

Descriptor groups	Descriptor	Number
Amino acid composition	Amino acid composition	20
	Dipeptide composition	400
	Tripeptide composition	8000
	Normalized Moreau-Broto	240 <sup>a</sup>
Autocorrelation	Moran	240 <sup>a</sup>
	Geary	240 <sup>a</sup>
	Composition	21
	Transition	21
CTD	Distribution	105
	Conjoint Triad	343
Quasi-sequence-order	Sequence-order-coupling number	60 <sup>a</sup>
	Quasi-sequence-order descriptors	100 <sup>a</sup>
Pseudo-amino acid composition	Type I	50 <sup>a</sup>
	Type II	80 <sup>a</sup>
Proteochemometric descriptors	Principal components analysis (amino acid properties based)	175 <sup>b</sup>
	Principal components analysis (2D and 3D molecular descriptors based)	4025 <sup>b</sup>
	Factor analysis (amino acid properties based)	175 <sup>b</sup>
	Factor analysis (2D and 3D molecular descriptors based)	4025 <sup>b</sup>
	Multidimensional scaling (amino acid properties based)	175 <sup>b</sup>
	Multidimensional scaling (2D and 3D molecular descriptors based)	4025 <sup>b</sup>
	BLOSUM and PAM matrix-derived descriptors	175 <sup>b</sup>
PSSM	PSSM profile	—

<sup>a</sup>The number of descriptor values depends on the choice of the number of properties of amino acid and the choice of the parameter

<sup>b</sup>The number of descriptor values depends on the choice of the number of components and the choice of the lag parameter