

FILE S1**Supporting Data**

File S1 is available for download as a compressed folder at <http://www.genetics.org/cgi/content/full/genetics.111.128025/DC1>.

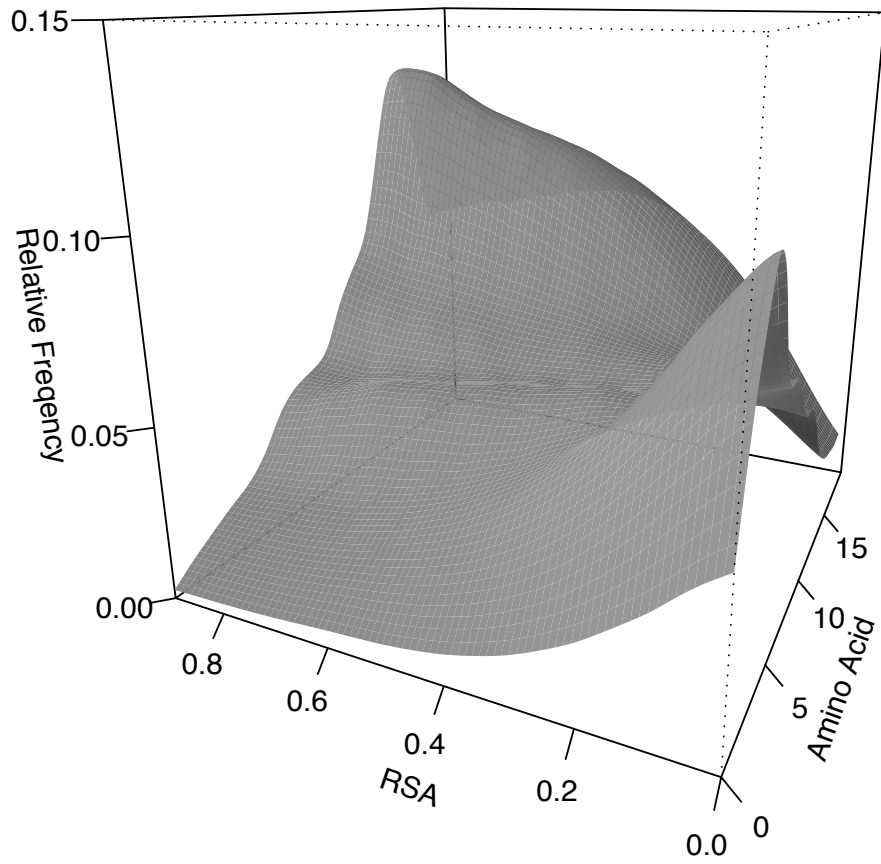
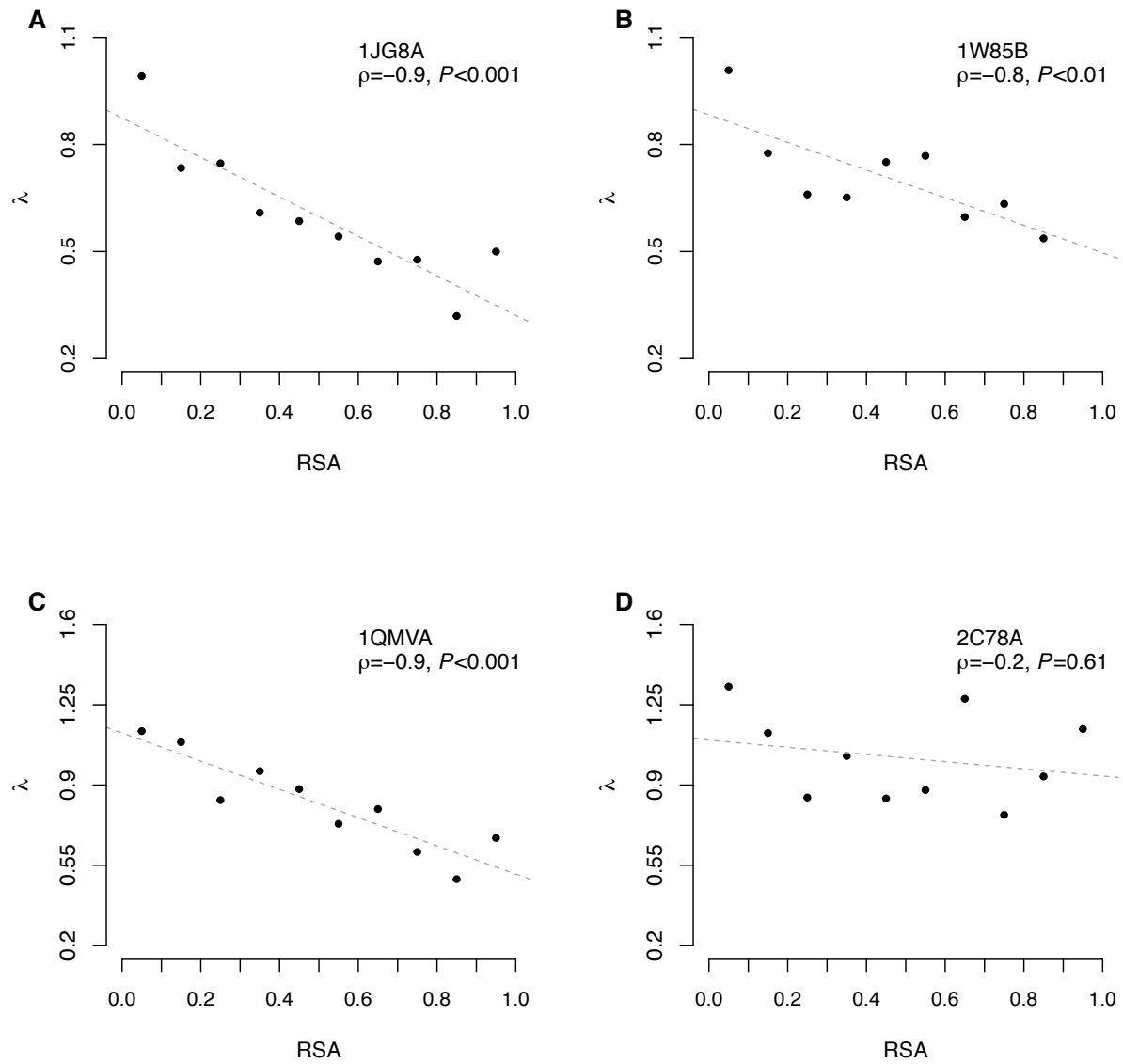


FIGURE S1.—Relative frequencies of amino acids across 525 yeast proteins, binned by RSA and ordered by decreasing hydrophobicity. Cysteine and methionine are omitted. Due to their specialized function, their average frequencies across all bins were 1.5% and 0.6%, respectively.

FIGURE S2.—The exponential parameter λ as a function of RSA, for four different protein structures.

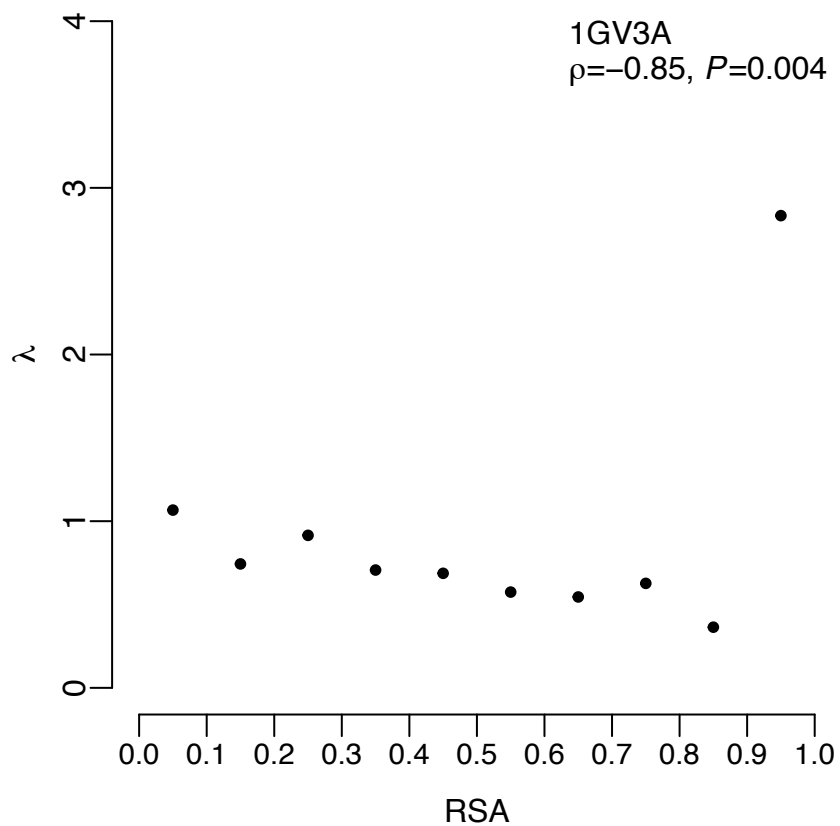


FIGURE S3.—The protein manganese superoxide dismutase (PDF identifier 1GV3, chain A) shows a clear linear decrease of λ with RSA, except for the highest RSA bin. Even though a non-parametric correlation analysis shows a strong negative correlation, a linear model (dashed line) infers a positive slope because of the one outlying data point.

TABLE S1**Fitted constants c_1 and c_2 for protein structures highlighted in Figure 5.**

PDB id	c_1	c_2
1GV3A	0.53	0.68
1JG8A	0.90	-0.55
1QMVA	1.16	-0.62
1S4OA	0.90	-0.57
1W85B	0.90	-0.39
2C78A	1.10	-0.16
2GLFA	0.87	-0.37