

Table 1: Parameter settings for indel-Seq-Gen, ROSE, and SIMPROT used in this study.

Set	Parameter	indel-Seq-Gen	ROSE	SIMPROT
GPCR	Partitions ^a	EC ₁ . . . EC ₄ , four extracellular partitions TM ₁ . . . TM ₇ , seven transmembrane partitions CY ₁ . . . CY ₄ , four cytosolic partitions	Single Sequence, I+ γ = 1.0, <i>except for</i> : TM region I+ γ = 0.7 Conserved region I+ γ = 0.0	15 partitions: corresponds with iSG partitions
	Amino acid frequencies	gpcr-cy, gpcr-tm, and gpcr-ec	gpcr-all	JTT frequencies
	Substitution matrix	JTT	PAM	JTT
	Guide tree ^b – <i>Topology</i> – <i>Branch lengths</i>	Same as template, for all partitions Calculated for each partition	Same as template Same as template	Same as template, for all partitions Calculated for each partition
	Indels: – <i>Maximum size</i>	EC ₁ : 8, EC ₂ : 2, EC ₃ : 6, EC ₄ : 6 TM ₁ . . . TM ₇ : 2 CY ₁ : 2, CY ₂ : 2, CY ₃ : 6, CY ₄ : 8	4 corresponding partition	Same as iSG for each
	– <i>Frequency</i> (<i>per substitution</i>)	TM ₁ . . . TM ₇ = 0.01 CY ₄ , EC ₁ = 0.04, EC ₄ = 0.03 CY ₁ . . . CY ₃ , EC ₂ , EC ₃ = 0.02	0.00065 ^c	Same as iSG for each corresponding partition
	– <i>Length distribution</i>	Zipfian	Zipfian	Zipfian
Lipocalin	Partitions	B ₁ . . . B ₄ , four beta partitions B ₅₆₇₈ , a region of four betas N-term, C-term, terminal regions C, concatenated coil regions.	Single sequence, I+ γ = 1, <i>except for</i> : Beta-strands I+ γ = 0.9 Terminal regions I+ γ = 1.2	Same as iSG for each corresponding partition
	Amino acid frequencies	beta, alpha, and JTT frequencies	lipo-all	JTT frequencies
	Substitution matrix	JTT	PAM	JTT
	Guide tree ^b – <i>Topology</i> – <i>Branch lengths</i>	Same as template, for all partitions Calculated for each partition, <i>except for</i> : Regions B ₅₆₇₈ and C	Same as template Same as template	Same as template, for all partitions Calculated for each partition
	Indels: – <i>Maximum size</i>	B ₁ . . . B ₄ : 0, B ₅₆₇₈ : 1 C: 2, N-term, C-term: 4	4	Same as iSG for each corresponding partition
	– <i>Frequency</i> (<i>per substitution,</i> <i>insertion/deletion</i>)	B ₁ . . . B ₇ = 0.0, B ₅₆₇₈ = 0.01/0.0 Nterm, Cterm = 0.04 C = 0.02/0.01	0.00090/0.00045	B ₁ . . . B ₇ = 0.0, B ₅₆₇₈ = 0.01/0.01 Nterm, Cterm = 0.04 C = 0.02
	– <i>Length distribution</i>	Zipfian	Zipfian	Zipfian

^a For ROSE, each partition was specified by assigning different values in the I+ γ array as shown. Information regarding the conserved regions were taken from Fuchs et. al. (2001).

^b The guide tree was obtained from the template alignment as shown in figure 1.

^c The indel frequency was set so that the hidden Markov model generated by ROSE was reasonably diverged from the template sequence, i.e. the Viterbi scores shown in Supplementary figure 2 were similar to those generated by indel-Seq-Gen.