

Table A1 Final combinations of predictive features

Model	Feature
EASE-AA ₂	SIFT score
EASE-AA ₂	Δ hydrophobicity
EASE-AA ₂	Δ isoelectric point
EASE-AA ₂	helix tendency (mean, min, max)
EASE-AA ₂	Δ compressibility
EASE-AA ₂	sheet tendency (mean, min, max)
EASE-AA ₂	ΔM
EASE-AA ₂	Δ volume
EASE-AA ₂	Δ steric
<i>exposed</i>	Δ helix tendency
<i>exposed</i>	secondary structure probabilities
<i>exposed</i>	Δ steric parameter
<i>exposed</i>	SIFT score
<i>exposed</i>	flexibility (mean, min, max)
<i>exposed</i>	Δ compressibility
<i>exposed</i>	Δ hydrophobicity
<i>buried</i>	Δ hydrophobicity
<i>buried</i>	SIFT score
<i>buried</i>	Δ isoelectric point
<i>buried</i>	compressibility (mean, min, max)
<i>buried</i>	ΔM
<i>buried</i>	helix tendency (mean, min, max)
<i>buried</i>	hydrophobicity (mean, min, max)
<i>buried</i>	polarisability (mean, min, max)
<i>buried</i>	relative accessible surface area
<i>helix</i>	Δ helix tendency
<i>helix</i>	Δ hydrophobicity
<i>helix</i>	Δ volume
<i>helix</i>	SIFT score
<i>helix</i>	volume (mean, min, max)
<i>helix</i>	steric parameter (mean, min, max)
<i>helix</i>	secondary structure probabilities
<i>helix</i>	disorder probability
<i>sheet</i>	SIFT score
<i>sheet</i>	Δ sheet tendency
<i>sheet</i>	isoelectric point (mean, min, max)
<i>sheet</i>	Δ volume
<i>sheet</i>	Δ flexibility
<i>sheet</i>	Δ bulkiness
<i>sheet</i>	ionisation (mean, min, max)
<i>sheet</i>	relative accessible surface area
<i>coil</i>	hydrophobicity (mean, min, max)
<i>coil</i>	Δ bulkiness
<i>coil</i>	relative accessible surface area
<i>coil</i>	SIFT score
<i>coil</i>	Δ polarisability
<i>coil</i>	Δ steric parameter
<i>coil</i>	Δ sheet tendency
<i>coil</i>	Δ hydrophobicity

* Δ refers to the change between the introduced and deleted amino acids; (mean, min, max) was calculated for a window of six neighbouring residues