SCIENTIFIC REPORTS

OPEN

Prediction and interpretation of deleterious coding variants in terms of protein structural stability

Received: 10 July 2017
Accepted: 20 February 2018
Published online: 14 March 2018

François Ancien[1,2], Fabrizio Pucci[1,2], Maxime Godfroid[1,3] & Marianne Rooman[1,2]

The classification of human genetic variants into deleterious and neutral is a challenging issue, whose complexity is rooted in the large variety of biophysical mechanisms that can be responsible for disease conditions. For non-synonymous mutations in structured proteins, one of these is the protein stability change, which can lead to loss of protein structure or function. We developed a stability-driven knowledge-based classifier that uses protein structure, artificial neural networks and solvent accessibility-dependent combinations of statistical potentials to predict whether destabilizing or stabilizing mutations are disease-causing. Our predictor yields a balanced accuracy of 71% in cross validation. As expected, it has a very high positive predictive value of 89%: it predicts with high accuracy the subset of mutations that are deleterious because of stability issues, but is by construction unable of classifying variants that are deleterious for other reasons. Its combination with an evolutionary-based predictor increases the balanced accuracy up to 75%, and allowed predicting more than 1/4 of the variants with 95% positive predictive value. Our method, called SNPMuSiC, can be used with both experimental and modeled structures and compares favorably with other prediction tools on several independent test sets. It constitutes a step towards interpreting variant effects at the molecular scale. SNPMuSiC is freely available at https://soft.dezyme.com/.

Despite the large amounts of genomic data collected in the last decade and the multiple efforts to elucidate their links with phenotypic traits, an accurate and interpretative classification of the effects of genetic variants on various disorders remains a difficult goal to achieve. The high complexity of the problem is mainly due to the heterogeneity of the molecular mechanisms underlying the diseases, most of which have still to be deeper understood to be useful at the clinical level. For example, a change in a single DNA base pair that occurs in a coding region can be synonymous, *i.e.* change the codon but not the amino acid, and perturb the gene expression level or the efficiency of the translational mechanism. Non-synonymous mutations that change both the codon and the associated amino acid can modify the biophysical properties of the encoded protein, such as its thermodynamic stability or solubility, or its functional properties by affecting the active site or the binding affinity for ligands and proteins partners. Moreover, the impact of a given variant crucially depends on its context within the genome: its deleteriousness can depend on the presence of other variants in the same or in other genes. Its effect also varies according to the cellular context. Indeed, some variants occur in proteins that play an essential role for the cell or the organism, which cannot be performed by any other protein. In such case, even a slightly destabilizing variant can be strongly deleterious. In contrast, some proteins perform functions that can very well be performed by other proteins, so that inactivating mutations do not cause diseases.

Among all the possible types of variants, the non-synonymous single nucleotide variants in coding genes (SNV) play an important role since they constitute more than half of the mutations known to be associated in human inherited disorders[1]. They are directly related to a wide range of pathological conditions among which Parkinson's and Alzheimer's diseases, and are involved in complex diseases such as cancers, in which the accumulation of different types of genetic variants determine the tumor initiation and progression[2].

[1]Department of BioModeling, BioInformatics & BioProcesses, Université Libre de Bruxelles (ULB), CP 165/61, Roosevelt Avenue 50, 1050, Brussels, Belgium. [2]Interuniversity Institute of Bioinformatics in Brussels, ULB, CP 263, Triumph Bld, 1050, Brussels, Belgium. [3]Present address: Institute of General Microbiology, Kiel University, Am Botanischen Garten 11, 24118, Kiel, Germany. François Ancien and Fabrizio Pucci contributed equally to this work. Correspondence and requests for materials should be addressed to F.A. (email: fancien@ulb.ac.be) or F.P. (email: fapucci@ulb.ac.be) or M.R. (email: mrooman@ulb.ac.be)

It is thus of primary interest, especially in the context of personalized medicine, to have computational methods to classify SNVs and identify those that are disease causing. Classification methods developed in the literature usually integrate different kinds of biological data in view of gathering the complexity of the phenomenon, and require only the amino acid sequence as input. Some of them, such as Provean[3,4], SIFT[5] or the Evolutionary Action method[6,7], use as sole ingredient the evolutionary conservation in homologous proteins at the mutated position and in the neighborhood along the sequence. Though this information is represented by a single score, it corresponds to a mixture of several molecular effects, among which protein stability, solubility, function, and interactions. Understanding why a residue is conserved is impossible on the basis of the evolutionary data only. Other methods add further layers of information on top of this data, by considering for example structural alterations in the biophysical characteristics of proteins, which are usually predicted from the sequence and sometimes derived from the structure when available; these methods include Mutation Tasser[8], Mutation Assessor[9], CADD[10], Polyphen-2[11], and VIPUR[12]. The addition of contextual information about the protein-protein interaction networks, the protein essentiality index, and the pathway in which the protein is involved appear to significantly improve the performances, as shown in DEOGEN[13,14] and SuSPect[15]. These various computational tools use machine learning techniques to predict from the considered features the effect of missense mutations.

Unfortunately, the accuracy of the prediction methods remain limited, with often a quite high false-positive rate with the detrimental consequence that many of the predicted deleterious variants observed in clinical whole-exome sequencing turn out to be neutral[16]. Frequently moreover, the computational methods only focus on reaching the highest variant classification accuracy on a given dataset rather than predicting and understanding the modifications that occur at the molecular scale and are responsible for the altered phenotype; yet this information is a prerequisite for the rational design of drugs or treatments. As a matter of fact, although protein structure can help getting insight into in the molecular impact of mutations, it is rarely (fully) exploited by variant classification methods, except in a few cases[12,17–19].

The present analysis aims at deepening the understanding of the relation between protein structural stability and variant deleteriousness. To evaluate the stability of a given protein structure and its change upon single-site mutations, we used knowledge-based statistical mean-force potentials derived from a dataset of three-dimensional (3D) protein structures. Combinations of these potentials, performed with the help of different artificial and probabilistic neural network architectures that include the solvent accessibility of the mutated residue to modulate the importance of the energetic terms, were used to classify SNVs into neutral and deleterious. Note that the primary goal of this analysis is not to reach the highest prediction accuracy but rather to get insight into the functional and stability characteristics of protein variants and their relation with the phenotypic traits. Our objective also involves predicting with high accuracy the subset of mutations that are deleterious due to stability problems.

Methods

Non-synonymous SNV datasets for training and testing. The training dataset was built by combining the annotated, non-synonymous SNV data from three different databases: DbSNP[20], SwissVar[21] and HumSaVar[22]. In a first stage, we combined all SNVs while avoiding repeats. Note that variants occurring in more than one database with different neutral/deleterious annotation were discarded. The SNVs were characterized in these databases by the protein's UniProt code[22] and the variant's residue number, without any reference to a possible 3D structure. We had thus, in a second stage, to identify the subset of variants introduced in proteins that have an experimental structure, using the SIFTS webserver[23]. We only considered the subset of mutations introduced in:

- globular proteins, or cytoplasmic or extracellular domains of membrane proteins;
- proteins with an X-ray structure available in the Protein DataBank[24] (PDB) of resolution of 2.5 Å at most.

Some of the PDB protein structures have not exactly the same sequence as the corresponding entry in the UniProt database, but contain one or several mutations. We only retained the subset of SNVs of residues that are far from all these additional mutations by an inter-Cα distance of 10 Å at least, to avoid direct interactions between them. The final dataset is referred to as **S** and is reported in Table S1 of Supplementary Material. It contains 5,302 variants inserted in 1,016 different proteins, 1,301 of which are annotated as polymorphic and 4,001 as deleterious.

To test our predictors and compare their performances with commonly used classification tools, we applied them to four test sets used by VIPUR[12]: S_H (variants in Human proteins), S_{NH} (in Non-Human proteins), S_{CV} (from the ClinVar dataset[25]) and S_{SSC} (from the Simon Simplex Collection[26]). They are described in Supplementary Material (Tables S2–S5).

Statistical Potentials and other structural features. We used the statistical potential formalism to evaluate the stability of a protein and its change caused by non-synonymous SNVs, and to understand the link with their polymorphic or pathogenic effect. These potentials are knowledge-driven mean force potentials that are extracted from a dataset of well-resolved X-ray protein structures[27–31]. The energetic contribution ΔW associated to the sequence-structure association (s,c), where s and c are sequence and structure elements respectively, is obtained from the inverse Boltzmann law as:

$$\Delta W(s, c) = -k_B T \ \mathrm{Log}\left(\frac{P(s, c)}{P(c)P(s)}\right), \tag{1}$$

where k_B is the Boltzmann constant, T the absolute temperature, and $P(c,s)$, $P(c)$ and $P(s)$ are the probabilities of observing (c,s), (c) or (s) elements. These probabilities are approximated in terms of the number of occurrences of these elements in a reference dataset of protein structures. The sequence elements s are amino acid types and the structural motifs c can be interresidue distances, torsion angle domains or solvent accessibilities. Higher order potentials, in which different combinations of sequence and structure elements are considered, have also been utilized in this investigation. For example, a 3-term potential describing the association between one structure and two sequence elements is defined as:

$$\Delta W(s_1, s_2, c) = -k_B T \ \mathrm{Log}\left(\frac{P(s_1, s_2, c)P(s_1)P(s_2)P(c)}{P(s_1, c)P(s_2, c)P(s_1, s_2)}\right). \tag{2}$$

For sake of simplicity, such potentials are referred to as ΔW_{ssc} in what follows. Further generalizations and technical details can be found in our earlier papers[31,32].

We used here 13 statistical potentials, listed in Table S6, which can be grouped in three classes according to the type structural descriptor c: (1) distance potentials describing tertiary interactions, in which c is the distance d between the average side chain geometric centers of two amino acids; (2) solvent accessibility potentials, in which c is the solvent accessibility a of a given residue; (3) torsion potentials describing local interactions along the sequence, in which c is the main chain torsion angle domain t of a residue. Several combinations of these basic descriptors were also used. Besides these potentials, we also considered two biophysical characteristics: the solvent accessibility A of the mutated residues and the difference in volume ΔV between the wild type and the mutated residues[32].

These potentials and structural features were used to estimate, for each mutation of our SNV dataset **S**, the corresponding change in volume ΔV and in folding free energy $\Delta\Delta W_i$, with $i = 1, \ldots 13$. In a first step, the distributions of $\Delta\Delta W_i$, A and ΔV were compared between deleterious and neutral variants from **S**, in view of getting insight into the relation between the structure and stability characteristics and the pathogenicity of the variants. In a second stage, all the potentials and features were combined to set up a predictor of variant deleteriousness.

Utilizing neural network architectures for variant classification. *Pathogenicity prediction from thermodynamic and thermal stability changes.* The first approach for predicting on a structural basis which variants are disease-causing and which are not consists in using our PoPMuSiC[32] and HoTMuSiC[33] algorithms, which predict changes in thermodynamic and thermal stabilities upon single-site mutations from protein 3D structures. More precisely, PoPMuSiC uses a combination of the 13 statistical potentials and the two structural features listed in Table S6 to estimate the values of the folding free energy changes upon mutation:

$$\Delta\Delta G = \sum_{i=1}^{13}\alpha_i(A)\Delta\Delta W_i + \alpha_{14}(A)\Delta V_+ + \alpha_{15}(A)\Delta V_- + \alpha_{16}(A), \tag{3}$$

where ΔV_+ and ΔV_- are two volume terms defined as: $\Delta V_{\pm} = \theta(\pm\Delta V)||\Delta V||$, with θ being the Heaviside function, which take into account the potentially destabilizing effect due to the creation of holes or stress in the protein interior. The $\alpha_i(A)$ functions are sigmoids that depend on the solvent accessibility A of the mutated residues:

$$\alpha_i(A) = \frac{\omega_i}{1 + exp^{-\nu_i(A-\xi_i)}} + \varphi_i, \tag{4}$$

with ω_i, ν_i, ξ_i and ϕ_i parameters that were identified so as to minimize the mean square deviation between predicted and observed stability changes in a dataset of mutations with experimentally characterized $\Delta\Delta G$s; note that this dataset has a negligible overlap with **S** (one mutation). For further technical details about the choice of parameters or the construction of the model, see the original PoPMuSiC papers[32,34].

The classification of the variants of the **S** dataset into deleterious and neutral was performed on the basis of the computed $\Delta\Delta G$ values: SNVs with a predicted $\Delta\Delta G$ higher than a threshold value, which correspond to the most destabilizing mutations, were predicted as deleterious, whereas those with lower $\Delta\Delta G$ were predicted as neutral. The value of the threshold was identified so as to optimize the values of the balanced accuracy (BACC):

$$BACC = \frac{TP}{2(TP + FN)} + \frac{TN}{2(TN + FP)}, \tag{5}$$

where TP, TN, FP, FN are the true positive, true negative, false positive and false negative predictions, respectively. Positive is chosen to correspond to deleterious variants and negative to neutral ones.

Similarly, we also used the thermal stability predictor HoTMuSiC[33] for deleterious/neutral classification. This algorithm predicts changes in melting temperature ΔT_m upon mutations rather than changes in folding free energy $\Delta\Delta G$. These two quantities are anti-correlated only in a first approximation, and yield two different informations about protein stability[35]. HoTMuSiC uses a distinct combination of the statistical potentials:

$$\Delta T_m = \frac{1}{aN_r + b}\sum_{i=1}^{9}\beta_i(A)\Delta\Delta W_i + \beta_{10}(A)\Delta V_+ + \beta_{11}(A)\Delta V_- + \beta_{12}(A), \tag{6}$$

where the $\beta_i(A)$ functions are sigmoids of the form specified in Eq. (4), and N_r is the number of residues in the protein. The parameters in the $\beta_i(A)$ functions, as well as a and b, where identified using as cost function the mean

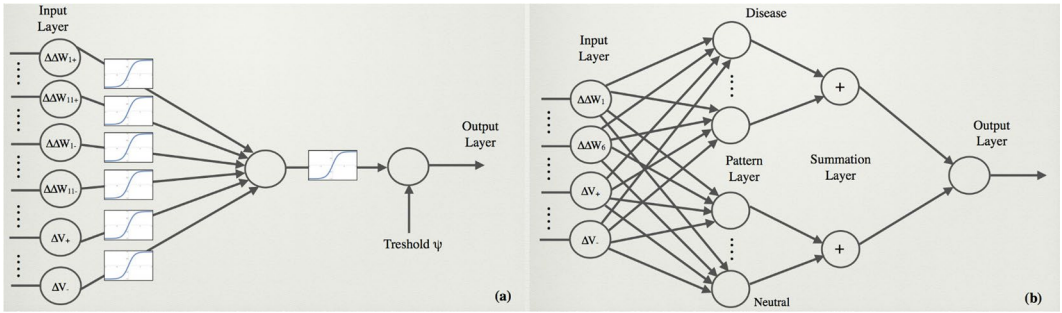

Figure 1. Schematic representation of: (**a**) the artificial neural network and (**b**) the probabilistic neural network used in the classification of the variants.

square deviation between experimental and predicted ΔT_m s on a dataset of characterized mutants, that displays only a very small intersection with **S** (3 mutations only). Note that the number of energy terms has been restricted to nine because of the smaller size of the ΔT_m learning dataset, to avoid overfitting.

Using a similar procedure as for the $\Delta\Delta G$ predictions, a given variant of the **S** dataset is considered as deleterious if its ΔT_m value is smaller than a given threshold, which means that the variant provokes a strong thermal destabilization; otherwise it is predicted as neutral. Again, the threshold value was identified using the BACC score as cost function.

Stability-based pathogenicity index using an artificial neural network. Using a different combination of the same potentials and two structural features listed in Table S6, we set up a specific predictor of the deleteriousness of a mutation based on stability criteria. It has the peculiarity to relax the hypothesis assumed in the previous section that only destabilizing mutations are likely to be deleterious. Indeed, although most deleterious mutations are destabilizing, some stabilizing mutations can also cause diseases, usually because they affect the protein function, as observed earlier[36,37]. To take this into account, we considered the stabilizing and destabilizing statistical potential contributions separately, and separated the potential terms into their positive and negative parts:

$$\overline{\Delta\Delta W}_{i_\pm} = \theta(\pm\Delta\Delta W_i)||\Delta\Delta W_i||, \tag{7}$$

and used them to define the stability-based pathogenicity index *I*:

$$I = \sum_{i=1}^{11}\phi_i^+(A)\overline{\Delta\Delta W}_{i_+} + \sum_{i=1}^{11}\phi_i^-(A)\overline{\Delta\Delta W}_{i_-} + \phi_{12}(A)\Delta V_+ + \phi_{13}(A)\Delta V_-. \tag{8}$$

The functions $\phi_i^{\pm}(A)$ are solvent accessible-dependent sigmoids of the form of Eq. (4). The number of parameters related to the potentials are here almost doubled compared to the PoPMuSiC model structure of Eq. (3).

The model's parameters were optimized using the double layer artificial neural network (ANN) schematically depicted in Fig. 1a, using as learning set the annotated SNVs from the **S** dataset and as cost function the mean square deviation between *I* and the annotated variant effect, with deleterious defined as 1 and neutral as 0. The mutations were then considered as deleterious when the pathogenicity index is larger than a given threshold value ψ and neutral otherwise. The threshold value was optimized so as to maximize the BACC score (Eq. (5)) of the classification. To evaluate the performance of the method, we used a 5-fold cross-validation procedure. We also tested other validation strategies, among which 10-fold cross validation, with comparable results.

By construction, this new model allows predicting as deleterious both stabilizing and destabilizing mutations. It is a striking illustration that elucidating the biophysical mechanisms underlying the variant effects can be used to guide the model architecture and contribute to improve the classification performances and the understanding of the data.

Stability-based pathogenicity index using a probabilistic neural network. Probabilistic neural networks (PNN)[38] have architectures that are very different from ANNs, and are utilized for classification purposes. The idea behind their construction consists in a Bayes classification strategy through the estimation of the probability density function (PDF) for each considered category, based on a number of characteristics of the elements to be classified. PNNs are composed of four layers, as shown in Fig. 1b. The input layer contains the features characterizing the considered variant, chosen here to be: $\vec{X} = (\Delta\Delta W_i, \Delta V_+, \Delta V\_, A)$, with $i = 1, \ldots, 6$ (see Supplementary Material for details). The second layer, called pattern layer, contains as many nodes as there are samples in the training set. Each node, labeled as *j*, contains a probability value computed from the comparison between the feature vectors of the input variant ($\vec{X}$) and of the j^{th} sample ($\vec{X}_j$), and estimated via a multivariate Gaussian distribution centered on $\vec{X}_j$:

$$P_j(\vec{X}) = \frac{1}{(2\pi)^{p/2}(\det\Sigma)^{1/2}} e^{-\frac{1}{2}(\vec{X}-\vec{X}_j)^T\Sigma^{-1}(\vec{X}-\vec{X}_j)}, \tag{9}$$

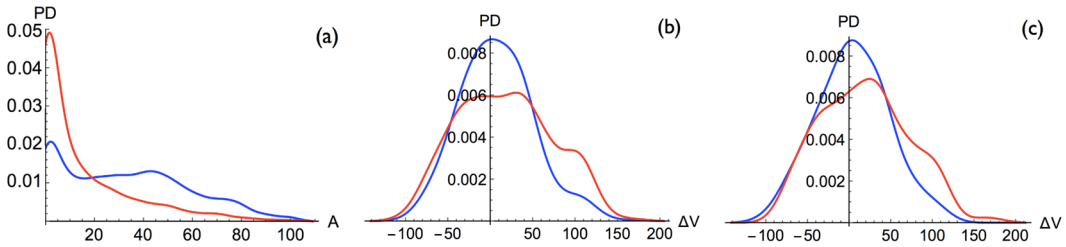

Figure 2. Probability density distributions of deleterious mutations (red curve) and neutral mutations (blue curve) for: (**a**) the solvent accessibility A (0–100%) of the mutated residues; (**b**) the change in volume ΔV (in Å^3) for residues with $A \leq 60\%$; (**c**) the change in volume ΔV for residues with $A > 60\%$; in our conventions, mutations of smaller into larger residues have a positive ΔV value.

where p is the number of features in the input layer and Σ the $p \times p$-dimensional covariance matrix of the PDF, considered here to be diagonal. Usually, PNNs use equal covariances for all features, which we do not assume here, hence allowing every feature to contribute with a different weight to the pathogenicity of a mutation. The elements of this Σ-matrix are parameters to be optimized. In the summation layer, the PDFs of the samples that belong to the same category - here neutral or deleterious - are summed, thus defining the category PDFs P_D and P_N:

$$P_{\mathrm{D}}(\overrightarrow{X}) = \frac{\eta_1}{n_D} \sum_{j\in\{deleterious\}} P_{\mathrm{j}}(\overrightarrow{X}) \; ; \qquad P_{\mathrm{N}}(\overrightarrow{X}) = \frac{\eta_2}{n_N} \sum_{j\in\{neutral\}} P_{\mathrm{j}}(\overrightarrow{X}), \tag{10}$$

where n_D and n_N are the number of disease-causing and neutral variants in the training dataset, respectively, and η_1 and η_2 two parameters to be optimized. The output layer contains the result of the classification according to the rule that a variant with feature vector $\overrightarrow{X}$ is deleterious if $P_{\mathrm{D}}(\overrightarrow{X}) > P_{\mathrm{N}}(\overrightarrow{X})$ and neutral otherwise. The parameters of the model, *i.e.* η_1, η_2 and the p diagonal elements of the covariance matrix Σ, are identified so as to minimize the BACC score on the variants of the **S** training set. Again, all performances are evaluated in 5-fold cross validation.

Combining stability-based pathogenicity index with evolutionary information. Structural stability and evolutionary information are entangled but distinct notions. On the one hand, biophysical constraints, including stability but also solubility, aggregation propensity, interactions and function, act on natural selection, thus affecting the evolution of proteins, while on the other hand, natural selection guides mutations, which in turn have an impact on the biophysical properties of proteins.

In our last predictor, both types of information are joined to obtain a more complete picture of the variant effects. We would like to stress that, in contrast to the black-box machine learning approach usually employed in variant classification, we tested concretely whether the stability change due to a residue substitution impacts directly on the variant's deleteriousness, taking also into account the role of the evolutionary pressure.

The model structure employed is a simple combination of the Provean[3,4] score, noted PRO, with the pathogenicity index I defined in Eq. (8). The new index, called $\mathcal{J}$, is defined as:

$$\mathcal{J} = \gamma_i + \gamma_2 I + \gamma_3 \mathrm{PRO}. \tag{11}$$

The three coefficients γ_i as well as the classification threshold were identified by optimizing the BACC score on the **S** dataset. Again, all tests were performed in strict 5-fold cross-validation. We call this predictor SNPMuSiC.

Results and Discussion

Analysis of the statistical potentials and other structural features. Let us start analyzing the solvent accessibility A of the mutated residues, which is one of the structural attributes known to be related to variant deleteriousness[18]. Indeed, mutations in the core are usually more stabilizing or more destabilizing with respect to surface mutations, since buried residues play a special role both in the early folding stages and in the stability of the folded structure. As a consequence, core residue substitutions are more likely to cause structural rearrangements and/or the loss of protein function, and have thus a higher probability to be deleterious for the organism. This is what we observe in Fig. 2a: variants have a higher probability to be disease-causing if the mutation is in a buried region, defined as $A < 20\%$, while the neutral variant distribution shows only a mild dependence on A with a weak decrease of the probability for large A-values.

The second structural feature we considered is the change in volume ΔV upon residue substitution. We found that in the totally and partially buried regions, up to a solvent accessibility $A \leq 60\%$, the substitutions from smaller into larger amino acids (positive ΔV s) have a higher probability to be deleterious, and so are, albeit to a much smaller extent, the substitutions from larger into smaller amino acids (negative ΔV s), as shown in Fig. 2b. In other words, substitutions that create stress or holes in the protein interior are more likely to be deleterious; holes are, however, easier to manage through limited structural rearrangements than stress. On the protein surface, defined as $A > 60\%$, mutations from smaller to larger residues still have a higher probability to be disease-causing, but less than in the core (Fig. 2c), whereas no difference is observed for substitutions from larger to smaller residues. Note that the differences between deleterious and neutral variant distributions, both for the solvent

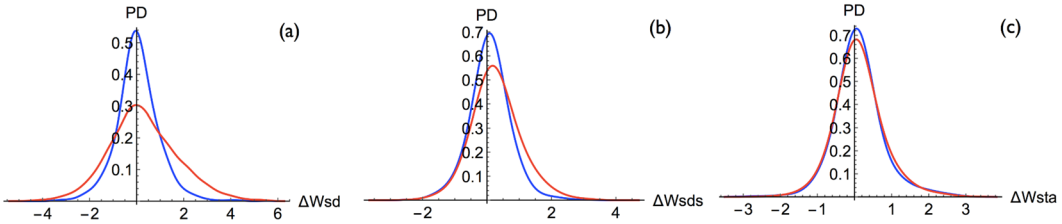

Figure 3. Probability density distributions of deleterious mutations (red curve) and neutral mutations (blue curve) for the changes in folding free energy $\Delta\Delta W$ (in kcal/mol) computed with the following statistical potentials: (**a**) the distance potential ΔW_{sd}; (**b**) the distance potential ΔW_{sds}; (**c**) the torsion angle and solvent accessibility potential ΔW_{sta}; in our conventions, positive $\Delta\Delta W$ values correspond to destabilizing mutations.

Method	Sensitivity	Specificity	PPV	NPV	BACC	Threshold	AUROC
PoPMuSiC	0.65	0.62	0.84	0.37	0.63	0.75 kcal/mol	0.68
HoTMuSiC	0.59	0.65	0.84	0.34	0.62	−1.8 °C	0.66
Solvent Accessibility	0.71	0.66	0.86	0.42	0.68	18.0%	0.72
PNN	0.69	**0.72**	0.88	0.43	0.71	—	0.76
ANN	0.70	**0.72**	**0.89**	0.44	0.71	0.74	0.77
Provean	**0.85**	0.58	0.86	**0.56**	0.72	−2.5	0.80
SNPMuSiC	0.79	0.71	**0.89**	0.52	**0.75**	0.66	**0.83**

Table 1. Performance of the different prediction methods in mutation-based 5-fold cross validation on the learning set. Sensitivity is defined as $TP/(TP + FN)$, specificity as $TN/(TN + FP)$, positive predictive value (PPV) as $TP/(TP + FP)$, and negative predictive value (NPV) as $TN/(TN + FN)$. The scores and threshold values correspond to averages on the 5-fold cross-validation experiments. The values in bold indicate the highest scores in each category; the AUROC score in bold is statistically different from the other AUROC scores, as estimated by DeLong's test.

accessibility and the volume change, are statistically significant, as measured by a very low Kolmogorov-Smirnov (KS) test P-value shown in Table S6.

Next we compared the probability density functions of deleterious and neutral mutations for the changes in folding free energy $\Delta\Delta W$ computed by each of the 13 statistical potentials listed in Table S6. We can classify the results into three classes, whose typical behavior is depicted in Fig. 3a–c. In the first class, disease-causing variants are associated with both highly stabilizing and highly destabilizing values. In class 2, the deleterious variants are more frequently associated with destabilizing mutations but not with stabilizing mutations. In the last class, the difference between deleterious and neutral distributions is not statistically significant, as shown by a KS-test P-value larger than 0.05. Note that the biophysical interpretation of the neutral and deleterious variant distributions is here less obvious than for solvent accessibility and volume changes, since statistical potentials describe complex interactions between sequence and structure factors.

Classification Performances in Cross Validation. The structural and stability features analyzed in the previous section were used, alone or in combination, to classify single-site mutations into deleterious or neutral, as described in the Methods section. The performance of this classification was evaluated in 5-fold cross validation at the mutation level by training the model on 4/5 randomly chosen mutations that belong to the **S** dataset and applying it on the remaining entries. This procedure was repeated five times, considering each of the five subsets in turn as test set. The average scores are reported in Table 1. Another 5-fold cross-validation was performed at the protein level, by dividing the proteins rather than the mutations into five subsets; the results are shown in Table S7. No statistically significant differences are observed between the two types of cross-validation scores, as estimated by the DeLong test[39].

Strikingly, the variant classification based solely on solvent accessibility gives quite good results, with a BACC score of 0.68, as expected by the large dissimilarity between deleterious and neutral variant distributions shown in Fig. 2a. This result again demonstrates the important role of core residues in folding, structure and stability, and their sensitivity to mutations.

The classifications based on the prediction of thermodynamic and thermal stability changes upon mutations computed by PoPMuSiC (Eq. (3)) and HoTMuSiC (Eq. 6)), respectively, yield BACC scores of 0.62 and 0.63; the neutral and deleterious variant distributions for PoPMuSiC are shown in Fig. 4a. There is thus a significant, but limited, correlation between destabilization and deleteriousness. Two reasons can be invoked to explain why this correlation is not higher. The first is that deleteriousness can be caused by destabilization but also by other factors. The second explanation is that not only destabilizing but also stabilizing mutations can be deleterious, as seen in Fig. 3a. To illustrate this important point, we describe some disease-causing variants associated with a gain in structural stability in the next subsection.

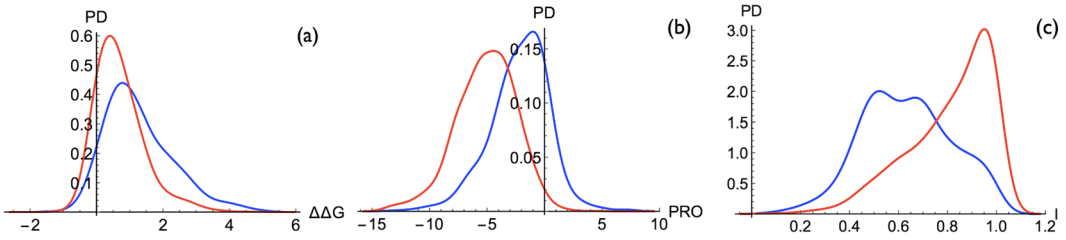

Figure 4. Probability density distributions of deleterious mutations (red curve) and neutral mutations (blue curve) for (**a**) the change in folding free energy $\Delta\Delta G$ computed by PoPMuSiC (Eq. (3)) (in kcal/mol), (**b**) the Provean score[3,4], and (**c**) the pathogenicity index I computed by the ANN model (Eq. (8)).

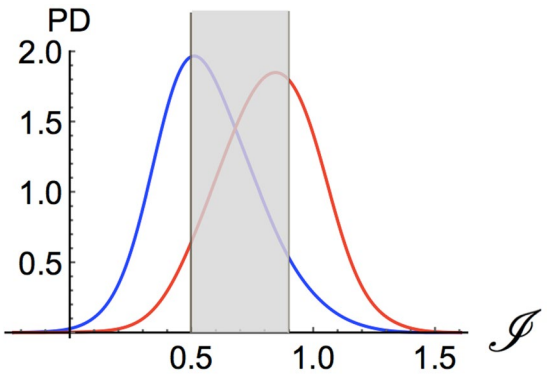

Figure 5. Probability density distribution of deleterious mutations (red curve) and neutral mutations (blue curve) for the pathological index $\mathcal{J}$ (Eq. (11)) computed by SNPMuSiC. The distribution curves in the high confidence intervals, which lie from either side of the two vertical lines, are depicted on a white background.

To take into account that both stabilizing and destabilizing mutations can cause diseases, we set up two new classifiers that use the same structural features and statistical potentials as PoPMuSiC and HoTMuSiC but are based on two different neural network architectures, an ANN (Eq. (8)) and a PNN (Eq. (10)). These two classifiers both reach a BACC score of 0.71; the neutral and deleterious variant distributions of the ANN model are shown in Fig. 4c. As expected, these two models take more properly into account the weight of the different energy contributions in the determination of the variant deleteriousness. The fact that the ANN and PNN model structures yield the same score suggests that it could correspond to the maximum performance that can be obtained from the features considered.

Quite interestingly, the Provean classifier[3,4] reaches the same score as our ANN and PNN models (0.07 higher), though the former is based on evolutionary information that mixes stability with other biophysical characteristics such as solubility and activity while the latter purely exploits stability information. The variant distributions obtained with this predictor are depicted in Fig. 4b. A closer analysis shows, however, some important differences. Our ANN and PNN prediction methods have a sensitivity of about 0.70 that is much lower than the Provean value of 0.85, and a specificity of 0.72 that is much higher than the Provean value of 0.58 (Table 1). The high specificity and low sensitivity of our ANN and PNN classifiers come from their number of false positives being reduced by nearly 50% compared to Provean, as well as a larger false negative rate. This result can be explained as follows: variants predicted as strongly stabilizing or destabilizing are usually well classified as disease-causing as they are very likely to have an impact on the phenotype; this accounts for the low false positive rate. Conversely, our ANN and PNN predictors are unable, by construction, to predict the deleterious effects due to biophysical characteristics other than stability; this rationalizes the high number of false negatives.

Our last and best predictor, that we call SNPMuSIC, combines the deleteriousness index I of the ANN model (Eq. (8)) with the Provean score, thus defining the $\mathcal{J}$ index (Eq. (11)). This model reaches the highest BACC score, *i.e.* 0.75, and the highest Area Under the Receiver Operating Characteristic curve (AUROC score), *i.e.* 0.83. The associated distributions of neutral and deleterious variants are quite well separated as shown in Fig. 5. This predictor has BACC and AUROC scores that are more than 3% higher than both Provean and ANN, which is statistically significant (P-value $< 10^{-4}$ using a bootstrap test). It performs even better than more complete classification methods, based on a whole range of sequence and structural features, such as Polyphen-2[11]. Strikingly, SNPMuSiC increases the sensitivity of the ANN and PNN models up to 0.79, and almost maintains their high specificity value (0.71) (Table 1).

Classification Performances on the Test Sets. To analyze in more detail the performances of our predictors, we applied them to four different test sets $\mathbf{S_H}$, $\mathbf{S_{NH}}$, $\mathbf{S_{CV}}$ and $\mathbf{S_{SSC}}$ taken from[12] and specified in Tables S2–S5. In contrast to our learning set, these sets do not contain only well resolved X-ray structures but also structures

Method	Sensitivity	Specificity	PPV	NPV	BACC	AUROC
Human Variants						
VIPUR	0.87	0.55	**0.83**	0.62	**0.71**	**0.77**
Polyphen-2	**0.95**	0.33	0.78	0.71	0.64	0.70
Provean	0.94	0.39	0.80	**0.73**	0.67	0.72
SNPMuSiC	0.76	**0.60**	**0.83**	0.49	0.68	**0.74**
Non-Human Variants						
Polyphen-2	**0.96**	0.30	0.77	**0.78**	0.63	**0.73**
Provean	0.94	0.34	0.77	0.70	0.64	0.70
SNPMuSiC	0.78	**0.54**	**0.80**	0.51	**0.66**	0.70

Table 2. Comparison of the performances of the different predictors on the test sets $\mathbf{S_H}$ and $\mathbf{S_{NH}}$. These datasets have no overlap with the training sets of Polyphen-2 and SNPMuSiC. For VIPUR, the scores for $\mathbf{S_H}$ are those labelled as being cross validated in[12], while for the $\mathbf{S_{NH}}$ set, no cross validated scores are available. For the sequence-based method Provean, the dataset overlap has not been considered, although it plays a role in the identification of the threshold values. See Table 1 for further details. The values in bold indicate the highest scores in each category; the AUROC scores that are not significantly different from the highest score (as estimated by a DeLong test P-value ≥ 0.05) are also in bold.

obtained by comparative modeling. This test amounts thus to evaluating the performance of our predictors on low-resolution structures. They were compared with the commonly used deleteriousness predictors Polyphen-2[11], Provean[3,4], CADD[10], SIFT[5] and VIPUR[12], which use sequence and sometimes structural features but no contextual (or systems biology) information. The results of these predictors were taken from VIPUR's article[12].

The first two test sets correspond to human variants from HumVar[11] and to non-human variants, respectively. We excluded from the former set the variants present in the training sets of our method and of Polyphen-2, as well as the proteins whose 3D structure was modeled from templates that have less than 30% identity with respect to the target sequence to avoid incorrectly modeled structures. The results are shown in Table 2. SNPMuSiC has the highest specificity and PPV on both test sets, and the highest BACC on the non-human set S_{NH}. In contrast, Polyphen-2, Provean and VIPUR have better sensitivity and NPV.

We also evaluated the performances on the two other sets $\mathbf{S_{CV}}$ and $\mathbf{S_{SSC}}$, which contain neutral and disease-causing human variants from ClinVar[25] and *de novo* missense mutations of the Simon Simplex Collection (SSC)[26], respectively. The results are shown in Table S8. Here, we could not filter out incorrect 3D models, as we do not have access to the sequence identity between the target and template sequences. This could explain why sequence-based predictors sometimes perform better than structure-based ones on $\mathbf{S_{CV}}$. In spite of this, the specificity and PPV scores of SNPMuSiC are the highest on both sets.

This leads us to the conclusion that SNPMuSiC is also applicable to mutations in proteins with low-resolution (modeled) structures, which drastically increases its application possibilities. Its strength is rooted in its high PPV and specificity for all test sets analyzed, compared to other commonly used sequence- or structure-based classification tools. Hence, the deleterious mutations predicted by SNPMuSiC have a large chance of being so. This is in accordance with its construction: it only predicts mutations that are deleterious because of stability reasons, and misses some other deleterious mutations that are instead caused by other biophysical mechanisms.

Link between protein stabilization and disease. While destabilizing mutations can clearly be deleterious as they are likely to cause structural modification which can negatively impact on protein function, it is less obvious that variants with a stabilizing effect can also be disease causing. To clarify this point we analyzed the ten deleterious mutations of the variant dataset **S** that are predicted to be the most thermodynamically stabilizing by PoPMuSiC; they are reported in Table 3. All the mutated residues are located in the protein core, with an accessibility lower than 15%. Note that the imperfect correlation of thermodynamic and thermal stabilities[35] is reflected in the imperfect anticorrelation between PoPMuSiC's $\Delta\Delta G$ and HoTMuSiC's ΔT_m prediction values.

Obviously, the PoPMuSiC and HoTMuSiC-based classifiers wrongly predict these ten mutations as neutral, since only destabilizing mutations have a chance to be classified as deleterious by these two models. In contrast, both our ANN-based predictor and SNPMuSiC, which are designed to be able to predict stabilizing disease-causing mutations, correctly classify all ten mutations as deleterious.

It is instructive to briefly analyze the biophysical mechanisms that are involved in the pathogenic phenotype of these SNVs caused by protein stabilization. These are taken from the annotations of the variants reported in the Uniprot database[22]. As shown in Table 3, the protein stabilization leads to a decrease of the enzymatic activity in the majority of the cases, which can be explained by the mutation being close to the protein active site or inducing inhibiting allosteric effects. Other mutations cause the change in affinity for ligands or protein partners[40,41] or affect post-translational modifications sites[42].

In summary, this investigation shows that biophysical knowledge at the molecular level - in this case, the fact that both strongly stabilizing and destabilizing mutations are likely to cause diseases - can guide the design of the model structures, improve the classification performances and lead to a deeper understanding of the variant effects.

Protein	Chain	Mutation	Biophysical effect	$\Delta\Delta G$ PoPMuSiC (kcal/mol)	ΔT_m HoTMuSiC (°C)	I index ANN	$\mathcal{J}$ index SNPMuSiC
1m6i	A	E493V	Increase of NADH affinity	−2.24	1.7	1.00	4.8
2wzb	A	D163V	Decrease of enzymatic activity	−1.62	1.0	1.00	4.1
3hcn	A	T283I	Decrease of enzymatic activity	−1.46	1.1	0.92	1.9
2izz	A	G206W	Loss of function	−1.35	2.8	1.00	1.4
2nt0	A	D399Y	Decrease of enzymatic activity	−1.31	0.8	1.00	4.4
3f9m	A	G385V	Decrease of enzymatic activity	−1.23	−0.2	0.84	1.3
4do4	A	R329W	Decrease of enzymatic activity	−0.99	0.3	0.99	2.3
1aly	A	G227V	Loss of ligand binding	−0.96	1.8	0.99	1.7
4az3	A	S23Y	Loss of phosphorylation	−0.95	1.8	1.00	0.7
2nt0	A	D380H	Decrease of enzymatic activity	−0.86	0.3	1.00	3.1

Table 3. List of the 10 mutations from the **S** dataset which are annotated as disease causing and are predicted as the most stabilizing (*i.e.* with the most negative $\Delta\Delta G$ values) by PoPMuSiC. All variants are predicted as neutral by the PoPMuSiC- and HoTMuSiC-based classifiers while they are predicted as deleterious by ANN and SNPMuSiC.

Reliability of structural stability prediction. Our SNPMuSiC method reaches the very high positive predicted value (PPV) of 89%, as shown in Table 1, which indicates that when a mutation is predicted as deleterious, it has a large probability of being so. The negative predictive value is much lower, 52%, indicating that when a mutation is predicted as neutral, it has still 50% chance of being disease-causing, albeit for a reason that is not related to stability.

The idea here is thus to identify a subset of mutations, whose pathogenicity index $\mathcal{J}$ falls above a certain confidence threshold and whose prediction reaches an even higher level of accuracy. For these mutations, the protein stability is the main contributing factor to the variant's deleteriousness. We also identified another subset of mutations, with $\mathcal{J}$ below another confidence threshold, which are predicted as neutral with a good accuracy. This strategy is quite useful in the perspective of combining SNPMuSiC with other variant effect predictors that are based on different biophysical characteristics such as aggregation propensities or flexibility.

Setting the $\mathcal{J}_D$ confidence threshold equal to 0.9, we find that the PPV reaches 95% on the subset $\mathbf{S}_{high}$ with $\mathcal{J} \geq \mathcal{J}_D$. This subset contains about 1,500 mutations, which corresponds to 27% of the whole **S** set. Here we can be almost sure of the deleteriousness of the mutations and of its molecular cause. Indeed, combining the SNPMuSiC prediction with the outputs of PoPMuSiC and HoTMuSiC, we can determine if they are deleterious due to thermodynamic or thermal stabilization or destabilization.

With the $\mathcal{J}_N$ threshold for neutral mutations equal to 0.5, we obtain an NPV of 72% on the subset $\mathbf{S}_{low}$ for which $\mathcal{J} \leq \mathcal{J}_N$. This subset contains about 700 mutations, thus about 13% of the total **S** set. For these mutations, we have a good indication of their neutral impact on the phenotype.

The BACC score computed on the subsets characterized by $\mathcal{J} \geq \mathcal{J}_D$ and $\mathcal{J} \leq \mathcal{J}_N$, representing 40% of the **S** set, is also significantly improved and reaches 0.87. The SNPMuSiC distribution for deleterious and neutral variants at both sides of the confidence thresholds are shown in Fig. 5.

Conclusion

Here we made a further step towards the prediction of the biophysical causes of the deleteriousness of non-synonymous SNVs. We focused on protein stability which guides a series of crucial biophysical mechanisms that encompass folding, interactions, and function. Our analysis can be extended and generalized to all other types of biophysical characteristics that play a role at the protein level and could be at the basis of deleteriousness, such as solubility, aggregation, allostery, flexibility, and catalytic activity. This will lead to a complete view of the relation between the effect of SNVs at the molecular level and disease, and pave the way towards personalized medecine.

Let us give a flavor of the next stages of our quest. We will include the changes caused by mutations on the protein aggregation properties, which are not directly related to protein stability[43] even though they are frequently considered as correlated. Other factors that we will add to this analysis is the SNV effect on protein activity and on protein-protein interaction networks[40,44], which are undoubtedly important factors in the initiation of diseases.

Finally we would like to underline the importance of using the 3D structures of the proteins in which SNV occur to predict and interpret their biophysical and disease-causing effects. This data is essentially overlooked in most commonly used deleteriousness classification methods, which drastically limits their interpretative power. Two reasons can be invoked to explain this shortcoming. On the one hand, protein sequences are much easier to handle than protein structures, and on the other hand, many proteins have no available experimental structure. This last issue has, however, to be relativized, as many protein structures can be obtained through comparative modeling. Given that our methods use coarse-grained representations of protein structure, they can be applied on protein models with only a small loss of accuracy[45]. We would like to end by stressing the broad applicability of structure-based predictors, as it has been estimated that almost half of all structured proteins have either an experimental or reliably modeled structure[46].

References

1. Stenson, P. D. *et al*. The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Human Genetics* **136**, 665–677 (2017).
2. Niu, B. *et al*. Protein-structure-guided discovery of functional mutations across 19 cancer types. *Nature Genetics* **48**, 827–837 (2016).
3. Choi, Y., Sims, G. E., Murphy, S., Miller, J. R. & Chan, A. P. Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS One* **7**, e46688 (2012).
4. Choi, Y. & Chan, A. P. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **31**, 2745–2747 (2015).
5. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4**, 1073–1081 (2009).
6. Katsonis, P. & Lichtarge, O. A formal perturbation equation between genotype and phenotype determines the evolutionary action of protein-coding variations on fitness. *Genome Research* **24**, 2050–2058 (2014).
7. Gallion, J. *et al*. Predicting phenotype from genotype: Improving accuracy through more robust experimental and computational modeling. *Human Mutation* **38**, 569–580 (2017).
8. Schwarz, J. M., Rödelsperger, C., Schuelke, M. & Seelow, D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* **7**, 575–576 (2010).
9. Reva, B., Antipin, Y. & Sander, C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res.* **39**, e118 (2011).
10. Kircher, M. *et al*. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014).
11. Adzhubei, I. A. *et al*. A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
12. Baugh, E. H. *et al*. Robust classification of protein variation using structural modelling and large-scale data integration. *Nucleic Acids Res.* **44**, 2501–2513 (2016).
13. Raimondi, D., Gazzo, A. M., Rooman, M., Lenaerts, T. & Vranken, W. F. Multilevel biological characterization of exomic variants at the protein level significantly improves the identification of their deleterious effects. *Bioinformatics* **32**, 1797–1804 (2016).
14. Raimondi, D. *et al*. DEOGEN2: prediction and interactive visualization of single amino acid variant deleteriousness in human proteins. *Nucleic Acids Res.* **45**, W201–W206 (2017).
15. Yates, C. M., Filippis, I., Kelley, L. A. & Sternberg, M. J. SuSPect: enhanced prediction of single amino acid variant (SAV) phenotype using network features. *J Mol Biol.* **426**, 2692–701 (2014).
16. Miosge, L. A. *et al*. Comparison of predicted and actual consequences of missense mutations. *Proc. Natl. Acad. Sci. USA* **112**, E5189–98 (2015).
17. Pires, D. E. V., Chen, J., Blundell, T. L. & Ascher, D. B. In silico functional dissection of saturation mutagenesis: Interpreting the relationship between phenotypes and changes in protein stability, interactions and activity. *Scientific Reports* **6**, 19848 (2016).
18. Saunders, C. T. & Baker, D. Evaluation of structural and evolutionary contributions to deleterious mutation prediction. *J. Mol. Biol.* **322**, 891–901 (2002).
19. Stefl, S., Nishi, H., Petukh, M., Panchenko, A. R. & Alexov, E. Molecular mechanisms of disease-causing missense mutations. *J. Mol. Biol.* **425**, 3919–3936 (2013).
20. Sherry, S. T. *et al*. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
21. Mottaz, A., P.A. David, F., Veuthey, A.-L. & Yip, Y. L. Easy retrieval of single amino-acid polymorphisms and phenotype information using SwissVar. *Bioinformatics* **26**, 851–852 (2010).
22. Wu, C. H. *et al*. The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res.* **34**, D187–D191 (2006).
23. Velankar, S. *et al*. SIFTS: Structure Integration with Function, Taxonomy and Sequences resource. *Nucleic Acids Res* **41**, D483–D489 (2013).
24. Berman, H. M. *et al*. The Protein Data Bank. *Nucleic Acids Research* **28**, 235–242 (2000).
25. Landrum, M. J. *et al*. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res* **42**, D980–5 (2014).
26. O'Roak, B. J. *et al*. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **485**, 246–250 (2012).
27. Miyazawa, S. & Jernigan, R. L. Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation. *Macromolecules* **18**, 534–552 (1985).
28. Sippl, M. J. Calculation of conformational ensembles from potentials of mena force. *J. Mol. Biol.* **213**, 859–883 (1990).
29. Rooman, M., Kocher, J. P. & Wodak, S. Prediction of protein backbone conformation based on seven structure assignments: Influence of local interactions. *J. Mol. Biol.* **221**, 961–979 (1991).
30. Kocher, J. P., Rooman, M. J. & Wodak, S. J. Factors influencing the ability of knowledge-based potentials to identify native sequence-structure matches. *J. Mol. Biol.* **235**, 1598–1613 (1994).
31. Dehouck, Y., Gilis, D. & Rooman, M. A new generation of statistical potentials for proteins. *Biophys. J.* **90**, 4010–4017 (2006).
32. Dehouck, Y. *et al*. Fast and accurate predictions of protein stability changes upon mutations using statistical potentials and neural networks: PoPMuSiC-2.0. *Bioinformatics* **25**, 2537–2543 (2009).
33. Pucci, F., Bourgeas, R. & Rooman, M. Predicting protein thermal stability changes upon point mutations using statistical potentials: Introducing HoTMuSiC. *Sci. Rep.* **6**, 23257 (2016).
34. Dehouck, Y., Kwasigroch, J. M., Gilis, D. & Rooman, M. PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. *BMC Bioinformatics* **12**, 151 (2011).
35. Pucci, F., Bourgeas, R. & Rooman, M. High-quality thermodynamic data on the stability changes of proteins upon single-site mutations. *Journal of Physical and Chemical Reference Data* **45**, 023104 (2016).
36. Tokuriki, N. & Tawfik, D. S. Stability effects of mutations and protein evolvability. *Curr Opin Struct Biology* **19**, 596–604 (2009).
37. Worth, C. L., Preissner, R. & Blundell, T. L. SDM - a server for predicting effects of mutations on protein stability and malfunction. *Nucleic Acids Res* **39**, W215–W222 (2011).
38. Specht, D. F. Probabilistic neural networks. *Neural Netw.* **3**, 109–118 (1990).
39. DeLong, E. R., DeLong, D. M. & Clarke-Pearson, D. L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* **44**, 837–845 (1988).
40. Vidal, M., Cusick, M. E. & Barabasi, A. L. Interactome Networks and Human Disease. *Cell.* **144**, 986–998 (2011).
41. Pires, D. E., Blundell, T. L. & Ascher, D. B. mCSM-lig: quantifying the effects of mutations on protein-small molecule affinity in genetic disease and emergence of drug resistance. *Sci. Rep.* **6**, 29575 (2016).
42. Reimand, J., Wagih, O. & Bader, G. D. Evolutionary constraint and disease associations of post-translational modification sites in human genomes. *PLoS Genet* **11**, e1004919 (2015).
43. Ganesan, A. *et al*. Structural hot spots for the solubility of globular proteins. *Nat Commun* **24**, 10816 (2016).
44. Yates, C. M. & Sternberg, M. J. The Effects of Non-Synonymous Single Nucleotide Polymorphisms (nsSNPs) on Protein-Protein Interactions. *Journal of Molecular Biology* **425**, 3949–3963 (2013).
45. Gonnelli, G., Rooman, M. & Dehouck, Y. Structure-based mutant stability predictions on proteins of unknown structure. *J Biotechnol* **161**, 287–93 (2012).
46. Fiser, A. Template-Based Protein Structure Modeling. *Methods Mol Biol.* **673**, 73–94 (2010).

Acknowledgements
We thank Georges Coppin, Suchsia Chao and Thanh Phuong Pham for participating in the early stages of this project, and Jean Marc Kwasigroch, Daniele Raimondi and Andrea Gazzo for help in setting up the training dataset.

Author Contributions
F.P. and M.R. conceived the experiments, F.A., F.P. and M.G. conducted the experiments, F.A., F.P. and M.R. analyzed the results, F.P. and M.R. wrote the paper. All authors reviewed the manuscript.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-018-22531-2.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.