PSORT II Variables

Last updated: $Date: 2006/08/29 00:53:33 $

Name	Description
psg	Score for the presence of signal peptides; the algorithm was basically developed by McGeoch and was expanded by Nakai and Kanehisa in 1992 (GIW96)
gvh	Original weight-matrix score of von Heijne for the recognition of cleavage sites of eukaryotic signal peptides subtracted by 7.5
alm	Discriminant score, calculated from the most hydrophobic 17 residue segment, given by the algorithm of Klein, Kanehisa, DeLisi
top	Score for predicting the topology of membrane proteins, by Hartmann, Rapoport, Lodish's method. It is essentially the net charge difference of 15 residues flanking the most N-terminal transmembrane segment on both sides
tms	Predicted number of transmembrane segments (except the cleavable signal peptide), given by Klein, Kanehisa, DeLisi's algorithm
mit	Score for the presence of N-terminal mitochondrial targeting signal, which is calculated from the amino acid composition of the N-terminal 20 residues (developed by Nakai and Kanehisa)
mip	Predicted position of the cleavage sites of mitochodrial targetign signals by Gavel and von Heijne's method
nuc	Discriminant score for being a nuclear protein, calculated from the presence of NLS motif, bipartite motif, and the amino acid composition, by Nakai and Kanehisa. NLSs are detected by the two rules: 4 residue pattern (called 'pat4') composed of 4 basic amino acids (K or R), or composed of three basic amino acids (K or R) and either H or P; the other (called 'pat7') is a pattern starting with P and followed within 3 residues by a basic segment containing 3 K/R residues out of 4. The bipartite pattern (called 'bipartite') is: 2 basic residues, 10 residue spacer, and another basic region consisting of at least 3 basic residues out of 5 residues.
erl	"1" if the "(K/H)DEL" motif exists on its C-terminus; "0" otherwise(for the prediction of ER luminal proteins)
erm	Score of being an ER membrane protein, calculated from the presence of the retention signals and the membrane topology (PSORT II original)
pax	Prediction of peroxisomal proteins from the presence of "SKL"-like motifs (PTS1) mainly on the C-terminus
px2	Prediction of peroxisomal proteins from the presence of the "(R/K)(L/I)xxxxx(H/Q)L" motif (see this review)

vac	Presence of a possible/unreliable vacuolar targeting motif "(T/I/K)LP(L/K/I)"
rnp	Presence of a motif of RNA-binding proteins "[RK]G[^EDRKHPCG][AGSCI][FY][LIVA].[FYM]" (taken from PROSITE)
act	Presence of 2 motifs for actinin-type actin-binding (taken from PROSITE) in the hope to detect some cytoskeletal proteins
caa	Presence of the CaaX motif ("C[^DENQ][LIVM].$") on the C-teminus for predicting isoprenylated proteins (review)
yqr	Presence of the "YQRL" motif in membrane proteins; this motif is used for the transport from the cell surface to Golgi
tyr	An arbitrary score calculated from the presence of the tyrosine-containing motif in the cytoplasmic tail of membrane proteins, which can be important for the selective inclusion in clathrin-coated vesicles (endocytosis) and lysosomal targeting
leu	An arbitrary score calculated from the presence of the dileucine motif "LL" in the cytoplasmic tail of membrane proteins, which can be important for the selective inclusion in clathrin-coated vesicles (endocytosis) and lysosomal targeting
gpi	Prediction of GPI-anchored proteins based on empirical knowledge that most of them are the type Ia membrane proteins with very short cytoplasmic tail (within 10 residues) (Nakai and Kanehisa)
myr	Presence of some features around the N-teminus for predicting N-myristoylated/palmitylated proteins
dna	Presence of 63 PROSITE motifs related to DNA binding, which may be useful to distinguish nuclear proteins
rib	Presence of 71 motifs for ribosomal proteins obtained from PROSITE, which may be necessary for prediction because the sorting process of ribosomal proteins are quite complex
bac	Presence of 33 prokaryotic DNA binding motifs from PROSITE, which might be useful for the prediction of even eukaryotic nuclear proteins
m1a	"1" if it is predicted to be a membrane protein with the type 1a topology (having a cleavable signal sequence and one transmembrane segment); "0", otherwise; prediction is the combination of variaous methods (psg, gvh, alm, and top)
m1b	"1" if it is predicted to be a membrane protein with the type 1b topology; "0", otherwise; prediction is the combination of variaous methods (psg, gvh, alm, and top)
m2	"1" if it is predicted to be a membrane protein with the type 2 topology; "0", otherwise; prediction is the combination of variaous methods (psg, gvh, alm, and top)
mNt	"1" if it is predicted to be a membrane protein with the N-tail topology (having an

	uncleavable signal peptide and one transmembrane segment near its C-terminus); "0", otherwise; prediction is the combination of variaous methods (psg, gvh, alm, and top)
m3a	"1" if it is predicted to be a membrane protein with the type 3a topology (multiple transmembrane regions with its N-teminus facing the cytosolic? side); "0", otherwise; prediction is the combination of variaous methods (psg, gvh, alm, and top)
m3b	"1" if it is predicted to be a membrane protein with the type 3b topology (multiple transmembrane regions with its N-terminus facing the extra-cytosolic? side); "0", otherwise; prediction is the combination of variaous methods (psg, gvh, alm, and top)
m_	"1" if it is predicted to be a membrane protein with any exceptional topology; "0", otherwise; prediction is the combination of variaous methods (psg, gvh, alm, and top)
ncn	Score (called "NNCN") for discriminating the tendency to be at either the nucleus or the cytoplasm, calculated based on the amino acid composition with the neural network constructed by Reinhardt and Hubbard
lps	Predicted number of residues in the coiled-coil structure; prediction method is according to Lupas et al.
len	Length of the input sequence