

Supplement: Learning Peptide Recognition Rules 1
for a Low-Specificity Protein 2

Lucas C. Wheeler^{1,2,3}, Arden Perkins^{1,2}, Caitlyn E. Wong^{1,2} 3
and Michael J. Harms^{1,2*}

1. Institute of Molecular Biology, University of Oregon, Eugene OR 97403 4

2. Department of Chemistry and Biochemistry, University of Oregon, Eugene 5
OR 97403 6

3. Department of Ecology and Evolutionary Biology, University of Colorado, 7
Boulder CO 80309 8

Table S1: Features used for supervised machine learning. Features denoted (CIDER) were calculated using the CIDER software package (1). Other features were calculated using our own software package (HOPS: <https://github.com/harmslab/hops>). Features were organized into the combined feature group categories indicated in the second column.

feature	Training feature group	Pooled feature group	ref
κ	CIDER	polar	(1)
Δ	CIDER	polar	(1)
Ω	CIDER	polar	(1)
FER	CIDER	polar	(1)
Σ	CIDER	polar	(1)
dmax	CIDER	polar	(1)
Δ max	CIDER	polar	(1)
NCPR	CIDER	polar	(1)
F_+	CIDER	polar	(1)
F_-	CIDER	polar	(1)
FCR	CIDER	polar	(1)
mean hydrophathy	CIDER	hydrophobic	(1)
White Interface scale	HOPS	hydrophobic	(2)
Engleman scale	HOPS	hydrophobic	(3)
% buried in structures	HOPS	hydrophobic	(4)
Kyte/Doolittle scale	HOPS	hydrophobic	(5)
Octanol scale	HOPS	hydrophobic	(6)
Hopp-Woods scale	HOPS	hydrophobic	(7)
Uversky scale	HOPS	hydrophobic	(1)
cumulative mean hydrophathy	HOPS	hydrophobic	(1)
side chain accessible area	HOPS	hydrophobic	(8)
main chain accessible area	HOPS	hydrophobic	(8)
Chou-Fasman, β	HOPS	sec. structure	(9)
Chou-Fasman, α	HOPS	sec. structure	(9)
Chou-Fasman, turn	HOPS	sec. structure	(9)
fraction poly-proline II	HOPS	sec. structure	(1)
num. hbond acceptors	HOPS	polar	—
num. hbond donors	HOPS	polar	—

predicted charge at pH 4	HOPS	polar	—
predicted charge at pH 5	HOPS	polar	—
predicted charge at pH 6	HOPS	polar	—
predicted charge at pH 7	HOPS	polar	—
predicted charge at pH 8	HOPS	polar	—
predicted charge at pH 9	HOPS	polar	—
num. positive amino acids	HOPS	polar	—
num. neutral amino acids	HOPS	polar	—
num. negative amino acids	HOPS	polar	—
net charge	HOPS	polar	—
isoelectric point	HOPS	polar	—
knob main chain, b	HOPS	geometry	(10)
socket main chain, x	HOPS	geometry	(10)
socket main chain, y	HOPS	geometry	(10)
socket main chain, h	HOPS	geometry	(10)
knob side chain, b	HOPS	geometry	(10)
socket side chain, x	HOPS	geometry	(10)
socket side chain, y	HOPS	geometry	(10)
socket side chain, h	HOPS	geometry	(10)
side chain volume	HOPS	geometry	(11)
molecular weight	HOPS	geometry	—
aromatic	HOPS	geometry	—

References

- [1] Holehouse AS, Ahad J, Das RK, Pappu RV. CIDER: Classification of Intrinsically Disordered Ensemble Regions. *Biophysical Journal*. 2015 Jan;108(2):228a.
- [2] Hessa T, Kim H, Bihlmaier K, Lundin C, others. Recognition of Transmembrane Helices by the Endoplasmic Reticulum Translocon. *Nature*. 2005;433(7024):377.
- [3] Engelman DM, Steitz TA, Goldman, A . Identifying Nonpolar Transbilayer Helices in Amino Acid Sequences of Membrane Proteins. *Annual Review of Biophysics and Biophysical Chemistry*. 1986;15(1):321–353.
- [4] Schein CH. Solubility as a Function of Protein Structure and Solvent Components. *Nature Biotechnology*. 1990 Apr;8(4):308–317.
- [5] Kyte J, Doolittle RF. A Simple Method for Displaying the Hydrophobic Character of a Protein. *Journal of Molecular Biology*. 1982 May;157(1):105–132.
- [6] Wimley WC, White SH. Experimentally Determined Hydrophobicity Scale for Proteins at Membrane Interfaces. *Nature Structural & Molecular Biology*. 1996 Oct;3(10):842–848.
- [7] Hopp TP, Woods KR. Prediction of Protein Antigenic Determinants from Amino Acid Sequences. *Proceedings of the National Academy of Sciences of the United States of America*. 1981 Jun;78(6):3824–3828.
- [8] Hubbard SJ, Thornton JM. Naccess. Computer Program, Department of Biochemistry and Molecular Biology, University College London. 1993;2(1).

- [9] Chou PY, Fasman GD. Empirical Predictions of Protein Conformation. Annual Review of Biochemistry. 1978;47:251–276. 35 36
- [10] Joo H, Tsai J. An Amino Acid Code for β -Sheet Packing Structure. Proteins. 2014 Sep;82(9):2128–2140. 37 38
- [11] Richards FM. Areas, Volumes, Packing and Protein Structure. Annual Review of Biophysics and Bioengineering. 1977;6:151–176. 39 40