Table S1 Tools evaluated for predicting protein stability change.

Tool	Published	Execution time	Availability	Integrated
I-Mutant2.0	2005	1 sec.	Stand alone	V
MUpro	2006	1 sec.	Online	V
CUPSAT	2006	1 sec.	Online	V
iPTREE-STAB	2007	1 sec.	Online	V
PoPMuSiC	2009	2 min.	Online	V
AUTO-MUTE2.0	2010	1 sec.	Stand alone	V
MuStab	2010	NA	Online	
SDM	2011	1 min.	Online	V
mCSM	2014	1 min.	Online	V
DUET	2014	1 min.	Online	V
MAESTRO	2015	1 sec.	Stand alone	V
INPS	2015	>20min.	Online	
EASE-MM	2016	>1hr.	Online	
ELASPIC	2016	>10min.	Online	
SDM2	2017	1min.	Online	V
TopologyNet	2017	>1hr.	Online	
TML	2017	>1hr.	Online	
DynaMut	2018	>10min.	Online	

Table S2 Features from the prediction results using different protein sequence tools.

Tool	Prediction results	Feature bits
I-Mutant2.0_SEQ	Prediction, ddG value, Reliability index	3
MUpro_SVM	Prediction, Confidence score	2
MUpro_NN	Prediction, Confidence score	2
iPTREE-STAB	Prediction, ddG value	2

I-Mutant2.0_SEQ: I-Mutant 2.0 prediction strategy using protein sequence.
MUpro_SVM: MUpro SVM prediction model.
MUpro_NN: MUpro NN prediction model.
Prediction: decrease or increase.

Table S3 Features from the prediction results using different protein structure tools.

Tool	Prediction results	Feature bits
I-Mutant2.0_PDB	Prediction, ddG value, RSA, Reliability index	4
CUPSAT	Prediction, ddG value, RSA	3
PoPMuSiC	Prediction, ddG value, SS, RSA	10
AUTO-MUTE2.0_ SVM	Prediction, ddG value, RSA, SS, Confidence score	9
AUTO-MUTE2.0_ RF/TR	Prediction, ddG value, RSA, SS, Confidence score	9
SDM	Prediction, ddG value	2
DUET	Prediction, ddG value	2
mCSM	Prediction, ddG value	2
MAESTRO	Prediction, ddG value, ddG Confidence, Score, delta_score	5
SDM2	Prediction, ddG value, Wild-type_SS, Wild-type_RSA, Wild-type_depth, Wild-type_OSP, Wild-type_H-hond, Mutant_SS, Mutant_RSA, Mutant_depth, Mutant_OSP, Mutant_ H-hond	32

I-Mutant2.0_PDB: I-Mutant 2.0 prediction strategy using PDB ID.
AUTO-MUTE2.0_RF/TR: AUTO-MUTE 2.0 Random Forest prediction model; Prediction and Confidence scores from Random Forest, ddG value from TR (Tree Regression: REPTree).
AUTO-MUTE2.0_SVM: AUTO-MUTE 2.0 SVM prediction model; Prediction and Confidence score from SVM, ddG value from SVR.
Prediction: decrease or increase.
RSA: relative solvent accessible area.
SS: secondary structure.

Table S4 20-bit amino acid encoding.

Amino acid	1-Letter	Encoding
Alanine	A	00000000000000000001
Cysteine	C	00000000000000000010
Aspartic acid	D	00000000000000000100
Glutamic acid	E	00000000000000001000
Phenylalanine	F	00000000000000010000
Glycine	G	00000000000000100000
Histidine	H	00000000000001000000
Isoleucine	I	00000000000010000000
Lysine	K	00000000000100000000
Leucine	L	00000000001000000000
Methionine	M	00000000010000000000
Asparagine	N	00000000100000000000
Proline	P	00000001000000000000
Glutamine	Q	00000010000000000000
Arginine	R	00000100000000000000
Serine	S	00001000000000000000
Threonine	T	00010000000000000000
Valine	V	00100000000000000000
Tryptophan	W	01000000000000000000
Tyrosine	Y	10000000000000000000
Gap	-	00000000000000000000

Table S5 William *et al.* encoding.

Amino acid	Polarity	Secondary structure	Molecular size or volume	Codon diversity	Electrostatic charge
A	-0.591	-1.302	-0.733	1.570	-0.146
C	-1.343	0.465	-0.862	-1.020	-0.255
D	1.050	0.302	-3.656	-0.259	-3.242
E	1.357	-1.453	1.477	0.113	-0.837
F	-1.006	-0.59	1.891	-0.397	0.412
G	-0.384	1.652	1.330	1.045	2.064
H	0.336	-0.417	-1.673	-1.474	-0.078
I	-1.239	-0.547	2.131	0.393	0.816
K	1.831	-0.561	0.533	-0.277	1.648
L	-1.019	-0.987	-1.505	1.266	-0.912
M	-0.663	-1.524	2.219	-1.005	1.212
N	0.945	0.828	1.299	-0.169	0.933
P	0.189	2.081	-1.628	0.421	-1.392
Q	0.931	-0.179	-3.005	-0.503	-1.853
R	1.538	-0.055	1.502	0.440	2.897
S	-0.228	1.399	-4.760	0.670	-2.647
T	-0.032	0.326	2.213	0.908	1.313
V	-1.337	-0.279	-0.544	1.242	-1.262
W	-0.595	0.009	0.672	-2.128	-0.184
Y	0.260	0.830	3.097	-0.838	1.512
-	0.000	0.000	0.000	0.000	0.000

Table S6 Mathura *et al.* encoding.

Amino Acid	Hydrophobicity	Side chain length	α-helix propensity	Number of codons	β-strand propensity
A	0.008	0.134	-0.475	-0.039	0.181
R	0.171	-0.361	0.107	-0.258	-0.364
N	0.255	0.038	0.117	0.118	-0.055
D	0.303	-0.057	-0.014	0.225	0.156
C	-0.132	0.174	0.070	0.565	-0.374
Q	0.149	-0.184	-0.030	0.035	-0.112
E	0.221	-0.280	-0.315	0.157	0.303
G	0.218	0.562	-0.024	0.018	0.106
H	0.023	-0.177	0.041	0.280	-0.021
I	-0.353	0.071	-0.088	-0.195	-0.107
L	-0.267	0.018	-0.265	-0.274	0.206
K	0.243	-0.339	-0.044	-0.325	-0.027
M	-0.239	-0.141	-0.155	0.321	0.077
F	-0.329	-0.023	0.072	-0.002	0.208
P	0.173	0.286	0.407	-0.215	0.384
S	0.199	0.238	-0.015	-0.068	-0.196
T	0.068	0.147	-0.015	-0.132	-0.274
W	-0.296	-0.186	0.389	0.083	0.297
Y	-0.141	-0.057	0.425	-0.096	-0.091
V	-0.274	0.136	-0.187	-0.196	-0.299
-	0.000	0.000	0.000	0.000	0.000

Table S7 Comparison of the machine learning algorithms in the classification models trained with S3568.

Model	Type	Method	Sn	Sp	Acc	MCC
Structure-based	Tree	XGBoost	0.758	0.964	0.908	0.758
	Function	SVM	0.737	0.961	0.901	0.738
	Meta	EnsembleSelection	0.733	0.960	0.900	0.735
	Meta	RandomSubSpace	0.707	0.969	0.900	0.734
	Rules	FURIA	0.715	0.966	0.899	0.732
	Tree	SysFor	0.751	0.947	0.895	0.724
	Tree	SimpleCart	0.733	0.954	0.895	0.723
	Tree	ForestPA	0.707	0.964	0.896	0.723
	Tree	BFTree	0.725	0.955	0.894	0.719
	Meta	Decorate	0.767	0.937	0.892	0.717
	Tree	LMT	0.730	0.951	0.892	0.715
Sequence-based	Tree	XGBoost	0.670	0.953	0.877	0.672
	Function	SVM	0.674	0.946	0.874	0.662
	Tree	ForestPA	0.641	0.956	0.872	0.655
	Meta	RandomSubSpace	0.599	0.969	0.871	0.650
	Meta	EnsembleSelection	0.586	0.967	0.866	0.636
	Rules	MODLEM	0.619	0.954	0.865	0.635
	Tree	SysFor	0.660	0.936	0.863	0.633
	Meta	Decorate	0.685	0.923	0.860	0.630
	Rules	FURIA	0.604	0.957	0.863	0.628
	Lazy	LWL	0.551	0.975	0.862	0.626
	Tree	LADTree	0.601	0.957	0.862	0.626
	Bayes	BayesianLogisticRegression	0.655	0.933	0.859	0.624

Table S8 Comparison of the machine learning algorithms in the regression models trained with S3568.

Model	Type	Method	PCC
Structure-based	Tree	XGBoost	0.864
	Tree	RandomForest	0.833
	Meta	RandomSubSpace	0.832
	Tree	M5P	0.828
	Meta	EnsembleSelection	0.823
	Tree	AlternatingModelTree	0.821
	Rule	M5Rules	0.821
	Function	PaceRegression	0.819
	Meta	RandomCommittee	0.814
	Meta	MultiSearch	0.812
	Function	GaussianProcesses	0.810
Sequence-based	Tree	XGBoost	0.818
	Tree	RandomForest	0.798
	Meta	RandomCommittee	0.779
	Meta	RandomSubSpace	0.752
	Tree	M5P	0.713
	Function	PaceRegression	0.707
	Tree	AlternatingModelTree	0.706
	Function	GaussianProcesses	0.704
	Meta	MultiSearch	0.701
	Function	LinearRegression	0.700
	Rule	M5Rules	0.692

Table S9 Ranking of important features with f-scores greater than 100 in the structure-based classification model.

Feature name	F-score	Feature name	F-score
MAESTRO_ddG	682	PBP_W-M_α-helix propensity	188
MAESTRO_ddG_Conf.	568	PBP_W-M_β-strand propensity	178
AUTO-MUTE2.0_FR	456	I-Mutant2.0_PDB_Reliability index	176
AUTO-MUTE2.0_FR_Conf.	442	PBP_W-M_Polarity	174
PoPMuSiC_ddG	432	I-Mutant2.0_PDB	166
DUET_ddG	412	PBP_W-M_Hydrophobicity	156
pH	404	SDM2_Mutant_OSP	142
TEMPERATURE	374	PBP_W-M_Codon diversity	138
I-Mutant2.0_SEQ_ddG	360	PBP_W-M_Side chain length	136
AUTO-MUTE2.0_SVM_Conf	336	PBP_W-M_Secondary structure	134
AUTO-MUTE2.0_SVM_DDG	330	PBP_W-M_Electrostatic charge	134
MUpro_NN_Conf.	320	NetSurfP_W13_ASA	122
CUPSAT_ddG	290	NetSurfP_W3_RSA	120
MUpro_SVM_Conf.	286	NetSurfP_W1_Coil	116
iPTREE-STAB_ddG	266	NetSurfP_W8_ASA	112
I-Mutant2.0_PDB_ddG	266	PBP_M_ Side chain length	110
MAESTRO_delta_score	256	NetSurfP_W8_Z-score	110
SDM_ddG	248	PBP_W-M_ Number of codons	110
mCSM_ddG	228	NetSurfP_W18_RSA	108
MAESTRO_score	210	NetSurfP_Alpha-Helix[*]	102
AUTO-MUTE2.0_TR_ddG	208	NetSurfP_W15_ASA	102
NetSurfP_Coil[*]	198	I-Mutant2.0_SEQ_Reliability index	100
NetSurfP_Buried[*]	194		

[*]we calculated the frequency of the three secondary structures in the entire protein.
OSP: residue occluded surface.
PBP: physicochemical and biochemical properties.
Tool name: decrease or increase.
RSA: relative surface accessibility.
ASA: absolute surface accessibility.
W-M: differences between the attributes of the wild-type and mutant amino acids.
W+number: number = {1, 2, 3, 4, 5, 6, 7, 8, 9, mutant, 10, 11, 12, 13, 14, 15, 16, 17, 18} amino acids position.
F-score: the number of times a feature is used to split the data across all trees (XGBoost).

Table S10 Ranking of important features with f-scores greater than 100 in the sequence-based classification model.

Feature name	F-score
pH	624
MUpro_SVM_Conf.	616
I-Mutant2.0_SEQ_ddG	600
iPTREE-STAB_ddG	438
TEMPERATURE	434
MUpro_NN_Conf.	402
I-Mutant2.0_SEQ_Reliability index	258
PBP_W-M_ Hydrophobicity	220
PBP_W-M_β-strand propensity	188
PBP_W-M_ Side chain length	170
PBP_W-M_Electrostatic charge	168
PBP_W-M_ Number of codons	162
NetSurfP_Buried*	134
PBP_W-M_Secondary structure	130
iPTREE-STAB	122
PBP_W-M_Codon diversity	118
NetSurfP_W17_RSA	114
PBP_W-M_Molecular size or volume	114
PBP_W-M_Polarity	112
NetSurfP_Alpha-Helix*	110
NetSurfP_W17_Z-score	106
NetSurfP_W5_RSA	100

*we calculated the frequency of the three secondary structures in the entire protein.
PBP: physicochemical and biochemical properties.
Tool name: decrease or increase.
RSA: relative surface accessibility.
W-M: differences between the attributes of the wild-type and mutant amino acids.
W+number: number = {1, 2, 3, 4, 5, 6, 7, 8, 9, mutant, 10, 11, 12, 13, 14, 15, 16, 17, 18} amino acids position.
F-score: the number of times a feature is used to split the data across all trees (XGBoost).

Table S11 Ranking of important features with f-scores greater than 100 in the structure-based regression model.

Feature name	F-score
AUTO-MUTE2.0_SVM_ddG	394
iPTREE-STAB_ddG	153
I-Mutant2.0_PDB_ddG	140
AUTO-MUTE_TR_ddG	137
TEMPERATURE	131
MAESTRO_ddG	130
DUET_ddG	119
mCSM_ddG	116
AUTO-MUTE2.0_SVM_Conf.	109
PoPMuSiC_ddG	107
AUTO-MUTE2.0_FR_Conf.	106

F-score: The number of times a feature is used to split the data across all trees (XGBoost).

Table S12 Ranking of important features with f-scores greater than 100 in the sequence-based regression model.

Feature name	F-score
I-Mutant2.0_SEQ_ddG	331
iPTREE-STAB_ddG	329
PBP_W-M_Hydrophobicity	228
pH	196
MUpro_NN_Conf.	187
PBP_W-M_β-strand propensity	161
TEMPERATURE	158
NetSurfP_Beta-strand[*]	150
PBP_W-M_Number of codons	135
PBP_W-M_α-helix propensity	125
PBP_W-M_Secondary structure	119
PBP_W-M_Polarity	116
NetSurfP_Buried[*]	111
NetSurfP_W19_Z-score	109
PBP_W-M_Molecular size or volume	106

[*]we calculated the frequency of the three secondary structures in the entire protein.
PBP: physicochemical and biochemical properties.
W-M: differences between the attributes of the wild-type and mutant amino acids.
F-score: the number of times a feature is used to split the data across all trees (XGBoost).

Table S13 Ranking of important features with f-scores greater than 100 in iStable2.0_PBD_SASC.

Feature name	F-score
MAESTRO_ddG	516
AUTO-MUTE2.0_FR_Conf	410
AUTO-MUTE2.0_FR	328
MAESTRO_ddG_Conf.	286
TEMPERATURE	282
AUTO-MUTE2.0_SVM_Conf.	258
I-Mutant2.0_SEQ_ddG	244
AUTO-MUTE2.0_SVM_ddgG	232
I-Mutant2.0_PDB_ddG	216
MAESTRO_delta_score	206
NetSurfP_Coil*	152
I-Mutant2.0_SEQ_Reliability index	150
MAESTRO_score	148
PBP_W-M_α-helix propensity	134
pH	124
AUTO-MUTE2.0_FR_ddG	124
PBP_W-M_Polarity	120
I-Mutant2.0_PDB	106
PBP_W-M_Number of codons	102
PBP_W-M_Codon diversity	100

*we calculated the frequency of the three secondary structures in the entire protein.
PBP: physicochemical and biochemical properties.
Tool name: decrease or increase.
W-M: differences between the attributes of the wild-type and mutant amino acids.
F-score: the number of times a feature is used to split the data across all trees (XGBoost).

Table S14 Ranking of important features with f-scores greater than 100 in iStable2.0_SEQ_SASC.

Feature name	F-score	Feature name	F-score
pH	2740	PBP_W8_Electrostatic charge	160
I-Mutant2.0_SEQ_ddG	2228	NetSurfP_W16_ASA	158
TEMPERATURE	1988	NetSurfP_Beta-strand*	158
I-Mutant2.0_SEQ_Reliability index	1344	NetSurfP_W17_ASA	156
PBP_W-M_Hydrophobicity	724	NetSurfP_W4_RSA	152
PBP_W-M_Secondary structure	716	PBP_W7_Number of codons	152
PBP_W-M_β-strand propensity	708	NetSurfP_W7_ASA	150
PBP_W-M_Electrostatic charge	700	NetSurfP_W13_ASA	148
PBP_W-M_Side chain length	592	NetSurfP_W3_ASA	146
PBP_W-M_Codon diversity	576	NetSurfP_W2_RSA	146
PBP_W-M_Molecular size or volume	554	NetSurfP_W9_ASA	142
PBP_W-M_Number of codons	548	PBP_W18_Secondary structure	142
PBP_W-M_α-helix propensity	474	NetSurfP_W1_Coil	138
PBP_W-M_Polarity	440	NetSurfP_W14_ASA	138
I-Mutant2.0_SEQ	420	NetSurfP_W4_Z-score	138
PBP_Mutant_Secondary structure	394	NetSurfP_W9_Coil	138
NetSurfP_Buried*	392	NetSurfP_W14_RSA	136
PBP_Mutant_β-strand propensity	360	NetSurfP_W8_ASA	134
PBP_Mutant_Polarity	350	PBP_W18_α-helix propensity	130
PBP_Mutant_Number of codons	338	NetSurfP_W19_Zscore	130
PBP_Mutant_Molecular size or volume	312	PBP_W8_Molecular size or volume	130
PBP_Mutant_Electrostatic charge	308	PBP_W9_β-strand propensity	130
NetSurfP_W5_RSA	270	NetSurfP_W14_Coil	128
NetSurfP_W5_Z-score	268	PBP_W9_β-strand propensity	128
NetSurfP_W12_ASA	264	NetSurfP_W11_RSA	124
NetSurfP_W15_ASA	260	PBP_W19_α-helix propensity	124
PBP_Mutant_Hydrophobicity	252	NetSurfP_W18_ASA	122
NetSurfP_Alpha-Helix*	248	NetSurfP_W2_ASA	120
NetSurfP_W10_RSA	246	NetSurfP_W19_RSA	120
PBP_Mutant_Side chain length	246	PBP_W3_Molecular size or volume	120
NetSurfP_W15_Zscore	240	PBP_W3_Codon diversity	120
PBP_Mutant_Codon diversity	236	PBP_W8_Polarity	118
NetSurfP_W5_ASA	236	NetSurfP_W12_Coil	118
NetSurfP_W10_ASA	232	PBP_W4_Hydrophobicity	116
NetSurfP_W17_Z-score	232	PBP_W15_Codon diversity	114
NetSurfP_W11_Z-score	230	PBP_W16_Codon diversity	114

NetSurfP_W6_Z-score	230	PBP_W5_Hydrophobicity	114
NetSurfP_W14_Z-score	224	PBP_W5_Secondary structure	110
NetSurfP_W7_RSA	222	PBP_W2_Number of codons	110
NetSurfP_W8_Z-score	220	NetSurfP_W7_Coil	108
PBP_Mutant_α-helix propensity	218	PBP_W7_β-strand propensity	108
NetSurfP_W9_Z-score	218	NetSurfP_W13_RSA	106
NetSurfP_W1_RSA	218	PBP_W7_α-helix propensity	106
NetSurfP_W7_Z-score	210	NetSurfP_W19_ASA	106
NetSurfP_W10_Z-score	204	NetSurfP_W18_RSA	104
NetSurfP_W2_Z-score	202	PBP_W17_Codon diversity	104
NetSurfP_W1_ASA	200	PBP_W6_Codon diversity	104
NetSurfP_Coil*	200	PBP_W5_Codon diversity	104
NetSurfP_W9_RSA	200	NetSurfP_W12_Z-score	104
NetSurfP_W1_Z-score	196	PBP _W6_Polarity	104
NetSurfP_W13_Z-score	194	PBP _W1_Polarity	102
NetSurfP_W16_Z-score	192	NetSurfP_W6_RSA	102
NetSurfP_W15_RSA	190	PBP _W11_Molecular size or volume	102
NetSurfP_W8_RSA	190	PBP _W4_Molecular size or volume	100
NetSurfP_W4_ASA	188	NetSurfP_W7_Beta-strand	100
NetSurfP_W12_RSA	182	PBP _W17_Molecular size or volume	100
NetSurfP_W18_Z-score	178	NetSurfP_W6_ASA	100
NetSurfP_W16_RSA	174	PBP _W19_Molecular size or volume	100
NetSurfP_W3_Z-score	172	PBP _W9_Molecular size or volume	100
NetSurfP_W17_RSA	170	PBP _W1_Number of codons	100
NetSurfP_W3_RSA	162		

*we calculated the frequency of the three secondary structures in the entire protein.
PBP: physicochemical and biochemical properties.
Tool name: decrease or increase.
RSA: relative surface accessibility.
ASA: absolute surface accessibility.
W-M: differences between the attributes of the wild-type and mutant amino acids.
W+number: number = {1, 2, 3, 4, 5, 6, 7, 8, 9, mutant, 10, 11, 12, 13, 14, 15, 16, 17, 18} amino acids position.
F-score: the number of times a feature is used to split the data across all trees (XGBoost).

Table S15 Ranking of important features with f-scores greater than 100 in iStable2.0_PDB_Regression_SASC

Feature name	F-score
I-Mutant2.0_PDB_ddG	554
MAESTRO_ddG	533
AUTO-MUTE2.0_SVM_ddG	514
TEMPERATURE	437
MAESTRO_ddG_Conf.	429
I-Mutant2.0_SEQ_ddG	357
MAESTRO_delta_score	337
MAESTRO_score	322
AUTO-MUTE2.0_TR_ddG	311
AUTO-MUTE2.0_SVM_Conf.	287
pH	240
PBP_W-M_β-strand propensity	215
AUTO-MUTE2.0_FR_Conf.	208
PBP_W-M_Hydrophobicity	194
PBP_W-M_α-helix propensity	175
PBP_W-M_Number of codons	171
PBP_W-M_Codon diversity	169
NetSurfP_Buried[*]	166
PBP_W-M_Polarity	162
I-Mutant2.0_SEQ_Reliability index	147
PBP_W-M_Molecular size or volume	146
I-Mutant2.0_PDB_Reliability index	141
PBP_W-M_Side chain length	138
I-Mutant2.0_PDB_RSA	133
PBP_W-M_Electrostatic charge	127
NetSurfP_Beta-strand[*]	126
NetSurfP_W15_Z-score	106
PBP_W-M_Secondary structure	105

[*]we calculated the frequency of the three secondary structures in the entire protein.
PBP: physicochemical and biochemical properties.
RSA: relative surface accessibility.
W-M: differences between the attributes of the wild-type and mutant amino acids.
F-score: the number of times a feature is used to split the data across all trees (XGBoost).

Table S16 Ranking of important features with f-scores greater than 100 in iStable2.0_SEQ_Regression_SASC

Feature name	F-score
I-Mutant2.0_SEQ_ddG	690
TEMPERATURE	689
pH	580
PBP_W-M_β-strand propensity	326
PBP_W-M_Hydrophobicity	321
I-Mutant2.0_SEQ_Reliability index	259
PBP_W-M_Side chain length	256
PBP_W-M_Number of codons	248
PBP_W-M_Polarity	219
PBP_W-M_Secondary structure	213
PBP_W-M_Molecular size or volume	206
PBP_W-M_α-helix propensity	182
PBP _W-M_Codon diversity	162
PBP _W-M_Electrostatic charge	159
NetSurfP_W11_Z-score	138
NetSurfP_Beta-strand[*]	135
NetSurfP_Buried[*]	134
NetSurfP_W19_Z-score	133
PBP_M_Polarity	129
NetSurfP_W10_RSA	114
NetSurfP_W13_Z-score	111
NetSurfP_W9_RSA	108
NetSurfP_W1_ASA	108
NetSurfP_W10_ASA	105

[*]we calculated the frequency of the three secondary structures in the entire protein.
PBP: physicochemical and biochemical properties.
RSA: relative surface accessibility.
ASA: absolute surface accessibility.
W-M: differences between the attributes of the wild-type and mutant amino acids.
W+number: number = {1, 2, 3, 4, 5, 6, 7, 8, 9, mutant, 10, 11, 12, 13, 14, 15, 16, 17, 18} amino acids position.
F-score: the number of times a feature is used to split the data across all trees (XGBoost).

Table S17 Comparison of feature usage in different models.

		iStable2.0_SEQ		iStable2.0_SEQ_SASC		iStable2.0_SEQ_Regression		iStable2.0_SEQ_Regression_SASC	
		Total	Used	Total	Used	Total	Used	Total	Used
Online Server Module	iPTREE-STAB	2	2	0	0	2	2	0	0
	MUpro_SVM	2	2	0	0	2	1	0	0
	MUpro _NN	2	2	0	0	2	1	0	0
Stand-alone Module	I-Mutant2.0_SEQ	3	3	3	3	3	3	3	3
Sequence Coding Module	pH	1	1	1	1	1	1	1	1
	Temperature	1	1	1	1	1	1	1	1
	AAC	400	162	400	275	400	121	400	270
	PBP	210	209	210	210	210	204	210	210
	SS, SA	271	135	271	151	271	125	271	136
	Total of SCM	883	508	883	638	883	452	883	618

AAC: amino acid coding.
PBP: physicochemical and biochemical properties.
SS, SA: secondary structure and surface accessibility.

Table S18 Performance of Majority Voting with S630.

Method	Strategy	Sn	Sp	Acc	MCC
Majority Voting	A	0.405	0.970	0.824	0.493
	B	0.258	0.981	0.794	0.383

Majority Voting: The 14 of voters contains DUET, SDM, SDM2, mCSM, CUPSAT, I-Mutant2.0_PDB, PoPMuSiC, AUTO-MUTE2.0_SVM, RF, MAESTRO, iPTREE-STAB, I-Mutant2.0_SEQ, MUpro_SVM and MUpro_NN.
A: Predicting stable mutant when obtained the same number of votes.
B: Predicting instable mutant when obtained the same number of votes.