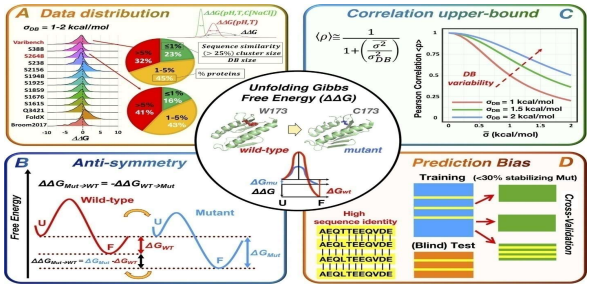


Fig.2. Protein stability variations upon mutations [78].

Machine learning techniques now consider reverse variations with inverse DDG signals during training. Only the methods listed in Table 2—SDM [33], DDGun(3D) [34], ThreeFoil [32], PROTS-RF [36], INPS [35], and DeepDDG [37]—address this aspect. DDGun(3D) applies anti-symmetric scoring functions, including evolutionary, hydrophobicity, and residue contact scores. PROTS-RF was the first to assess effectiveness by also considering reverse mutations. The simplest way to measure anti-symmetry bias is to compare the Pearson correlation between predictions for a variation set and its reverse. Correlation values further from 1 indicate a greater anti-symmetry bias (Table 3). Usmanova et al. [38] showed that to meet the anti-symmetry constraint, methods must generate the mutant's lowest energy structure from the native one and vice versa, which can be difficult due to unique program properties. For instance, when a residue is mutated, FoldX [19] preserves the side-chain and backbone of all residues except the neighbors. As a result, the modelled structure of the mutant (B) is expected to be less stable than the original.

Table 3: Performance of Method with their Correlation [78]

Method	Training Data	Test data	Correlation	Anti-symmetry
FoldX [19]	S339	S625	0.82	Biased (—0.38)
MUpro [76]	S1615	S1615, S388	0.13 – 0.76	Biased (—0.02)
CUPSAT [13]	S1538, S1603	S1538, S1603	0.55 – 0.78	Biased (—0.54)
I-Mutant(3.0/2.0) [25,26]	S1948	S1948	0.62 – 0.71	Biased (0.02)
iPTREE-STAB [40]	S1859	S1859	0.7	not evaluated
AUTO-MUTE (2.0) [56]	Subsets of S1948, S1615, S388, S1791, S1396, S2204	Subsets of S1948, S1615, S388,S1791, S1396	0.74 – 0.79	Biased (—0.06)
Prethermut [41]	S3366, S2156	S3366, S2156	0.67 – 0.72	not evaluated
POPMUSIC(3.1/2.1) [20]	S2648	S2648	0.63 – 0.79	Biased (—0.29); unbiased version POPMUSICsym (—0.77)
Pro-Maya [77]	S2156	S2156 + D180, D140 (27and 19 proteins)	0.62 – 0.86	Unbiased
iStable(2.0) [45,46]	S2648, S1948 v2.0: S3528	M1311, M1820	0.85 v2.0 : 0.67 – 0.71	Biased (—0.05) v2.0: not evaluated

DUET [27]	S2648	p53,S350	0.71 – 0.82	Biased (—0.21)
mCSM [80]	S2648, S1925	S350	0.51 – 0.82	Biased (—0.26)
EASE-MM [75]	1676 mutations from S1948	S543, S236	0.51 – 0.59	not evaluated
INPS(3D) [35,29]	S2648, p53	S2648, p53	0.53 –0.7	Unbiased
STRUM [4]	Q3421	S2648, S350, Q306 (subset of S2648)	0.4– 0.8	Biased (0.34)
ELASPIC [3]	S3463 (159 proteins)	S2636 (134 proteins), S2104 (79 proteins)	0.77	not evaluated
SAAFEC [23]	983 mutations fromProtherm	983 mutations from Protherm	0.61	not evaluated
MAESTRO(web) [5]	S2648, S1925 (from S1948), S1765, S2244	S2648, S350	0.63 – 0.76	Biased (—0.34)
SDM(2) [33]	None	S2648, S350, p53, S140	0.52 – 0.63	Biased (—0.75)
TML-MP [1]	S2648 S350, M233	2648 S350, M233	0.54 – 0.82	not evaluated
ThreeFoil [32]	Broom2017	Broom2017	0.73	not evaluated
DynaMut [47]	S2648	S351	0.58 – 0,70	Biased
DDGun [34]	Untrained	S2648, Ssym, p53, Myoglobin	0.45 – 0.71	Unbiased
DeepDDG [37]	Cao Train	Cao Test	0.56	not evaluated
ProTstab [42]	Varibench	Varibench	0.79	not evaluated

4. Issues and Best Practices in Prediction Assessment

4.1 Accuracy of prediction

Two samples are considered to have similar DDG values if their differences fall within the experimental error or uncertainty. However, since each dataset demonstrates inherent variability, it is necessary to determine whether an upper bound for prediction performance based on the dataset's standard deviation (and the uncertainty (r). The optimal predictor is defined by its similarity to another set of empirically determined DDG values, using measured DDG protein changes as a reference.

The underlying concept is that, when beginning with a normally distributed dataset in which the variability is of the same order of magnitude as the experimental uncertainty, the theoretical upper bounds for prediction performance may be lower than anticipated. This framework enables the derivation of an upper bound for the Pearson correlation coefficient as a function of both the noise and the distribution of the DDG data.

Based on the average data uncertainty (r) and the dataset standard deviation (rDB), previous studies [42, 43] have provided theoretical upper bounds for the Pearson correlation. For example, the widely used S2648 and VariBench datasets have an rDB of 2, yielding a Pearson correlation coefficient of 0.8 (an upper bound) and a root mean square error between experimental and predicted values of 1 kcal/mol (a lower bound). These findings indicate that methods reporting Pearson correlation values above 0.8 when using these datasets are likely overfitting during training.

4.2 Liability of Sequence Distinctiveness

The size of the training dataset directly impacts the reliability of machine learning techniques. It is essential to train and validate these methods using high-quality, consistent experimental data. Table 3 highlights the key elements shared by the approaches listed in Table 2. Careful consideration should be given to both the size and distribution of training data. Most datasets, with only a few hundred to a thousand cases, may be too small to generate effective descriptors for learning. Limited variability in training data can hinder the

generalization of prediction systems. Overrepresentation of certain descriptors may bias model weights and overlook broader predictive features. Additionally, many current approaches ignore similarities between training and test sets, leading to overfitting.

Proteins suitable for maintaining their specific tasks have been chosen by evolution. While evolutionary pressure affects protein structure, significant sequence variation can exist, with even highly dissimilar sequences adopting similar shapes. Homologues are proteins with different sequences but shared evolutionary or functional backgrounds. Quantifying these relationships is essential to prevent information leakage between data splits. Sequence identity, which measures the proportion of exact amino acid matches between aligned protein subsequences, is the primary metric used. For example, applying a 25% sequence identity threshold ensures no protein pairs in the training and test sets share more than 25% identical amino acids. Pires et al. [27] provide a clear illustration of the problem with comparable training and testing. When a random sample of variations is used, the Pearson correlation for the mCSM approach ranges between 0.7 and 0.8. When cross-validation uses "per site-" and "per protein-" clustering, the correlation falls to 0.54-0.51. Independent of residue substitution, different protein locations tend to be more or less sensitive to mutations [44]. As a result, using the same site or protein for training and testing yields inaccurate and overconfident results. Furthermore, methods developed under these conditions tend to generalize poorly. For these reasons, a straightforward random dataset split should be avoided.

The issue of similarity between training and test sets is even more pronounced with meta-predictors. Recently, stacked approaches such as DUET [27], iStable(2.0) [45,46], and Dynamut [47] have been developed by combining multiple predictors. Meta-predictors often outperform individual methods, even with simple majority voting, as each method has unique strengths and weaknesses. However, reported performance can be misleading. For example, ThreeFoil [32] aggregates predictions from 11 publicly available tools, all trained on different datasets derived from ProTherm.

The meta-predictor was developed using a training set of 605 mutations, with each tool's predictions weighted according to their performance relative to others predicting similar changes. The dataset was split into two halves for cross-validation to determine the Matthews Correlation Coefficient (MCC) values. One-half was used to calculate MCC-based weights, and the other half to test overall performance. However, several predictors had already used 60% of the proteins in this dataset during their own training, raising concerns about the similarity between the training and test sets as previously discussed.

Like Bioconcerns about similarity, DUET [27] and Dyna-mut integrate mutations into protein stability and dynamics to enhance prediction reliability. DUET combines SDM and mCSM [80], both trained on the S2648 dataset, which was also used by Dynamut [47] for training, cross-validation, and selecting 351 'blind' variants for evaluation. The authors of iStable 2.0 [45] introduced separate datasets for training and testing.

Of the 630 variants in the test set, 442 come from the same proteins as those in the training set. The training and test sets contain 703 and 77 recurring mutations, respectively. Only three proteins, representing 81 mutations in S630, can be used safely for testing based on sequence similarity. Therefore, it is essential to evaluate these techniques using data that does not share sequence similarity with the training set proteins.

4.3 Inconsistent, erroneous experimental datasets and Anti-symmetry

When training datasets are skewed, with some categories far outnumbering others, predictions can easily go astray. Imagine a dataset where 80% of the mutations are destabilising: the model will likely label most changes as destabilising, regardless of their true nature. This kind of imbalance throws off both classification and regression tasks. To get a fair measure of accuracy in these situations, metrics such as the Matthews correlation coefficient (MCC) and the F1 score, which balance precision and recall, are especially valuable.

Another key challenge is the anti-symmetric nature of free energy changes, which can trip up even the best algorithms. Many predictors stumble here, introducing bias that recent studies have highlighted. To tackle this, new datasets such as Ssym, Usmanova-DB, and Fang-DB were created to include both a variation and its reverse, such as swapping A for B and then B for A within the same protein. Interestingly, only INPS and PopMusic have proven reliable enough to respect this anti-symmetry. To spot technique bias, researchers

look at the Pearson correlation between predictions for a variation and its inverse. Table 3 highlights how different methods fare on this front, with a perfect score being a Pearson correlation of -1.

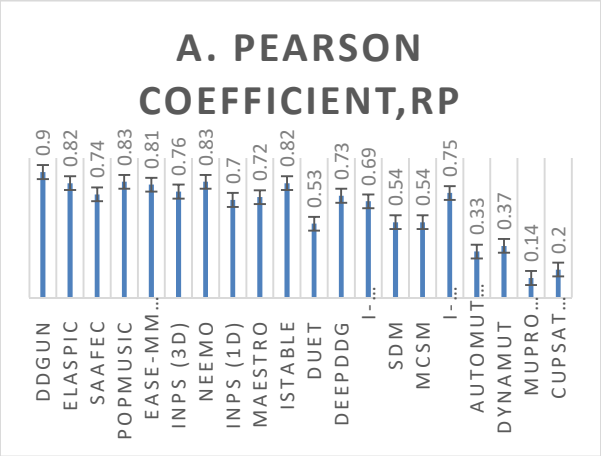
4.4 Discussion on performance study

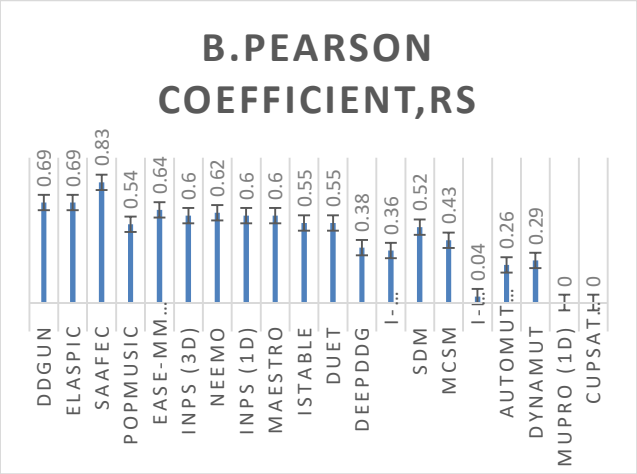
The Critical Assessment of Genome Interpretation (CAGI) is a community experiment aimed at fairly assessing the computational methods for genome interpretation [53].

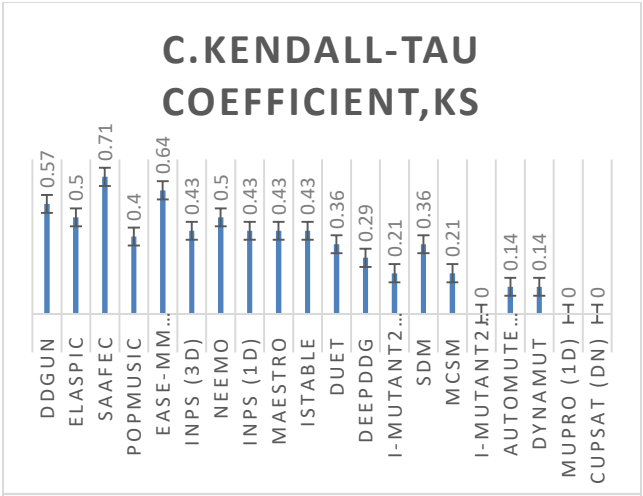
Table 4: Assessing cutting-edge methods across 8 variants of the CAGI FXN challenge

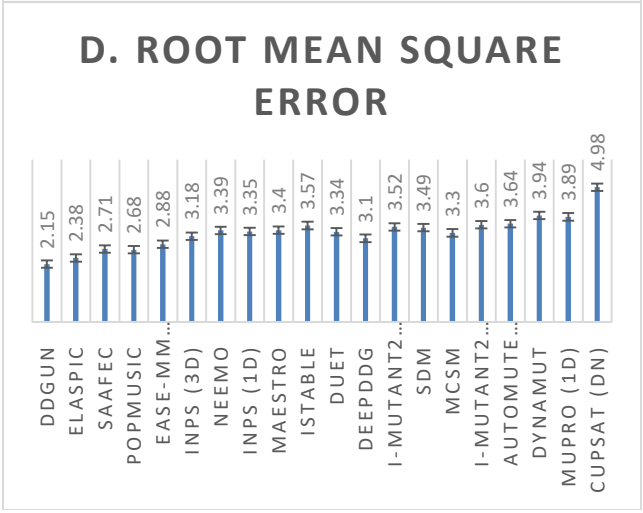
Method	r_p	r_s	Ks	RMSE	MAE	Q2 Accuracy	Matthews correlation coefficient	Area under the ROC Curve
DDGun[*34]*	0.90	0.69	0.57	2.15	1.93	0.70	0.45	0.80
ELASPIC [3]	0.82	0.69	0.50	2.38	1.71	0.80	0.60	0.93
SAAFEC[23]	0.74	0.83	0.71	2.71	2.08	0.83	0.75	0.93
PopMusic[20]	0.83	0.54	0.40	2.68	2.00	0.90	0.78	0.80
EASE-MM(1D)[75]	0.81	0.64	0.64	2.88	2.22	0.70	0.45	0.80
INPS(3D)[35]	0.76	0.60	0.43	3.18	2.36	0.80	0.60	0.87
NeEMO[79]	0.8	0.6	0.5	3.39	2.44	0.70	0.45	0.87
	3	2	0					
INPS(1D)[35]	0.70	0.60	0.43	3.35	2.52	0.80	0.60	0.87
MAESTRO[5]	0.72	0.60	0.43	3.40	2.48	0.70	0.45	0.67
iStable[45,46]	0.82	0.55	0.43	3.57	2.73	0.70	0.45	0.80
DUET[27]	0.53	0.55	0.36	3.34	2.46	0.70	0.45	0.73
DeepDDG [37]	0.73	0.38	0.29	3.10	2.43	0.53	0.07	0.67
I-Mutant2(1D) [25,26]	0.69	0.36	0.21	3.52	2.68	0.80	0.60	0.67
SDM [33]	0.54	0.52	0.36	3.49	2.73	0.63	0.26	0.67
mCSM[80]	0.54	0.43	0.21	3.30	2.49	0.37	-0.26	0.60
I-Mutant2(3D) [25,26]	0.75	0.04	-0.05	3.60	2.55	0.63	0.35	0.58
AUTOMUTE (TR) [56]	0.33	0.26	0.14	3.64	2.67	0.53	0.07	0.73
DynaMut[47]	0.37	0.29	0.14	3.94	2.89	0.60	0.29	0.67
MUpro (1D) [76]	0.14	-0.10	-0.07	3.89	2.80	0.70	0.45	0.53

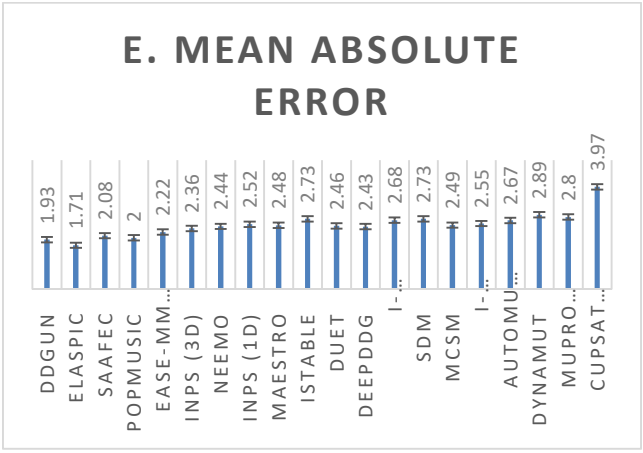
CUPSAT(DN) [13]	0.20	-0.12	-0.14	4.98	3.97	0.50	0.00	0.20

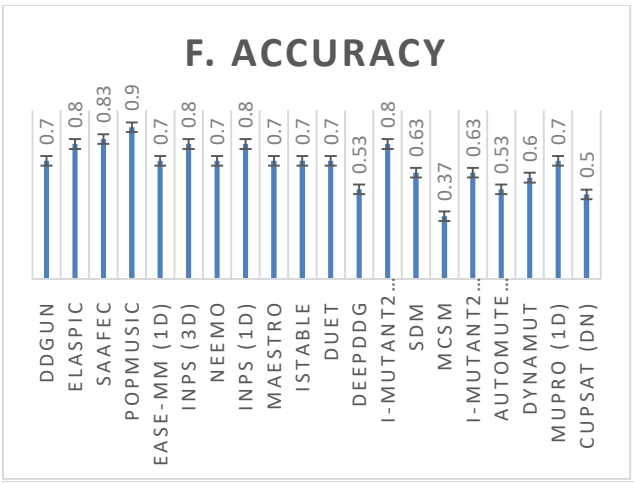


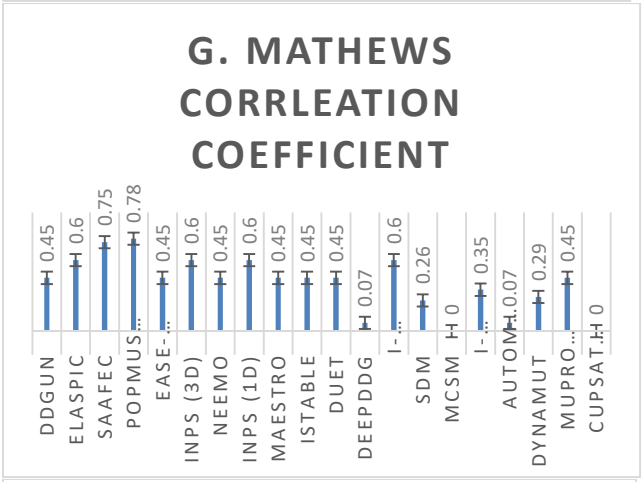












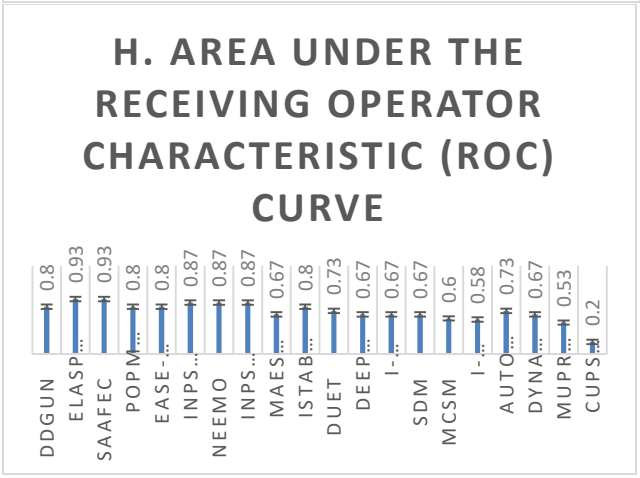


Fig. 3 A and B. Methods vs Perason Coefficients, C. Methods vs Kendall Tau coefficient,
D and E. Methods vs Errors, F. Methods vs Acciracy, G. Methods vs Mathews Correlation Coefficient, H. Methods vs Area Under ROC.

The most recent Critical Assessment of Genome Interpretation (CAGI 5) included measurements of the unfolding free energy for several variants of the highly conserved protein frataxin (FXN), which is essential for cellular iron homeostasis in prokaryotes and eukaryotes [53]. These results were used to compare the unfolding free energies of the variant and wild-type proteins at zero denaturant (DDGH2O) concentration.

The experimental dataset included eight amino acid changes [54] and was used to assess the performance of web-based tools in predicting DDGH2O values and classifying variants as destabilising or not-destabilizing [55]. Three metrics were used to evaluate classification performance: accuracy, Matthews Correlation Coefficient, and Area Under the Receiver Operating Characteristic Curve (AUC). Five metrics assessed regression predictions: Pearson and Spearman Correlation Coefficients, Kendall-Tau Coefficient, Root Mean Square Error, and Mean Absolute Error. A prediction threshold of —1 kcal/mol was set to distinguish destabilising (DDG —1 kcal/mol) from not-destabilizing (neutral, DDG 1 kcal/mol) variants [56].

This classification shows that five mutations destabilise the protein structure, while three do not. Each protein variant was submitted to the respective method's website for evaluation. Results indicate that DDGun and PopMusic, both optimised for the anti-symmetric property, are among the most effective approaches for predicting DDG in this challenge (Table 4; see Fig. 3). ELASPIC and INPS3D also performed well. Although the FXN dataset contains only eight variants, this is the first attempt to use a blind test set to validate DDG prediction methods.

5. Applications of Protein Stability Predictions in Bioinformatics

Sickle-cell anaemia is an autosomal recessive disorder caused by the E6V amino acid substitution in the beta-chain of human haemoglobin [59], where glutamic acid (E) is replaced by valine (V) at position 6. In 1949, it became the first disease linked to sequence variations and structural changes [57,58]. Today, databases such as the Human Gene Mutation Database (HGMD [60]), Online Mendelian Inheritance in Man (OMIM [61]), and the Catalogue of Somatic Mutations in Cancer (COSMIC [62]) collect thousands of single amino acid variants that are responsible for or associated with diseases.

Nevertheless, the pathogenic variants documented in these databases account for only

a small proportion of all potential disease-causing mutations present in the global human population.

Introducing mutations that modify protein stability presents significant opportunities for advancing precision medicine. Studies indicate that decreased protein stability is frequently a primary cause of inherited disorders, as mutations impacting thermodynamic stability are strongly associated with pathogenicity.

A deeper understanding of the relationships among protein sequence, structure, folding, and stability may facilitate patient-specific diagnoses and the development of novel therapeutic strategies. Casadio et al. [63] performed the first large-scale analysis of how alterations in protein stability contribute to human diseases by examining the Human Proteome [2]. Their findings demonstrated that changes in protein stability can be associated with disease at the proteome level by evaluating the probability of each amino acid mutation being disease-related. These probability indices were calculated using a dataset of 17,170 single amino acid variants in 5,305 proteins from Uni-ProtKB (version 2010_04), dbSNP (build 132), OMIM, and ProTherm [11]. Computational tools such as mCSM and DUET have been employed to estimate the effects of stability changes in various contexts, including cancer-related genes in the COSMIC database [64], inhibition of inosine-5'-monophosphate dehydrogenase, isoniazid and rifampicin resistance in Mycobacterium tuberculosis [65], phosphodiesterase mutations implicated in cancer and retinitis pigmentosa [66], and protein presenilin 2 associated with familial Alzheimer's disease [67].

Predicting how protein mutations affect function and distinguishing 'driver' mutations that cause disease or confer drug resistance from 'driver' mutations that are neutral or beneficial remains challenging in medical applications. This requires understanding the effects of missense mutations on protein function and mapping genetic variations to 3D protein structures [68]. Mutations at protein-protein interfaces tend to have greater global effects than those elsewhere on protein surfaces [2,68]. Pathogenic missense mutations often alter protein stability or cause improper folding. Protein stability is frequently linked to functional activity, and most missense mutations are harmful because they reduce stability [49]. However, some disorders result from missense mutations that increase protein stability [69]. The mutant kinase EGFR exemplifies both types of disease-causing modifications and highlights the complexity of treatment. The oncogenic Leu858 driver mutation [70], which stabilises the alpha C-helix in its active conformation, is present in 41% of lung cancer-associated EGFR mutations. The T790M mutation in EGFR stabilises the hydrophobic R-spine and destabilises the inactive conformation, leading to constitutive kinase activity. These mutations alter the equilibrium between conformational states. In summary, driver mutations may destabilise the inactive state, stabilise the active state, or exert both effects. These conformational changes, which affect the formation of asymmetric or symmetric dimers, are reflected in free-energy differences corresponding to the depths of the energy minima [71]. free energy changes as differences in the depths of energy minima [71].

Such information directly informs treatment strategies for lung cancer. For instance, the EGFR inhibitor Gefitinib is effective only in patients with specific genetic alterations. Immunogenomics is another area that may benefit from the design of stability-altering mutations. Immunotherapies have achieved notable success in treating various cancers. Recent strategies involve the adoptive transfer of autologous T cells genetically engineered with Chimeric Antigen Receptors (CAR) or T Cell Receptors (TCR) to recognize target peptides presented by Major Histocompatibility Complex class I (MHC-I). Consequently, computational methods are being developed to predict the immunogenicity of peptide-MHC-I complexes and to identify suitable candidates for CAR-T and TCR-T cell therapies.

Protein complex stability is recognized as a critical determinant in predicting the immunogenicity of peptide-MHC-I complexes (see, for example, NetMHCstab [72] and NetMHCstabpan [73]). These computational techniques are expected to expedite the initiation of patient-specific therapies. The influence of protein fold stability on immunogenicity has broad potential applications, including vaccine development. Early examples of stability-optimized vaccines have been reported [74], and substantial progress in protein design for novel therapeutics is anticipated in the near future.

6. Conclusion and Outlook

Accurate prediction of protein stability changes in response to variation is essential for

protein design and precision medicine. While current approaches remain suboptimal, they have demonstrated sufficient performance to support experimental studies [55]. However, several key challenges must be addressed to further enhance predictive accuracy, including:

1. Enhancing the quality and size of existing datasets by incorporating additional rigorously controlled experimental data;
2. Developing methodologies that are inherently antisymmetric, such that DDG (A to B) equals negative DDG (B to A);
3. Ensuring low sequence identity between training and testing sets during model development in deep learning neural network methodologies. Similarity-free cross-validation folds for the most relevant datasets are provided in the Supplementary Materials to support the scientific community in this effort.

If these indicators are taken into account, the next generation of predictors can achieve more reliable, more accurate predictions.

References

1. Cang Z, Wei GW. Analysis and prediction of protein folding energy changes upon mutation by element specific persistent homology. Bioinformatics 2017;33:3549–57. https://doi.org/10.1093/bioinformatics/btx460.
2. Martelli PL, Fariselli P, Savojardo C, Babbi G, Aggazio F, Casadio R. Large scale analysis of protein stability in OMIM disease related human protein variants. BMC Genomics 2016;17(Suppl 2):397. https://doi.org/10.1186/s12864-016-2726-y
3. Witvliet DK, Strokach A, Giraldo-Forero AF, Teyra J, Colak R, Kim PM. ELASPIC web-server: proteome-wide structure-based prediction of mutation effects on protein stability and binding affinity. Bioinformatics 2016;32:1589–91. https://doi.org/10.1093/bioinformatics/btw031.
4. Quan L, Lv Q, Zhang Y. STRUM: structure-based prediction of protein stability changes upon single-point mutation. Bioinformatics 2016;32:2936–46. https://doi.org/10.1093/bioinformatics/btw361.
5. Laimer J, Hofer H, Fritz M, Wegenkittl S, Lackner P. MAESTRO–multi agent stability prediction upon point mutations. BMC Bioinf 2015;16:116. https://doi.org/10.1186/s12859-015-0548-6.
6. Hartl FU. Protein Misfolding Diseases 2017;86:21–6. https://doi.org/10.1146/annurev-biochem-061516-044518
7. Ó'Fágáin C. Protein stability: enhancement and measurement. Methods MolBiol 2017;1485:101–29. https://doi.org/10.1007/978-1-4939-6412-3_7.
8. Compiani M, Capriotti E. Computational and theoretical methods for protein folding. Biochemistry 2013;52:8601–24. https://doi.org/10.1021/bi4001529.
9. Pucci F, Rooman M. Stability curve prediction of homologous proteins usingtemperature-dependent statistical potentials. PLoSComputBiol 2014;10:. https://doi.org/10.1371/journal.pcbi.1003689e1003689.
10. Sarkar A, Yang Y, Vihinen M. Variation benchmark datasets: update, criteria,quality and applications. Database (Oxford) 2020;2020. https://doi.org/10.1093/database/baz117.
11. Kumar MDS, Bava KA, Gromiha MM, Prabakaran P, Kitajima K, et al. ProThermand ProNIT: thermodynamic databases for proteins and protein-nucleic acid interactions. Nucleic Acids Res 2006;34:D204–6. https://doi.org/10.1093/nar/gkj103.
12. Pitera JW, Kollman PA. Exhaustive mutagenesis in silico: multicoordinatefree energy calculations on proteins and peptides. Proteins 2000;41:385–97.
13. Parthiban V, Gromiha MM, Schomburg D. CUPSAT: prediction of protein stability upon point mutations. Nucleic Acids Res 2006;34:W239–42. https://doi.org/10.1093/nar/gkl190
14. Sasidharan Nair P, Vihinen M. VariBench: a benchmark database for variations. Hum Mutat 2013;34:42–9. https://doi.org/10.1002/humu.22204.
15. Huang L-T, Gromiha MM, Ho S-Y. iPTREE-STAB: interpretable decision tree based method for predicting protein stability changes upon mutations. Bioinformatics 2007;23:1292–3. https://doi.org/10.1093/bioinformatics/btm100.
16. Lazaridis T, Karplus M. Effective energy functions for protein structure prediction. CurrOpinStructBiol 2000;10:139–45. https://doi.org/10.1016/s0959-440x(00)00063-4.
17. Duan Y, Wang L, Kollman PA. The early stage of folding of villin headpiece subdomain observed in a 200-nanosecond fully solvated molecular dynamics simulation. Proc Natl AcadSci U S A 1998;95:9897–902. https://doi.org/10.1073/pnas.95.17.9897.
18. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. AccChem Res 2000;33:889–97. https://doi.org/10.1021/ar000033j.
19. Guerois R, Nielsen JE, Serrano L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J MolBiol2002;320:369–87. https://doi.org/10.1016/S0022-2836(02)00442-4.
20. Dehouck Y, Kwasigroch JM, Gilis D, Rooman M. PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. BMC Bioinf 2011;12:151. https://doi.org/10.1186/1471-2105-12-151.
21. Petukh M, Li M, Alexov E. Predicting binding free energy change caused by point mutations with knowledge-modified MM/PBSA method. PLoSComputBiol 2015;11:. https://doi.org/10.1371/journal.pcbi.1004276e1004276.

22. Wang E, Sun H, Wang J, Wang Z, Liu H, et al. End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. Chem Rev 2019;119:9478–508. https://doi.org/10.1021/acs.chemrev.9b00055.
23. Getov I, Petukh M, Alexov E. SAAFEC: predicting the effect of single point mutations on protein folding free energy using a knowledge-modified MM/ PBSA approach. Int J MolSci 2016;17:512. https://doi.org/10.3390/ijms17040512.
24. Capriotti E, Fariselli P, Casadio R. A neural-network-based method for predicting protein stability changes upon single point mutations.Bioinformatics2004;20(Suppl 1):i63–8. https://doi.org/10.1093/bioinformatics/bth928
25. Capriotti E, Fariselli P, Casadio R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. Nucleic Acids Res 2005;33. https://doi.org/10.1093/nar/gki375. W306–310
26. Capriotti E, Fariselli P, Rossi I, Casadio R. A three-state prediction of single point mutations on protein stability changes. BMC Bioinf 2008;9(Suppl 2):S6. https://doi.org/10.1186/1471-2105-9-S2-S6
27. Pires DEV, Ascher DB, Blundell TL. DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach. Nucleic Acids Res 2014;42:W314–9. https://doi.org/10.1093/nar/gku411.
28. Khan S, Vihinen M. Performance of protein stability predictors. Hum Mutat 2010;31:675–84. https://doi.org/10.1002/humu.21242.
29. Savojardo C, Martelli PL, Casadio R, Fariselli P. On the critical review of fivemachine learning-based algorithms for predicting protein stability changes upon mutation. Brief Bioinform 2019. https://doi.org/10.1093/bib/bbz168
30. Pierce NA, Winfree E. Protein design is NP-hard. Protein Eng 2002;15:779–82. https://doi.org/10.1093/protein/15.10.779.
31. Hurst JM, McMillan LEM, Porter CT, Allen J, Fakorede A, Martin ACR. The SAAPdb web resource: a large-scale structural analysis of mutant proteins. Hum Mutat 2009;30:616–24. https://doi.org/10.1002/humu.20898
32. Broom A, Jacobi Z, Trainor K, Meiering EM. Computational tools help improve protein stability but with a solubility tradeoff. J BiolChem 2017;292:14349–61. https://doi.org/10.1074/jbc.M117.784165
33. Pandurangan AP, Ochoa-Montaño B, Ascher DB, Blundell TL. SDM: a server for predicting effects of mutations on protein stability. Nucleic Acids Res 2017;45: W229–35. https://doi.org/10.1093/nar/gkx439
34. Montanucci L, Capriotti E, Frank Y, Ben-Tal N, Fariselli P. DDGun: an untrained method for the prediction of protein stability changes upon single and multiple point variations. BMC Bioinf 2019;20:335. https://doi.org/10.1186/s12859-019-2923-1.
35. Fariselli P, Martelli PL, Savojardo C, Casadio R. INPS: predicting the impact of non-synonymous variations on protein stability from sequence. Bioinformatics 2015;31:2816–21. https://doi.org/10.1093/bioinformatics/btv291
36. Li Y, Fang J. PROTS-RF: a robust model for predicting mutation-induced protein stability changes. PLoS ONE 2012;7:. https://doi.org/10.1371/journal.pone.0047247e47247.
37. Cao H, Wang J, He L, Qi Y, Zhang JZ. DeepDDG: predicting the stability change of protein point mutations using neural networks. J ChemInf Model 2019;59:1508–14. https://doi.org/10.1021/acs.jcim.8b00697
38. Usmanova DR, Bogatyreva NS, AriñoBernad J, Eremina AA, Gorshkova AA, et al. Self-consistency test reveals systematic bias in programs for prediction change of stability upon mutation. Bioinformatics 2018;34:3653–8. https://doi.org/10.1093/bioinformatics/bty340.
39. Montanucci L, Fariselli P, Martelli PL, Casadio R. Predicting protein thermostability changes from sequence upon multiple mutations. Bioinformatics 2008;24:i190–5. https://doi.org/10.1093/bioinformatics/btn166.
40. Huang LT, Gromiha MM. Reliable prediction of protein thermostability change upon double mutation from amino acid sequence. Bioinformatics 2009;25:2181–7. https://doi.org/10.1093/bioinformatics/btp370.
41. Tian J, Wu N, Chu X, Fan Y. Predicting changes in protein thermostability brought about by single- or multi-site mutations. BMC Bioinf 2010;11:370. https://doi.org/10.1186/1471-2105-11-370
42. Montanucci L, Savojardo C, Martelli PL, Casadio R, Fariselli P. On the biases in predictions of protein stability changes upon variations: the INPS test case. Bioinformatics 2019;35:2525–7. https://doi.org/10.1093/bioinformatics/bty979.
43. Benevenuta S, Fariselli P. On the upper bounds of the real-valued predictions. BioinformBiol Insights 2019;13. https://doi.org/10.1177/1177932219871263.
44. Single recorded mutation aids in predicting the effects of other mutations in the same amino acid site. Bioinformatics 2011;27:3286–92. https://doi.org/10.1093/bioinformatics/btr576.
45. Chen CW, Lin MH, Liao CC, Chang HP, Chu YW. iStable 2.0: Predicting protein thermal stability changes by integrating various characteristic modules. ComputStruct. Biotechnol J 2020;18:622–30. https://doi.org/10.1016/j.csbj.2020.02.021.
46. Chen CW, Lin J, Chu YW. iStable: off-the-shelf predictor integration for predicting protein stability changes. BMC Bioinf 2013;14(Suppl 2):S5. https://doi.org/10.1186/1471-2105-14-S2-S5

47. Rodrigues CH, Pires DE, Ascher DB. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. Nucleic Acids Res 2018;46:W350–5. https://doi.org/10.1093/nar/gky300
48. Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD. Bio3d: an R package for the comparative analysis of protein structures. Bioinformatics 2006;22:2695–6. https://doi.org/10.1093/bioinformatics/btl461
49. Hashimoto K, Rogozin IB, Panchenko AR. Oncogenic potential is related to activating effect of cancer single and double somatic mutations in receptor tyrosine kinases. Hum Mutat 2012;33:1566–75. https://doi.org/10.1002/humu.22145.
50. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 2020;21:6. https://doi.org/10.1186/s12864-019-6413-7
51. Pucci F, Bernaerts KV, Kwasigroch JM, Rooman M. Quantification of biases in predictions of protein stability changes upon mutations. Bioinformatics 2018;34:3659–65. https://doi.org/10.1093/bioinformatics/bty348.
52. Fang J. A critical review of five machine learning-based algorithms for predicting protein stability changes upon mutation. Brief Bioinform 2019. https://doi.org/10.1093/bib/bbz071.
53. Andreoletti G, Pal LR, Moult J, Brenner SE. Reports from the fifth edition ofCAGI: the critical assessment of genome interpretation. Hum Mutat 2019;40 (9):1197–201. https://doi.org/10.1002/humu.v40.910.1002/humu.23876
54. Petrosino M, Pasquo A, Novak L, Toto A, Gianni S, et al. Characterization of human frataxin missense variants in cancer tissues. Hum Mutat 2019;40:1400–13. https://doi.org/10.1002/humu.23789.
55. Savojardo C, Petrosino M, Babbi G, Bovo S, Corbi-Verge C, et al. Evaluating the predictions of the protein stability change upon single amino acid substitutions for the FXN CAGI5 challenge. Hum Mutat 2019;40:1392–9. https://doi.org/10.1002/humu.23843
56. Baldi P, Brunak S, Chauvin Y, Andersen CA, Nielsen H. Assessing the accuracy of prediction algorithms for classification: an overview. Bioinformatics 2000;16:412–24. https://doi.org/10.1093/bioinformatics/16.5.412.
57. Pauling L, Itano HA. Sickle cell anemia a molecular disease. Science 1949;110:543–8. https://doi.org/10.1126/science.110.2865.543.
58. Perutz MF, Mitchison JM. State of haemoglobin in sickle-cell anaemia. Nature 1950;166:677–9. https://doi.org/10.1038/166677a0.
59. Hunt JA, Ingram VM. Allelomorphism and the chemical differences of the human haemoglobins A, S and C. Nature 1958;181:1062–3. https://doi.org/10.1038/1811062a0.
60. Stenson PD, Mort M, Ball EV, Shaw K, Phillips A, Cooper DN. The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. Hum Genet 2014;133:1–9. https://doi.org/10.1007/s00439-013-1358-4.
61. Amberger JS, Bocchini CA, Schiettecatte F, Scott AF, Hamosh A. OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. Nucleic Acids Res 2015;43:D789–98. https://doi.org/10.1093/nar/gku1205.
62. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res 2019;47(D1):D941–7. https://doi.org/10.1093/nar/gky1015
63. Casadio R, Vassura M, Tiwari S, Fariselli P, Luigi MP. Correlating disease-related mutations to their effect on protein stability: a large-scale analysis of the human proteome. Hum Mutat 2011;32:1161–70. https://doi.org/10.1002/humu.21555.
64. Malhotra S, Alsulami AF, Heiyun Y, Ochoa BM, Jubb H, et al. Understanding the impacts of missense mutations on structures and functions of human cancer-related genes: a preliminary computational analysis of the COSMIC Cancer Gene Census. PLoS ONE 2019;14:. https://doi.org/10.1371/journal.pone.0219935e0219935 .
65. Singh V, Pacitto A, Donini S, Ferraris DM, Boros S, et al. Synthesis and Structure-Activity relationship of 1-(5-isoquinolinesulfonyl)piperazine analogues as inhibitors of Mycobacterium tuberculosis IMPDH. Eur J Med Chem 2019;174:309–29. https://doi.org/10.1016/j.ejmech.2019.04.027.
66. Munir A, Kumar N, Ramalingam SB, Tamilzhalagan S, Shanmugam SK, et al. Identification and characterization of genetic determinants of isoniazid and rifampicin resistance in mycobacterium tuberculosis in Southern India. Sci Rep 2019;9:10283. https://doi.org/10.1038/s41598-019-46756-x.
67. Maryam A, Vedithi SC, Khalid RR, Alsulami AF, Torres PHM, et al. The molecular organization of human cGMP specific phosphodiesterase 6 (PDE6): structural implications of somatic mutations in cancer and retinitis pigmentosa. ComputStructBiotechnol J 2019;17:378–89. https://doi.org/10.1016/j.csbj.2019.03.004
68. Glusman G, Rose PW, Prlic´ A, Dougherty J, Duarte JM, et al. Mapping genetic variations to three-dimensional protein structures to enhance variant interpretation: a proposed framework. Genome Med

2017;9:113. https://doi.org/10.1186/s13073-017-0509-y.
69. Chiang CH, Grauffel C, Wu LS, Kuo PH, Doudeva LG, et al. Structural analysis of disease-related TDP-43 D169G mutation: linking enhanced stability and caspase cleavage efficiency to protein accumulation. Sci Rep 2016;6:21581. https://doi.org/10.1038/srep21581
70. Shan Y, Eastwood MP, Zhang X, Kim ET, Arkhipov A, et al. Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. Cell2012;149:860–70. https://doi.org/10.1016/j.cell.2012.02.063.
71. Nussinov R, Jang H, Tsai CJ, Cheng F. Review: precision medicine and driver mutations: Computational methods, functional assays and conformational principles for interpreting cancer drivers. PLoSComputBiol 2019;15:. https://doi.org/10.1371/journal.pcbi.1006658e1006658
72. Jørgensen KW, Rasmussen M, Buus S, Nielsen M. NetMHCstab – predicting stability of peptide-MHC-I complexes; impacts for cytotoxic T lymphocyte epitope discovery. Immunology 2014;141:18–26. https://doi.org/10.1111/imm.12160.
73. Rasmussen M, Fenoy E, Harndahl M, Kristensen AB, Nielsen IK, et al. Pan- specific prediction of peptide-MHC class I complex stability, a correlate of T cell immunogenicity. J Immunol 1950;2016(197):1517–24. https://doi.org/10.4049/jimmunol.1600582.
74. Scheiblhofer S, Laimer J, Machado Y, Weiss R, Thalhamer J. Influence of protein fold stability on immunogenicity and its implications for vaccine design. Expert Rev Vaccines 2017;16:479–89. https://doi.org/10.1080/14760584.2017.1306441.
75. Folkman L, Stantic B, Sattar A, Zhou Y. EASE-MM: sequence-based prediction of mutation-induced stability changes with feature-based multiple models. J MolBiol 2016;428:1394–405. https://doi.org/10.1016/j.jmb.2016.01.012.
76. Cheng J, Randall A, Baldi P. Prediction of protein stability changes for single- site mutations using support vector machines. Proteins 2006;62:1125–32. https://doi.org/10.1002/prot.20810
77. Wainreb G, Wolf L, Ashkenazy H, Dehouck Y, Ben-Tal N. Protein stability: a single recorded mutation aids in predicting the effects of other mutations in the same amino acid site. Bioinformatics 2011;27:3286–92. https://doi.org/10.1093/bioinformatics/btr576
78. TizianaSanaviaa, Giovanni Biroloa, LudovicaMontanuccib, Paola Turinac, EmidioCapriottic, PieroFariselliа. Limitations and challenges in protein stability prediction upon genome variations: towards future applications in precision medicine. Computation and structural bio technology, 2020, 18: 1968-1979.
79. Giollo M, Martin AJM, Walsh I, Ferrari C, Tosatto SCE. NeEMO: a method using residue interaction networks to improve prediction of protein stability upon mutation. BMC Genomics 2014;15(Suppl 4):S7. https://doi.org/10.1186/1471- 2164-15-S4-S7
80. Pires DEV, Ascher DB, Blundell TL. mCSM: predicting the effects of mutations in proteins using graph-based signatures. Bioinformatics 2014;30:335–42. https://doi.org/10.1093/bioinformatics/btt691.