

## SUPPLEMENTARY INFORMATION

Figure S1. Single-residue scans. (A) Experimental (blue) and predicted (Eqn. (1) in main text; red) single-residue scans. (B) Experimental (blue) and predicted (red) symmetric pair-scans. See Table S1 for sequences.

Figure S2. Correlation between  $\Delta G_{app}^{pred}$  and  $\Delta G_{app}$  values. Measured  $\Delta G_{app}$  values for 357 designed H-segments (black) and 22 natural 19-residue sequences from multi- and single-spanning membrane proteins (red, blue) and soluble proteins (green; see Table S1) plotted against cross-validated  $\Delta G_{app}^{pred}$  values obtained using Eqn. (1). Only the 324 designed H-segments with measured  $\Delta G_{app}$  values between -1.5 and +1.0 kcal/mol were used in the  $\Delta G_{app}^{aa}$  optimization. The line  $\Delta G_{app} = \Delta G_{app}^{pred}$  and two lines encompassing 90% of the designed H-segments in the training set ( $\pm 0.45$  kcal/mol, dashed) are shown for reference. The estimated uncertainty in the measured  $\Delta G_{app}$  values for the training set is  $\pm 0.2$  kcal/mol (see Ref. 5 in the main text).

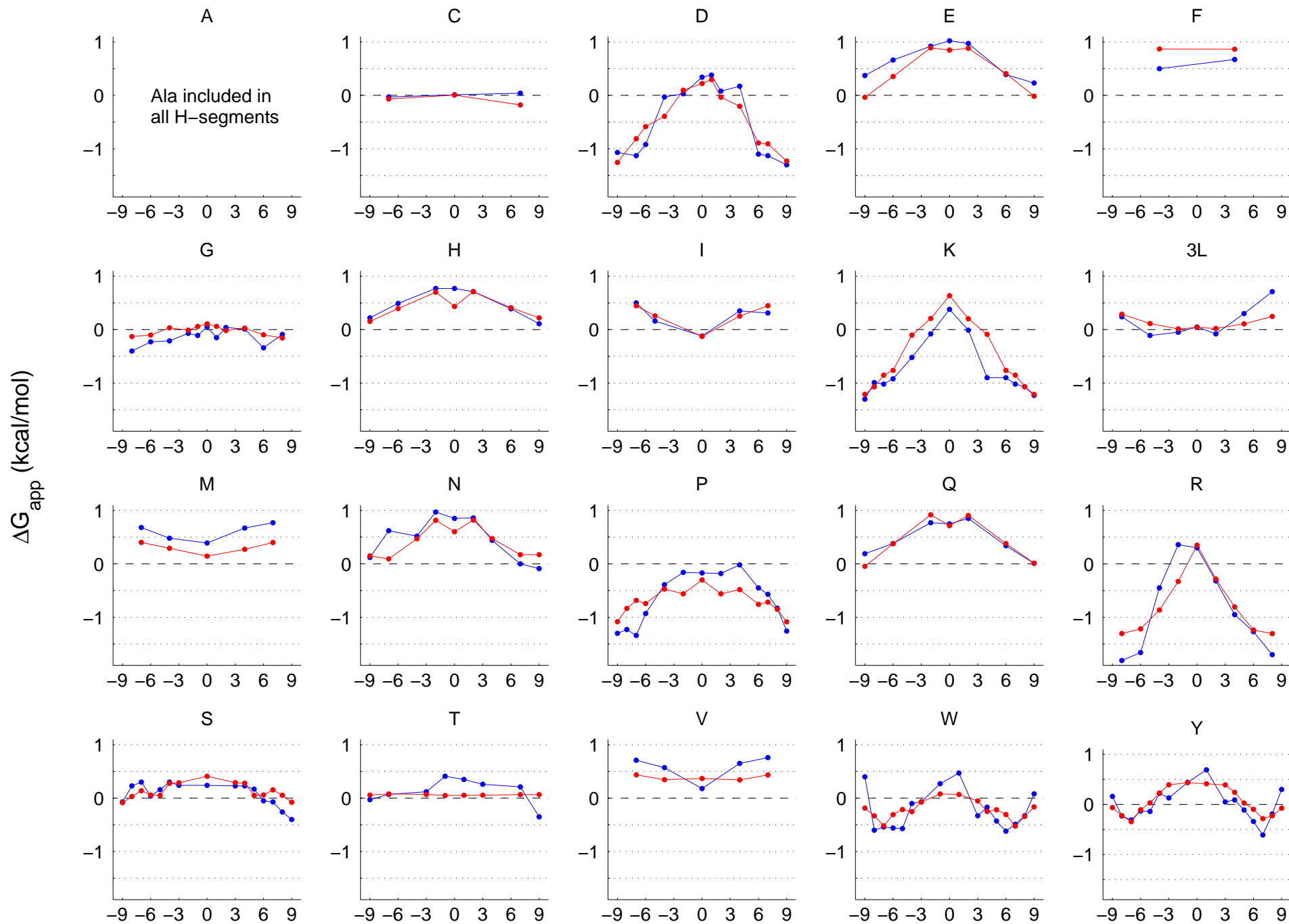
Figure S3. Length dependency of  $\Delta G_{app}$ . The slope of the lines in Fig. 3 near  $\Delta G_{app} = 0$  kcal/mol is plotted as a function of the total length ( $l = m+n$ ) at  $\Delta G_{app} = 0$  kcal/mol.

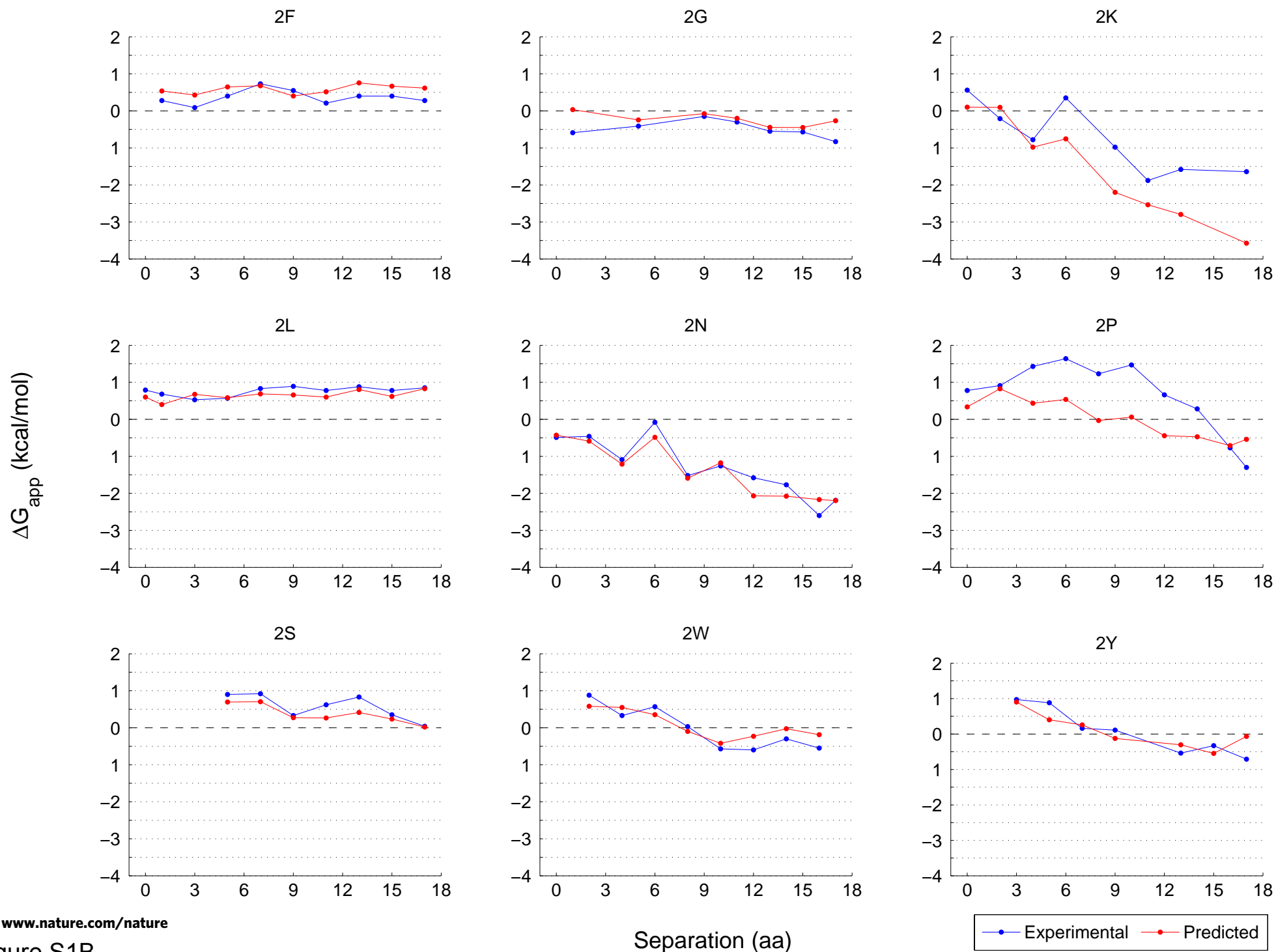
Figure S4. Single-residue and pair scans in H-segments of different lengths. (A) Single-residue Lys scan in a 15-residue (blue), 19-residue (red), and 25-residue (green) H-segment with GGPG...GPGG flanks. (B) Symmetric pair-scan with two Leu residues in a 19-residue (red) and 25-residue H-segment (blue). See Table S1 for sequences.

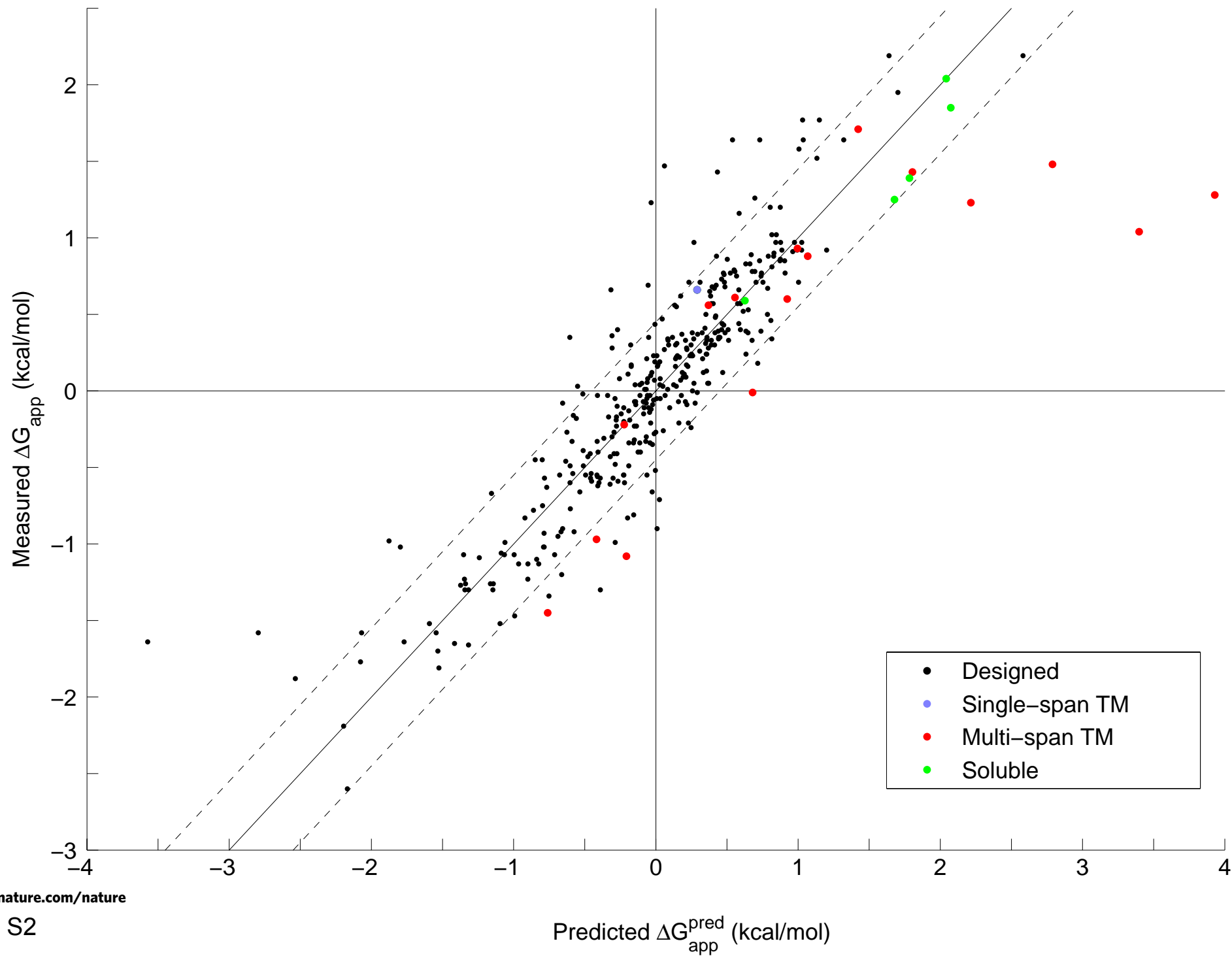
Figure S5. Contribution to  $\Delta G_{app}$  from flanking residues. Free-energy differences ( $\Delta\Delta G_{app}$ ) for different combinations of flanking residues relative to GGPG...GPGG flanks. (A) Upward-pointing triangles: 3L/16A (AAAALAAAALAAAALAAA) H-segments; downward-pointing triangles: 4L/15A (AAAALALAAAALALAAA) H-segments; blue bars: mean values.

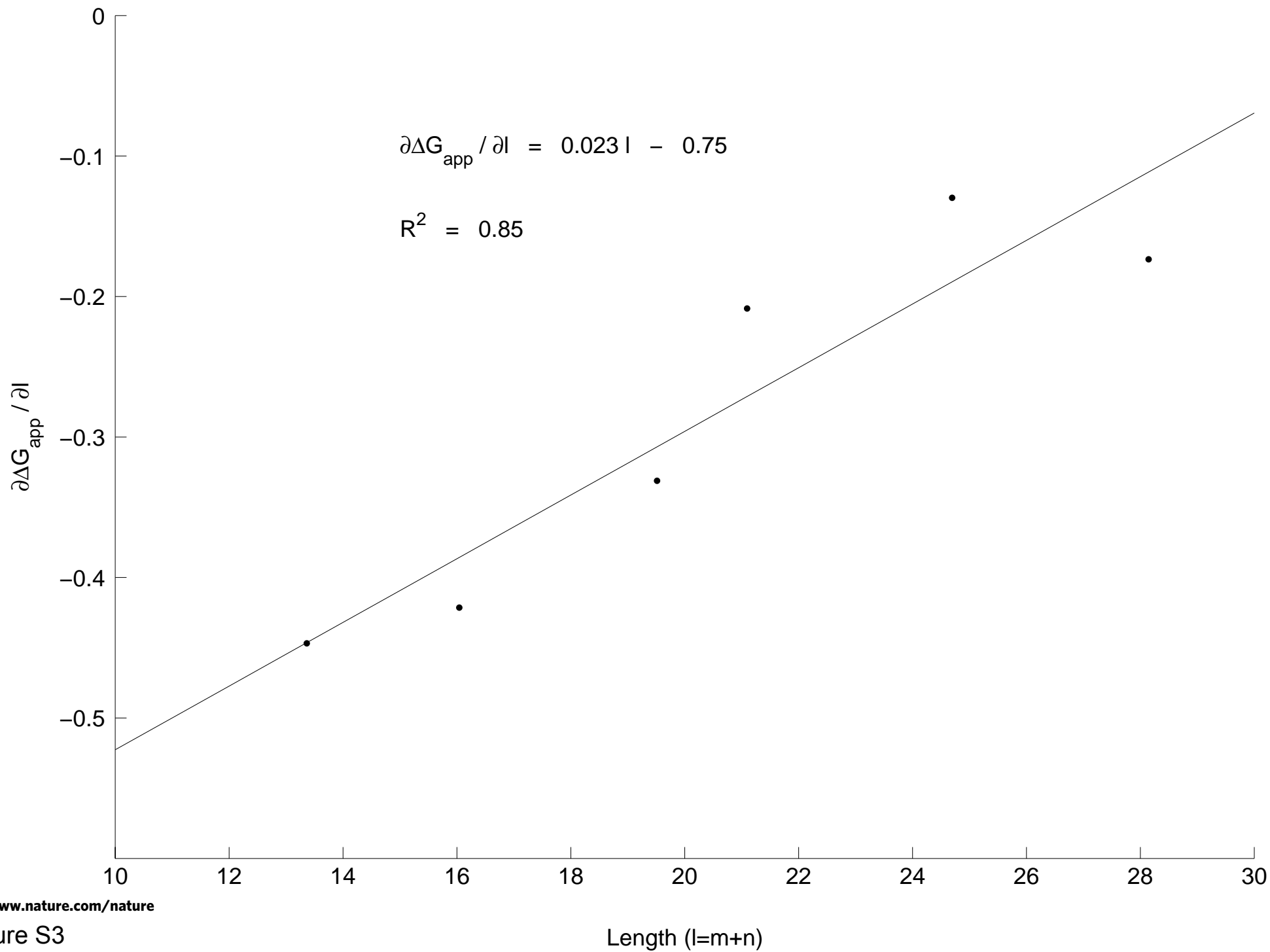
(B) Upward-pointing triangles: 10L (LLLLLLLLLLL) H-segments; downward-pointing triangles: 2L/23A (AAAAAAAAAALAAAAALAAAAA) H-segments; blue bars: mean values.

Figure S6. Comparison of different hydrophobicity scales. Distributions of  $\Delta G_{app}^{pred}$  for the most hydrophobic stretch in 1012 mammalian secreted proteins (black), 349 mammalian single-spanning membrane proteins (blue) and 508 TM helices from multispanning membrane proteins (red) using different hydrophobicity scales. The “full biological scale” includes the position-specific contributions  $\overline{\Delta G_{app}^{aa}}$  as well as the terms for length-dependence and hydrophobic moment described in Methods, whereas the “simple biological scale” is based only on the mean  $\Delta G_{app}^{aa}$  value for each residue and does not take length or hydrophobic moment into account. The number of data points in the overlap region between the secreted and single-spanning membrane proteins (SS overlap) and between the secreted and multi-spanning membrane proteins (MS overlap) is given in each panel. See Methods for details.

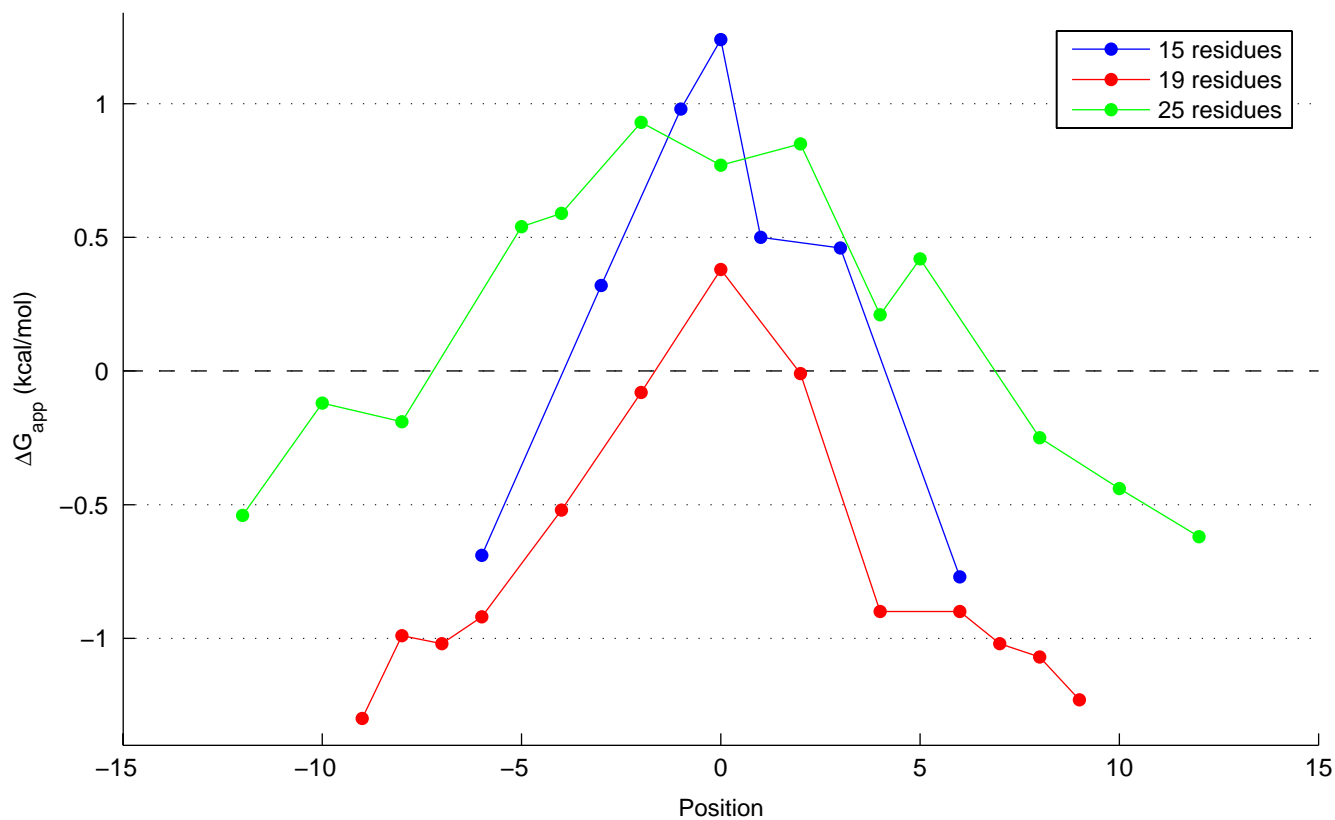




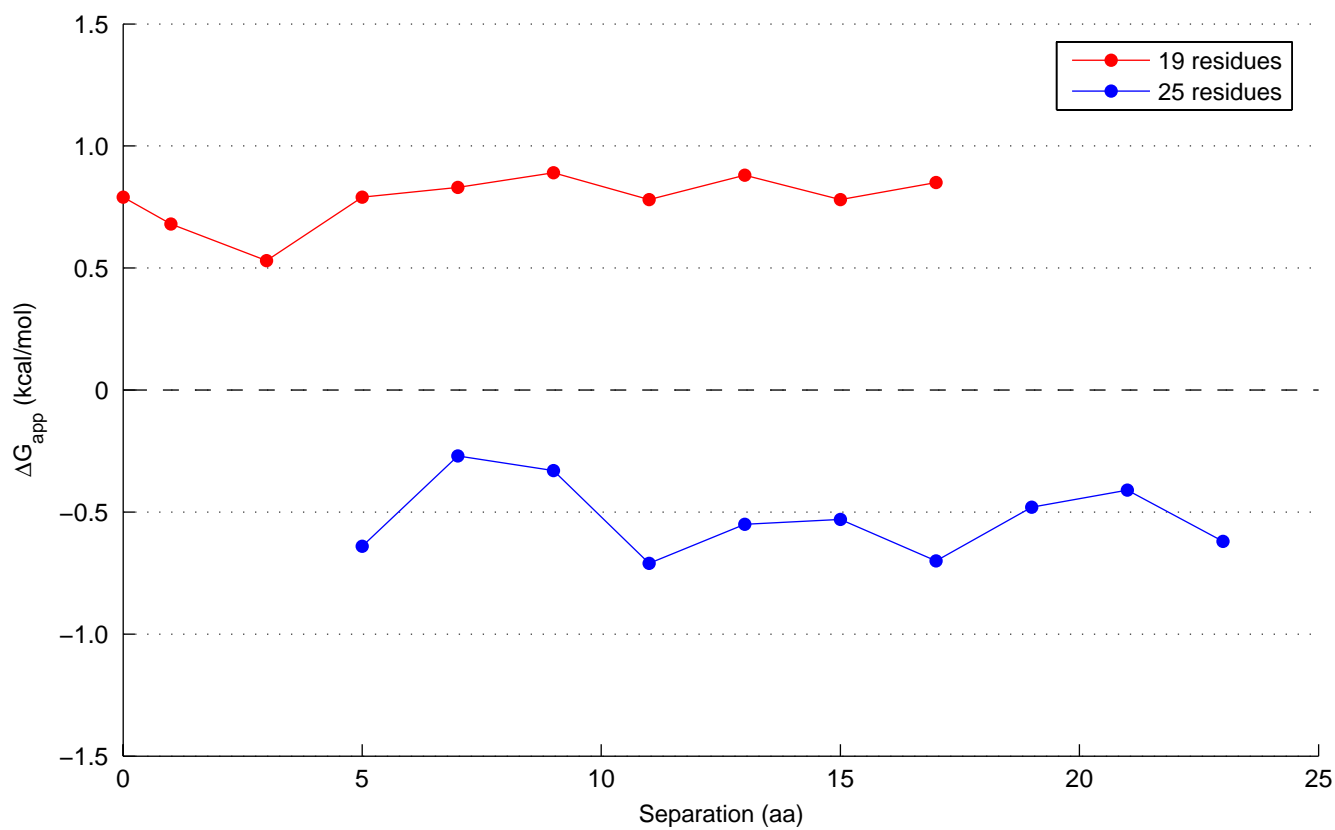


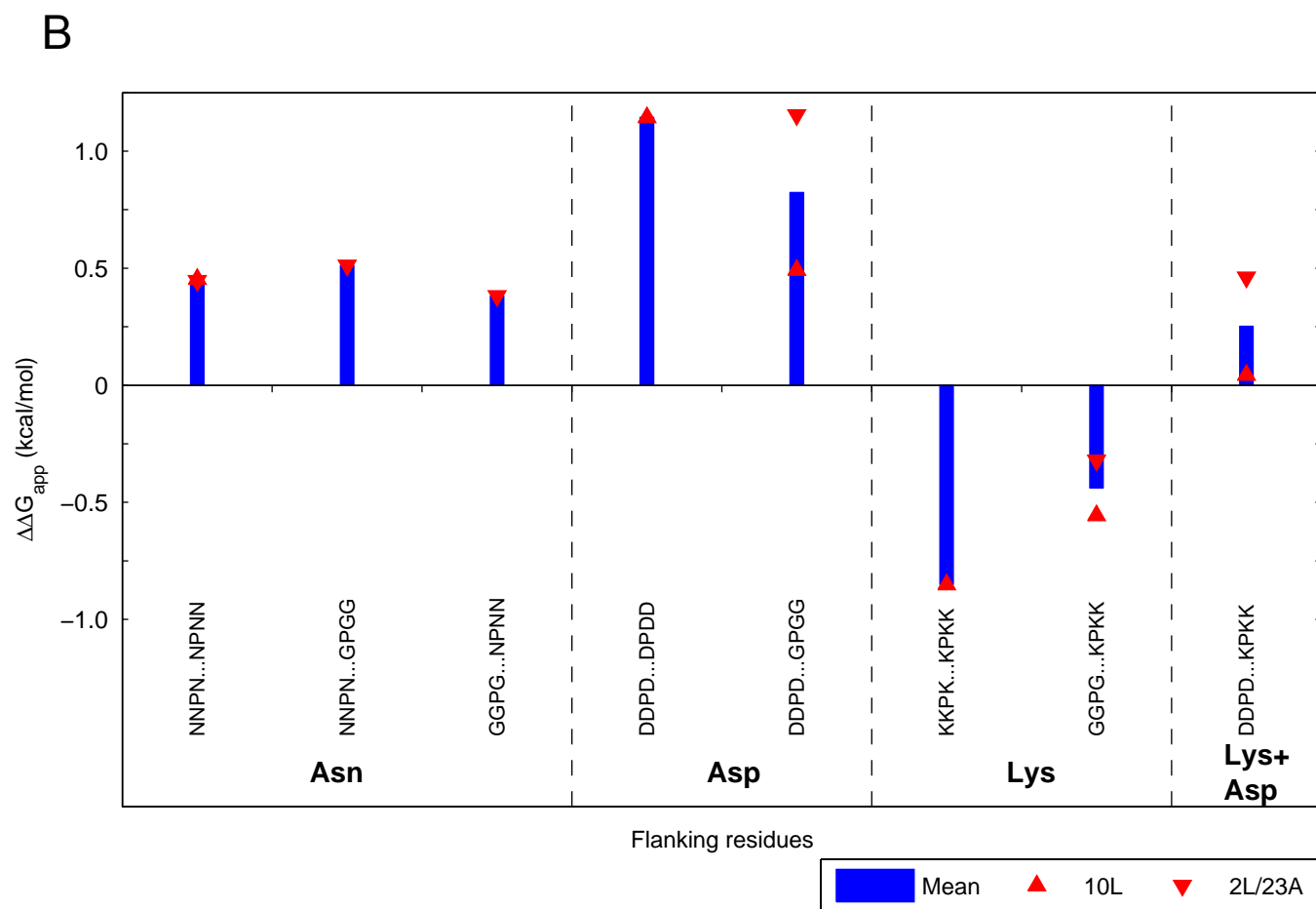
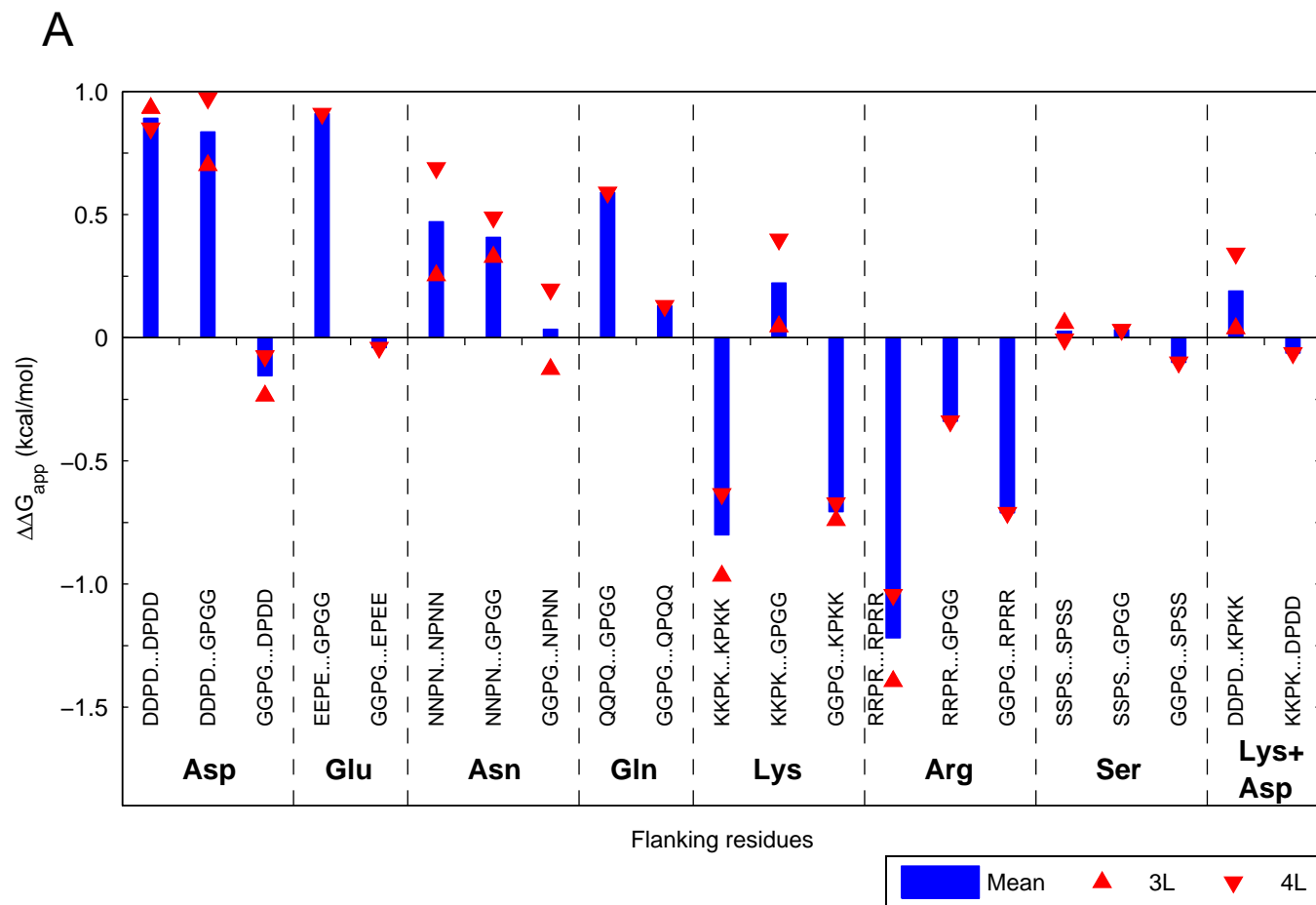


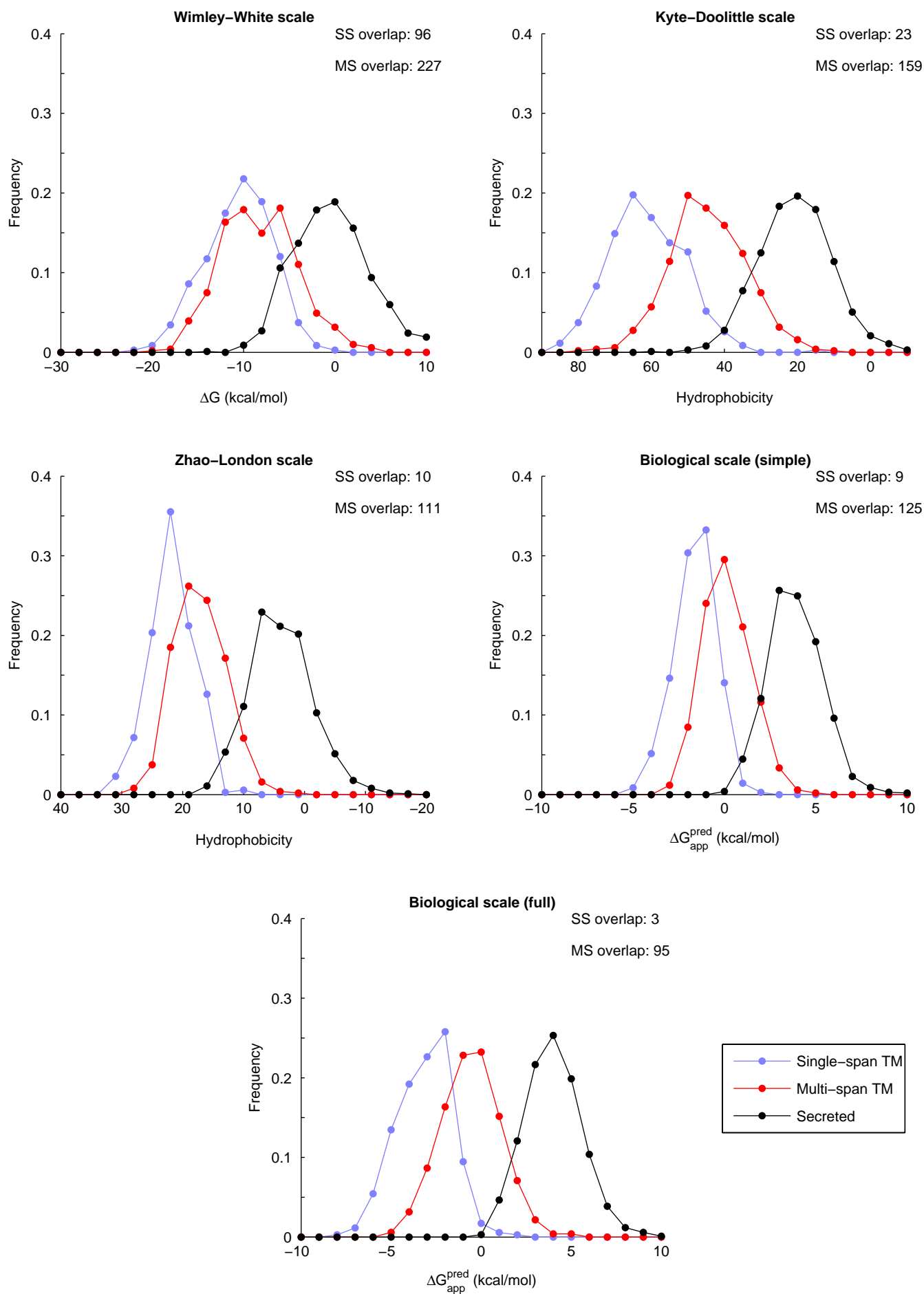
A



B







**Supplementary table S1:**  
**Measured and predicted  $\Delta G$  values (kcal/mol)**  
**for all sequences in the study**

**H-segments with G-flanks (length 19)**

<b>L</b>			
<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{Pred}$
1	GGPGLLLLLLLLLLLLLLLLLLGGPGG	0.24	0.29
2	GGPGAAALLLLLLLLLLLLLLGGPGG	-0.11	0.12
3	GGPGAAAAALLLLLLLLLLGGPGG	-0.05	0.01
4	GGPGAAAAAALLLLLLGGPGG	0.05	0.04
5	GGPGAAAAAALLLLLLGGPGG	-0.08	0.02
6	GGPGAAAAAALLLLLLGGPGG	0.30	0.11
7	GGPGAAAAAALLLLLLGGPGG	0.71	0.25
8	GGPGLAALALAAAAAALGGPGG	0.05	0.39
9	GGPGAAAAAALALALGGPGG	0.33	0.38
10	GGPGAAAAAALALALGGPGG	0.08	0.04
11	GGPGAAAAAALALALGGPGG	0.31	0.29
12	GGPGAAAAAALALALGGPGG	0.38	0.38
13	GGPGAAAAAALALALGGPGG	-0.06	0.03
14	GGPGAAALAAAAALAAAGPGG	0.19	0.06
15	GGPGAALAAAAALAAAGPGG	0.35	0.52
16	GGPGALAAAAALAAAGPGG	0.37	0.33
17	GGPGLAAAAALAAAGPGG	0.33	0.19
18	GGPGAAAAAALAAAAALGGPGG	0.79	0.60
19	GGPGAAAAAALALAAAAALGGPGG	0.68	0.40
20	GGPGAAAAAALAAAAALGGPGG	0.53	0.68
21	GGPGAAAAAALAAAAALGGPGG	0.57	0.58
22	GGPGAAAAAALAAAAALGGPGG	0.83	0.69
23	GGPGAAAAAALAAAAALGGPGG	0.89	0.66
24	GGPGAAALAAAAAALAAAGPGG	0.78	0.60
25	GGPGAALAAAAAALAAAGPGG	0.88	0.81
26	GGPGALAAAAAALAAAGPGG	0.78	0.62
27	GGPGLAAAAAALAAAGPGG	0.85	0.83
28	GGPGAAAAALALALAAAAALGGPGG	-0.40	-0.15
29	GGPGAAAAALALALAAAAALGGPGG	-0.60	-0.40
30	GGPGAAAAALALALAAAAALGGPGG	-0.66	-0.39
31	GGPGAAAAALALALAAAAALGGPGG	-0.55	-0.03
32	GGPGAAAAALALALAAAAALGGPGG	-0.55	-0.40
33	GGPGAAALALALAAAAALAAAGPGG	-0.59	-0.38
34	GGPGAALALALAAAAALAAAGPGG	-0.34	0.09
35	GGPGALALALAAAAALAAAGPGG	-0.49	-0.09
36	GGPGLALALAAAAALAAAGPGG	-0.05	-0.23
37	GGPGLAAAAAALAAAGPGG	0.81	0.73
38	GGPGALAAAAAALAAAGPGG	1.02	0.82
39	GGPGAAALALALALAAAAALGGPGG	-1.07	-0.80
40	GGPGAAAAAALAAAAALGGPGG	0.75	0.60
41	GGPGAAALALALALAAAAALGGPGG	-1.58	-1.54
42	GGPGALALALALALAAAGPGG	-1.64	-1.77
43	GGPGLLLLLLLLLLLLLLGGPGG	-2.67	-7.53

**K**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
44	GGPGKAAALALALALALALALAAAAGPGG	-1.30	-1.21
45	GGPGAKAALALALALALALALAAAAGPGG	-0.99	-1.07
46	GGPGAAKALALALALALALALAAAAGPGG	-1.02	-0.85
47	GGPGAAAKLALALALALALALAAAAGPGG	-0.92	-0.76
48	GGPGAAAALKLALALALALALAAAAGPGG	-0.52	-0.10
49	GGPGAAAALALKLALALALALAAAAGPGG	-0.08	0.21
50	GGPGAAAALALALKLALALALAAAAGPGG	0.38	0.63
51	GGPGAAAALALALALKLALALAAAAGPGG	-0.01	0.20
52	GGPGAAAALALALALALKLAAAAGPGG	-0.90	-0.09
53	GGPGAAAALALALALALALKAAAAGPGG	-0.90	-0.76
54	GGPGAAAALALALALALALAKAAGPGG	-1.02	-0.85
55	GGPGAAAALALALALALALAAKAGPGG	-1.07	-1.07
56	GGPGAAAALALALALALALAAAKGPGG	-1.23	-1.21
57	GGPGKAAAAALALALALAAAAAAGPGG	-0.20	-0.29
58	GGPGAKAAAALALALALAAAAAAGPGG	0.12	-0.20
59	GGPGAAAAAKLALALALAAAAAAGPGG	0.34	0.56
60	GGPGAAAALALAAKAALALAAAAGPGG	1.95	1.70
61	GGPGAAAAALALALALAKAAAAAGPGG	0.46	0.56
62	GGPGAAAAALALALALAAAKAAGPGG	-0.09	-0.19
63	GGPGAAAAALALALALAAAAAKGPGG	-0.11	-0.29
64	GGPGKAAAAALALALAAAAAAGPGG	0.23	0.09
65	GGPGAAAAALALALAAAAAKGPGG	0.21	0.09
66	GGPGAAALLLLLKKLLLLLAAAGPGG	0.56	0.10
67	GGPGAAALLLLLKKLLLLLAAAGPGG	-0.21	0.09
68	GGPGAAALLLLLKKLLLLLAAAGPGG	-0.78	-0.98
69	GGPGAAALLLLLKKLLLLLAAAGPGG	0.35	-0.76
70	GGPGAAALKLLLLLLLLLKLAAAGPGG	-0.98	-2.20
71	GGPGAAAKLLLLLLLLLLLLKAAAGPGG	-1.88	-2.53
72	GGPGAAKALLLLLLLLLLLAKAAGPGG	-1.58	-2.80
73	GGPGKAAALLLLLLLLLLAAAKGPGG	-1.64	-3.57
74	GGPGALAALALALKLALALALAGPGG	-0.81	-0.13
75	GGPGAAAALALALKAALALAAAAGPGG	0.92	1.09
76	GGPGKAAALAAAALAAAALAAAKGPGG	0.57	0.11
77	GGPGAAALLLLLKKLLLLLAAAGPGG	1.26	0.70
78	GGPGAAALLLLLKKLLLLLAAAGPGG	-0.27	-0.28
79	GGPGAAALLLLLKKLLLLLAAAGPGG	-0.75	-0.80
80	GGPGALLLLLLLLLKKLLLLLAAAGPGG	-1.02	-1.70

**D**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
81	GGPGDLAALALALAAALALALAGPGG	-1.07	-1.26
82	GGPGALDALALALAAALALALAGPGG	-1.13	-0.81
83	GGPGALADLALALAAALALALAGPGG	-0.92	-0.59
84	GGPGALAALDLALAAALALALAGPGG	-0.03	-0.39
85	GGPGALAALALDLAAALALALAGPGG	0.03	0.09
86	GGPGALAALALALDAALALALAGPGG	0.34	0.22
87	GGPGALAALALALADALALALAGPGG	0.38	0.29
88	GGPGALAALALALAADLALALAGPGG	0.08	-0.03
89	GGPGALAALALALAAALDLALAGPGG	0.17	-0.21
90	GGPGALAALALALAAALDALAGPGG	-1.10	-0.89
91	GGPGALAALALALAAALALADLAGPGG	-1.13	-0.91
92	GGPGALAALALALAAALALALDGP	-1.30	-1.23

93	GGPGALAAALALDLALALAAAGPGG	-0.55	-0.18
94	GGPGALAAALLLDLALALAAAGPGG	-0.55	-0.51
95	GGPGAAAALALDLALALAAAAGPGG	1.16	0.59
96	GGPGAAAALALDAALALAAAAGPGG	1.52	1.13
97	GGPGAAAALALAADAALALAAAAGPGG	2.19	1.64
98	GGPGAAALDLLLLLLDLLLLAAAGPGG	0.03	-0.40
99	GGPGAAALLLLLDDLLLLLLAAAGPGG	-0.24	0.23

**S**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
100	GGPGSAAAAALLLALAAAAAAGPGG	-0.07	-0.08
101	GGPGASAAAAALLLALAAAAAAGPGG	0.23	0.03
102	GGPGAASAAAAALLLALAAAAAAGPGG	0.30	0.14
103	GGPGAAASAAAAALLLALAAAAAAGPGG	0.04	0.06
104	GGPGAAAASALLLALAAAAAAGPGG	0.16	0.05
105	GGPGAAAAASALLLALAAAAAAGPGG	0.30	0.28
106	GGPGAAAAASLLLALAAAAAAGPGG	0.24	0.29
107	GGPGAAAAAALLSLLAAAAAAGPGG	0.24	0.41
108	GGPGAAAAAALLLALSAAAAAAGPGG	0.23	0.29
109	GGPGAAAAAALLLALASAAAAAAGPGG	0.23	0.28
110	GGPGAAAAAALLLALASAAAAAAGPGG	0.17	0.05
111	GGPGAAAAAALLLALAAASAAAGPGG	-0.05	0.06
112	GGPGAAAAAALLLALAAASAAAGPGG	-0.07	0.16
113	GGPGAAAAAALLLALAAASAAAGPGG	-0.26	0.05
114	GGPGAAAAAALLLALAAASAAAGPGG	-0.40	-0.08
115	GGPGAAAAASLLLALSAAAAAAGPGG	0.90	0.70
116	GGPGAAAAASALLLALASAAAAAAGPGG	0.92	0.70
117	GGPGAAAAASALLLALASAAAAAAGPGG	0.33	0.27
118	GGPGAAASAAALLLALAAASAAAGPGG	0.62	0.26
119	GGPGAASAAALLLALAAASAAAGPGG	0.83	0.41
120	GGPGASAAALLLALAAASAAAGPGG	0.35	0.24
121	GGPGSAAALLLALAAASAAAGPGG	0.04	0.02
122	GGPGAAAAASLLLLLSAAAAAAGPGG	-0.07	0.12
123	GGPGAAAALALAASAALALAAAAGPGG	0.27	0.17

**P**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
124	GGPGPAAALALALALALALAAAAGPGG	-1.30	-1.08
125	GGPGAPAALALALALALALAAAAGPGG	-1.23	-0.83
126	GGPGAAPALALALALALALAAAAGPGG	-1.34	-0.68
127	GGPGAAAPLALALALALALAAAAGPGG	-0.93	-0.74
128	GGPGAAAALPLALALALALALAAAAGPGG	-0.39	-0.47
129	GGPGAAAALPLALALALALALAAAAGPGG	-0.16	-0.56
130	GGPGAAAALALPLALALALALAAAAGPGG	-0.17	-0.30
131	GGPGAAAALALALALPLALALAAAAGPGG	-0.18	-0.56
132	GGPGAAAALALALALALPLAAAAGPGG	-0.02	-0.48
133	GGPGAAAALALALALALALPAAAAGPGG	-0.45	-0.76
134	GGPGAAAALALALALALALAPAAGPGG	-0.57	-0.72
135	GGPGAAAALALALALALALAAPAGPGG	-0.83	-0.85
136	GGPGAAAALALALALALALAAPAGPGG	-1.26	-1.09
137	GGPGAAAALLLPPLLLAAAAAAGPGG	0.78	0.34

138	GGPGAAAAALLPPLLPLAAAAAGPGG	0.91	0.83
139	GGPGAAAAALPLLLLPLAAAAAGPGG	1.43	0.43
140	GGPGAAAAAPLLLLLPLAAAAAGPGG	1.64	0.54
141	GGPGAAAAAPLLLAALLPLAAAAAGPGG	1.23	-0.03
142	GGPGAAAPALLLAALLLAPAAAAAGPGG	1.47	0.06
143	GGPGAAPALLLAALLLAAPAAAGPGG	0.66	-0.44
144	GGPGAPAAALLLAALLLAAAPAAAGPGG	0.28	-0.47
145	GGPGPAAAAALLLAALLLAAAAAPAGPGG	-0.77	-0.71
146	GGPGPAAAAALLLAALLLAAAAAPGPGG	-1.30	-0.54
147	GGPGAAAAALLLPPLLLLLAAAAAGPGG	-0.99	-0.39
148	GGPGAAAAALLLPPLLLLLLAAAAAGPGG	-1.20	-0.67
149	GGPGAAALLLPPLLLLLLAAAAAGPGG	-1.47	-0.92
150	GGPGAAAAALLLPALLLAAAAAGPGG	-0.43	-0.40
151	GGPGALALALALPAALALALAGPGG	-0.27	-0.63
152	GGPGALALALALPLALALALAGPGG	-1.26	-1.07
153	GGPGAAAAALALAAPALALAAAAAGPGG	1.64	0.73
154	GGPGAAAAALALALPIALALAAAAAGPGG	-0.03	-0.42

**N**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
155	GGPGNAAAAALALALALAAAAAGPGG	0.12	0.15
156	GGPGANAAALALALALAAAAAGPGG	0.62	0.09
157	GGPGAAAAANLALALALAAAAAGPGG	0.52	0.47
158	GGPGAAAAANLALALALAAAAAGPGG	0.97	0.82
159	GGPGAAAAALALNLALAAAAAGPGG	0.85	0.60
160	GGPGAAAAALALALNLAAAAAGPGG	0.86	0.82
161	GGPGAAAAALALALALNAAAAAGPGG	0.44	0.47
162	GGPGAAAAALALALALANAAGPGG	0.00	0.17
163	GGPGAAAAALALALALAAAANGPGG	-0.09	0.17
164	GGPGNAAAAALALALAAAAAGPGG	0.43	0.32
165	GGPGAAAAALALALAAAAANGPGG	0.38	0.33
166	GGPGAAALLLNLLLLLAAAAAGPGG	-0.49	-0.43
167	GGPGAAALLLNLLNLLLLAAAGPGG	-0.46	-0.59
168	GGPGAAALLLNLLLLNLLLLAAAGPGG	-1.09	-1.21
169	GGPGAAALLNLLLLLNLAAAAAGPGG	-0.08	-0.49
170	GGPGAAALNLLLLLNLAAAAAGPGG	-1.52	-1.59
171	GGPGAAANLLLLLNLAAAAAGPGG	-1.26	-1.18
172	GGPGAAANLLLLLNLAAAAAGPGG	-1.58	-2.07
173	GGPGAAANLLLLLNLAAAAAGPGG	-1.77	-2.08
174	GGPGANAAALLLLLLLAAAAAGPGG	-2.60	-2.17
175	GGPGNAAALLLLLLLAAAAAGPGG	-2.19	-2.19
176	GGPGAAALALALNLALALAAAAAGPGG	-0.28	0.02
177	GGPGALALALALNAALALALAGPGG	-0.48	-0.29
178	GGPGAAALALALNAALALAAAAAGPGG	0.86	0.46
179	GGPGAAALALANAALALAAAAAGPGG	1.64	1.04
180	GGPGAAALALALNIALALAAAAAGPGG	-0.27	-0.05
181	GGPGAAALLLNLLLLLAAAAAGPGG	-0.21	0.13
182	GGPGAAALLLNLLLLLAAAAAGPGG	-0.23	-0.17

183	GGPGAAALLLLLNLLLLLAAAGPGG	-1.13	-0.85
<b>R</b>			
<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
184	GGPGARAALALALALALALAAAAGPGG	-1.81	-1.30
185	GGPGAAARLALALALALALAAAAGPGG	-1.66	-1.22
186	GGPGAAALRLALALALALALAAAAGPGG	-0.45	-0.86
187	GGPGAAALALRLALALALALAAAAGPGG	0.36	-0.33
188	GGPGAAALALALRLALALALAAAAGPGG	0.30	0.35
189	GGPGAAALALALRLALALAAAAGPGG	-0.32	-0.28
190	GGPGAAALALALALRLALAAAAGPGG	-0.95	-0.81
191	GGPGAAALALALALALRAAAAGPGG	-1.27	-1.24
192	GGPGAAALALALALALAAARAGPGG	-1.70	-1.31
193	GGPGAAALALALRIALALAAAAGPGG	0.12	0.31
<b>G</b>			
<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
194	GGPGAGAALALAAAAALALAAAAGPGG	-0.40	-0.13
195	GGPGAAAGLALAAAAALALAAAAGPGG	-0.23	-0.10
196	GGPGAAALGLAAAAALALAAAAGPGG	-0.21	0.03
197	GGPGAAALALGAAALALAAAAGPGG	-0.07	-0.01
198	GGPGAAALALAGAALALAAAAGPGG	-0.11	0.06
199	GGPGAAALALAGAALALAAAAGPGG	0.06	0.11
200	GGPGAAALALAAAGALALAAAAGPGG	-0.15	0.06
201	GGPGAAALALAAAAGLALAAAAGPGG	0.04	-0.02
202	GGPGAAALALAAAALGLAAAAGPGG	0.01	0.03
203	GGPGAAALALAAAALALGAAAGPGG	-0.34	-0.10
204	GGPGAAALALAAAALALAAAGAGPGG	-0.09	-0.16
205	GGPGAAALALGLGLALALAAAAGPGG	-0.59	0.03
206	GGPGAAALGLALALGLALAAAAGPGG	-0.41	-0.24
207	GGPGAAAGLALALALALGAAAAGPGG	-0.15	-0.08
208	GGPGAAAGALALALALALAGAAAGPGG	-0.30	-0.20
209	GGPGAAGALALALALALAGAAGPGG	-0.55	-0.45
210	GGPGAGAAALALALALALAAAGAGPGG	-0.57	-0.45
211	GGPGGAAALALALALALAAAAGGPGG	-0.83	-0.27
212	GGPGAAAAALALALALAAAAGGGPGG	0.33	0.00
213	GGPGAAAAALALALALAAAGGAGPGG	0.69	-0.08
214	GGPGAAAAGAAAAGAAAAGAAAAGPGG	2.19	2.58
215	GGPGAAALALALGLALALAAAAGPGG	-1.52	-1.10
216	GGPGAAALAAAAGAAAALAAAAGPGG	1.77	1.15
217	GGPGGAAALALALALAAAAAGPGG	0.34	0.00
218	GGPGAAALAGAALAAGALAAAAGPGG	1.20	0.80
<b>H</b>			
<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
219	GGPGHAAAAALALALALAAAAAGPGG	0.22	0.15
220	GGPGAAHAALALALALAAAAAGPGG	0.49	0.39
221	GGPGAAAAALHLALALAAAAAGPGG	0.77	0.70
222	GGPGAAAAALALHLALAAAAAGPGG	0.77	0.43
223	GGPGAAAAALALALHLAAAAAGPGG	0.71	0.71
224	GGPGAAAAALALALALAAHAAAGPGG	0.39	0.41
225	GGPGAAAAALALALALAAAAAHGPGG	0.11	0.22

226	GGPGAAAALALHLALALAAAAGPGG	-0.19	-0.10
227	GGPGAAAALALLHAALALAAAAGPGG	0.30	0.36
228	GGPGAAAALALAAHAALALAAAAGPGG	1.20	0.87
229	GGPGAAAALALHIALALAAAAGPGG	-0.34	-0.16

**E**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
230	GGPGEAAAAALALALAAAAAAGPGG	0.37	-0.04
231	GGPGAAEAALALALAAAAAAGPGG	0.66	0.35
232	GGPGAAAAALELALALAAAAAAGPGG	0.92	0.89
233	GGPGAAAAALALELALAAAAAAGPGG	1.02	0.85
234	GGPGAAAAALALALELAAAAAAGPGG	0.97	0.88
235	GGPGAAAAALALALAAEAAAAGPGG	0.39	0.41
236	GGPGAAAAALALALAAAAAEGPGG	0.23	-0.02
237	GGPGAAAALALALELALALAAAAGPGG	0.44	0.02
238	GGPGALAALALALEAALALALAGPGG	-0.66	-0.10
239	GGPGAAAALALALEIALALAAAAGPGG	0.35	-0.08

**T**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
240	GGPGTAAAAALALALAAAAAAGPGG	-0.03	0.06
241	GGPGAATAAALALALAAAAAAGPGG	0.07	0.08
242	GGPGAAAAATLALALAAAAAAGPGG	0.12	0.07
243	GGPGAAAAAALTALALAAAAAAGPGG	0.41	0.05
244	GGPGAAAAAALALTALAAAAAAGPGG	0.35	0.05
245	GGPGAAAAAALALALTAAAAAAGPGG	0.26	0.06
246	GGPGAAAAAALALALAAATAAGPGG	0.21	0.07
247	GGPGAAAAAALALALAAAAATGPGG	-0.35	0.07
248	GGPGAAAALALATAALALAAAAGPGG	-0.05	-0.40

**Q**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
249	GGPGQAAAAALALALAAAAAAGPGG	0.19	-0.05
250	GGPGAAAQALALALAAAAAAGPGG	0.38	0.38
251	GGPGAAAAALQLALALAAAAAAGPGG	0.77	0.92
252	GGPGAAAAALALQLALAAAAAAGPGG	0.75	0.72
253	GGPGAAAAALALALQLAAAAAAGPGG	0.85	0.91
254	GGPGAAAAALALALAAQAAAAGPGG	0.34	0.38
255	GGPGAAAAALALALAAAAAQPGG	0.01	0.01
256	GGPGAAAALALALQLALALAAAAGPGG	0.10	0.12
257	GGPGALAALALALQAALALALAGPGG	-0.41	-0.16
258	GGPGAAAALALAAQAALALAAAAGPGG	1.77	1.03
259	GGPGAAAALALALQIALALAAAAGPGG	-0.03	0.05
260	GGPGAAAALALQLTLALALAAAAGPGG	0.23	-0.21

**W**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
261	GGPGWAAAAALALALAAAAAAGPGG	0.40	-0.19

262	GGPGAWAAAAALALALAAAAAAGPGG	-0.60	-0.33
263	GGPGAOWAAAAALALALAAAAAAGPGG	-0.54	-0.52
264	GGPGAAAOWAAALALALAAAAAAGPGG	-0.56	-0.31
265	GGPGAAAOWAALALALAAAAAAGPGG	-0.57	-0.21
266	GGPGAAAOWALALALAAAAAAGPGG	-0.10	-0.25
267	GGPGAAAOWLALALAAAAAAGPGG	-0.07	-0.07
268	GGPGAAAOWLWALALAAAAAAGPGG	0.27	0.08
269	GGPGAAAOWLALWALAAAAAAGPGG	0.47	0.07
270	GGPGAAAOWALALWAAAAAAGPGG	-0.33	-0.05
271	GGPGAAAOWALALWAAAAAAGPGG	-0.17	-0.25
272	GGPGAAAOWALALWAAAAAAGPGG	-0.43	-0.22
273	GGPGAAAOWALALWAAAAAAGPGG	-0.62	-0.31
274	GGPGAAAOWALALWAAAAAAGPGG	-0.49	-0.52
275	GGPGAAAOWALALWAAAAAAGPGG	-0.33	-0.34
276	GGPGAAAOWALALWAAAAAAGPGG	0.08	-0.16
277	GGPGAAAOWALLWAAAAAAGPGG	0.88	0.58
278	GGPGAAAOWALLWAAAAAAGPGG	0.33	0.55
279	GGPGAAAOWALLWAAAAAAGPGG	0.57	0.35
280	GGPGAAAOWALLWAAAAAAGPGG	0.03	-0.10
281	GGPGAAAOWALLWAAAAAAGPGG	-0.57	-0.42
282	GGPGAAAOWALLWAAAAAAGPGG	-0.60	-0.23
283	GGPGAAAOWALLWAAAAAAGPGG	-0.30	-0.03
284	GGPGAAAOWALLWAAAAAAGPGG	-0.55	-0.19
285	GGPGAAAOWALLWAAAAAAGPGG	-0.03	-0.23
286	GGPGAAAOWALLWAAAAAAGPGG	0.97	0.36
287	GGPGAAAOWALLWAAAAAAGPGG	0.71	0.08
288	GGPGAAAOWALLWAAAAAAGPGG	0.24	0.25
289	GGPGAAAOWALLWAAAAAAGPGG	-1.06	-0.97
290	GGPGAAAOWALLWAAAAAAGPGG	-0.67	-1.05
291	GGPGAAAOWALLWAAAAAAGPGG	-1.65	-1.42
292	GGPGAAAOWALLWAAAAAAGPGG	0.97	0.72
293	GGPGAAAOWALLWAAAAAAGPGG	0.07	0.16
294	GGPGAAAOWALLWAAAAAAGPGG	-0.41	-0.28
295	GGPGAAAOWALLWAAAAAAGPGG	-1.07	-0.55

**Y**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
296	GGPGYAAAAALALALAAAAAAGPGG	0.16	-0.06
297	GGPGAYAAAAALALALAAAAAAGPGG	-0.23	-0.22
298	GGPGAAYAAAAALALALAAAAAAGPGG	-0.31	-0.34
299	GGPGAAAYAAAAALALALAAAAAAGPGG	-0.13	-0.11
300	GGPGAAAAYAAALALALAAAAAAGPGG	-0.14	0.03
301	GGPGAAAAYALALALAAAAAAGPGG	0.22	0.23
302	GGPGAAAAYLALALAAAAAAGPGG	0.13	0.39
303	GGPGAAAAYLYLALAAAAAAGPGG	0.44	0.44
304	GGPGAAAAYLALYLAAAAAAGPGG	0.69	0.41
305	GGPGAAAAYLALALYAAAAAAGPGG	0.05	0.39
306	GGPGAAAAYLALALAYAAAAAAGPGG	0.09	0.24
307	GGPGAAAAYLALALAYAAAAAAGPGG	-0.11	0.03
308	GGPGAAAAYLALALAYAAAAAAGPGG	-0.34	-0.10
309	GGPGAAAAYLALALAYAAAAAAGPGG	-0.61	-0.28
310	GGPGAAAAYLALALAYAAAAAAGPGG	-0.19	-0.23
311	GGPGAAAAYLALALAYAAAAAAGPGG	0.30	-0.08
312	GGPGAAAAYLLLYAAAAAAGPGG	0.97	0.90
313	GGPGAAAAYALLLYAAAAAAGPGG	0.88	0.40
314	GGPGAAAAYALLLYAAAAAAGPGG	0.16	0.26
315	GGPGAAAAYALLLYAAAAAAGPGG	0.11	-0.13

316	GGPGAAAYAAAAALLLAAAAAYAGPGG	-0.54	-0.30
317	GGPGAYAAAAALLLAAAAAYAGPGG	-0.33	-0.55
318	GGPGYAAAAALLLAAAAAYGPGG	-0.71	-0.07
319	GGPGAAAALALAAYALALAAAAGPGG	-0.03	0.13
320	GGPGAAAALAAAAYAAAALAAAAGPGG	1.64	1.32

**F**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
321	GGPGAAAAFAAALAAAAAAGPGG	0.50	0.87
322	GGPGAAAAAALAAFAAAAAGPGG	0.67	0.86
323	GGPGAAAAAFLFAAAAAAGPGG	0.28	0.54
324	GGPGAAAAAFALAFAAAAAGPGG	0.09	0.43
325	GGPGAAAAFAALAAFAAAAAGPGG	0.40	0.65
326	GGPGAAAAFAAALAAFAAAAAGPGG	0.73	0.68
327	GGPGAAAAFAAALAAFAAAAAGPGG	0.55	0.40
328	GGPGAAAFAAAAALAAAFAAAGPGG	0.21	0.52
329	GGPGAAFAAAAAALAAAAFAAGPGG	0.40	0.76
330	GGPGAFAAAAALAAAAAFAGPGG	0.40	0.67
331	GGPGFAAAAAALAAAAAFGPGG	0.28	0.62
332	GGPGAAAALAAAFAAAAALAAAGPGG	0.17	0.34

**C**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
333	GGPGAACALAAAAALAAAAGPGG	-0.03	-0.07
334	GGPGAAAALAAALCAAAAALAAAGPGG	0.01	0.01
335	GGPGAAAALAAAAALAAACAGPGG	0.04	-0.18
336	GGPGAAAALAAAACAAAALAAAGPGG	0.33	0.39
337	GGPGAAAALALAACAALALAAAGPGG	-0.63	-0.64
338	GGPGCAAAAAALAAAAAACGPGG	1.58	1.01

**I**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
339	GGPGAAIAAAAAALAAAALAAAGPGG	0.50	0.45
340	GGPGAAAAIAAALAAAALAAAGPGG	0.16	0.26
341	GGPGAAAALAAAIAAALAAAGPGG	-0.12	-0.12
342	GGPGAAAALAAAALAAIAAAGPGG	0.35	0.25
343	GGPGAAAALAAAALAAAAIAAGPGG	0.31	0.45
344	GGPGIAAAAAALAAAAAIAAGPGG	0.71	0.84

**M**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
345	GGPGAAMAAAAAALAAAALAAAGPGG	0.68	0.40
346	GGPGAAAAAAMAAAALAAAALAAAGPGG	0.48	0.29
347	GGPGAAAALAAAAMAAAALAAAGPGG	0.39	0.14
348	GGPGAAAALAAAALAAAMAAAAGPGG	0.67	0.27
349	GGPGAAAALAAAALAAAAAMAAGPGG	0.77	0.40
350	GGPGMAAAAAALAAAAAAMGPGG	0.71	0.62

**V**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
351	GGPGAAVAAAAALAAAALAAAGPGG	0.71	0.44

352	GGPGAAAAVAAALAAAAAAGPGG	0.57	0.35
353	GGPGAAAAAALAAAVAAAAAAGPGG	0.18	0.37
354	GGPGAAAAAALAAALAAVAAAAAGPGG	0.65	0.34
355	GGPGAAAAAALAAALAAAAVAAGPGG	0.76	0.43
356	GGPGAAAAALVAAALAAAAAIAAGPGG	0.05	0.33
357	GGPGVAAAAAALAAAAAAAVGPGG	0.92	0.66

**H-segments of variable lengths**

<i>ID</i>	<i>Sequence</i>	$\Delta G_{app}$	$\Delta G_{app}^{pred}$
358	GGPGAAAAAAAAAAAAAAAAAAAAAAAAAGPGG	0.86	0.39
359	GGPGAAAAAAAAAAAAAAAAAAAAAAAAAGPGG	0.43	0.23
360	GGPGAAAAAAAAAAAAAAAAAAAAAAAAAGPGG	-0.32	-0.32
361	GGPGAAAAAAAAALAAAAAAAAAGPGG	0.57	0.59
362	GGPGAAAAAAAAALAAAAAAAAAGPGG	0.26	0.18
363	GGPGAAAAAAAAALAAAAAAAAAGPGG	-0.06	-0.17
364	GGPGAAAAAAAAALAAAAAAAAAGPGG	-0.70	-0.75
365	GGPGAAAAALAAAAAAGPGG	0.35	0.56
366	GGPGAAAAALAAAAAAGPGG	0.12	0.08
367	GGPGAAAAALAAAAAAGPGG	-0.37	-0.35
368	GGPGAAAAALAAAAAAGPGG	-0.87	-0.68
369	GGPGAAAAALAAALAAAAAGPGG	0.77	0.69
370	GGPGAAAAALAAALAAAAAGPGG	0.31	0.15
371	GGPGAAAAALAAALAAAAAGPGG	-0.55	-0.33
372	GGPGAAAAALAAALAAAAAGPGG	-1.38	-0.75
373	GGPGAAAAALLLLLAAAAAGPGG	0.53	0.03
374	GGPGAAAAALLLLLAAAAAGPGG	0.05	-0.26
375	GGPGAAAAALLLLLAAAAAGPGG	-0.57	-0.55
376	GGPGAAAAALLLLLAAAAAGPGG	-0.74	-0.83
377	GGPGALAALLLLLAAALAGPGG	0.16	0.07
378	GGPGALAALLLLLAAALAGPGG	-0.28	-0.28
379	GGPGALAALALAAALALALAGPGG	-1.64	-1.76
380	GGPGLLLLLLLGPGG	2.60	3.47
381	GGPGLLLLLLLGPGG	1.07	1.57
382	GGPGLLLLLLLGPGG	-0.41	-0.01
383	GGPGLLLLLLLGPGG	-1.47	-1.94
384	GGPGAAAAAAAAAAAAAAAAAAAAAGPGG	0.35	0.43
385	GGPGAAAAAAAAAAAAAAAAAAAAAGPGG	0.20	0.08
386	GGPGAAAAAAAAAAAAAAAAAAAAAGPGG	-0.57	-0.46
387	GGPGAAAAAALAAAAAAGPGG	-0.64	-0.69
388	GGPGAAAAAALAAAAAAGPGG	-0.27	-0.59
389	GGPGAAAAAALAAAAAAGPGG	-0.33	-0.56
390	GGPGAAAAAALAAAAAAGPGG	-0.71	-0.61
391	GGPGAAAAAALAAAAAAGPGG	-0.55	-0.47
392	GGPGAAAAAALAAAAAAGPGG	-0.53	-0.62
393	GGPGAAALAAAAAAGPGG	-0.70	-0.39
394	GGPGAAALAAAAAAGPGG	-0.48	-0.55
395	GGPGALAAAAAALAGPGG	-0.41	-0.35
396	GGPGLAAAAAALGPGG	-0.62	-0.41
397	GGPGLKLLALAAALALLALGPGG	-0.69	-0.46
398	GGPGLALLKLAALALLALGPGG	0.32	0.58
399	GGPGLALLLKAALALLALGPGG	0.98	1.06
400	GGPGLALLLAKALALLALGPGG	1.24	1.22
401	GGPGLALLLAAKLALLALGPGG	0.50	1.06
402	GGPGLALLLAAALKLLALGPGG	0.46	0.58

403	GGPGLALLLAAAALALLKLGPGG	-0.77	-0.46
404	GGPGKAAAAAAAAALAAAALAAAAAAAAAGPGG	-0.54	-0.35
405	GGPGAACKAAAAALAAAALAAAAAAAAAGPGG	-0.12	-0.15
406	GGPGAAAAKAAAAALAAAALAAAAAAAAAGPGG	-0.19	0.18
407	GGPGAAAAAAKALAAAALAAAAAAAAAGPGG	0.54	0.89
408	GGPGAAAAAAKALAAAALAAAAAAAAAGPGG	0.59	0.87
409	GGPGAAAAAAALKAAAALAAAAAAAAAGPGG	0.93	1.39
410	GGPGAAAAAAALAKAALAAAAAAAAAGPGG	0.77	1.21
411	GGPGAAAAAAALAAAALAAAAAAAAAGPGG	0.85	1.39
412	GGPGAAAAAAALAAAALKAAAAAAAAAGPGG	0.21	0.87
413	GGPGAAAAAAALAAAALAKAAAAAAAAAGPGG	0.42	0.89
414	GGPGAAAAAAALAAAALAAAKAAAAGPGG	-0.25	0.18
415	GGPGAAAAAAALAAAALAAAAKAAAGPGG	-0.44	-0.15
416	GGPGAAAAAAALAAAALAAAAAAKPGG	-0.62	-0.35
417	GGPGAWAAAAALLAAAAAAAAAGPGG	0.12	0.26
418	GGPGAWAAAAALLAAAAAAWAGPGG	0.22	0.12
419	GGPGAWAAAAALLAAAAAAWAGPGG	-0.25	-0.20
420	GGPGAAAAAAAAAAAAALAAAAAAAAAWAGPGG	-0.29	-0.43
421	GGPGAWAAAAAAAAAAAAALAAAAAAAAAWAGPGG	-0.61	-0.43
422	GGPGAWAAAAAAAAAAAAALAAAAAAAAAWAGPGG	-0.79	-0.70

**H-segments with other flanks**

ID	Sequence	$\Delta G_{app}$	$\Delta \Delta G_{app}$ (vs G-flanks)
423	DDPDAALAAAALAAAALAAAADPDD	0.79	0.85
424	DDPDAALAAAALAAAALAAAAGPGG	0.91	0.97
425	GGPGAAAALAAAALAAAALAAAADPDD	-0.13	-0.07
426	EEPEAAAALAAAALAAAALAAAAGPGG	0.85	0.91
427	GGPGAAAALAAAALAAAALAAAAPPEE	-0.09	-0.04
428	NNPNAAAALAAAALAAAALAAAANPNN	0.63	0.69
429	NNPNAAAALAAAALAAAALAAAAGPGG	0.43	0.49
430	GGPGAAAALAAAALAAAALAAAANPNN	0.14	0.20
431	QQPQAAAALAAAALAAAALAAAAGPGG	0.53	0.59
432	GGPGAAAALAAAALAAAALAAAQPPQ	0.07	0.13
433	KKPKAAAALAAAALAAAALAAAAPPK	-0.69	-0.63
434	KKPKAAAALAAAALAAAALAAAAGPGG	0.34	0.40
435	GGPGAAAALAAAALAAAALAAAAPPK	-0.73	-0.67
436	RPRRAAAAALAAAALAAAALAAAARPRR	-1.10	-1.04
437	GPGGAAAALAAAALAAAALAAAARPRR	-0.77	-0.71
438	RPRRAAAAALAAAALAAAALAAAAGPGG	-0.40	-0.34
439	SPSSAAAALAAAALAAAALAAAASPS	-0.07	-0.01
440	GPGGAAAALAAAALAAAALAAAASPS	-0.16	-0.10
441	SPSSAAAALAAAALAAAALAAAAGPGG	-0.03	0.03
442	KKPKAAAALAAAALAAAALAAAADPDD	-0.12	-0.06
443	DDPDAALAAAALAAAALAAAAPPK	0.28	0.34
444	GGPGAAAALALAAAALALAAAAGPGG	-0.46	0.00
445	NNPNAAAALALAAAALALAAAANPNN	-0.21	0.25
446	NNPNAAAALALAAAALALAAAAGPGG	-0.13	0.33
447	GGPGAAAALALAAAALALAAAANPNN	-0.59	-0.13
448	DDPDAALALAAAALALAAAADPDD	0.47	0.93
449	DDPDAALALAAAALALAAAAGPGG	0.24	0.70
450	GGPGAAAALALAAAALALAAAADPDD	-0.69	-0.23
451	KKPKAAAALALAAAALALAAAAPPK	-1.43	-0.97
452	KKPKAAAALALAAAALALAAAAGPGG	-0.41	0.05
453	GGPGAAAALALAAAALALAAAAPPK	-1.20	-0.74
454	KKPKAAAALALAAAALALAAAADPDD	-0.46	0.00
455	DDPDAALALAAAALALAAAAPPK	-0.42	0.04
456	RPRRAAAAALALAAAALALAAAARPRR	-1.85	-1.39
457	SPSSAAAALALAAAALALAAAASPS	-0.40	0.06

458	NNPNLLLLLLLLLLNPNN	0.04	0.45
459	DDPDLLLLLLLLLLDPDD	0.73	1.14
460	DDPDLLLLLLLLLLGPGG	0.08	0.49
461	KKPKLLLLLLLLLLKPKK	-1.26	-0.85
462	GGPGLLLLLLLLLLLKPKK	-0.97	-0.56
463	DDPDLLLLLLLLLLKPKK	-0.37	0.04
464	NNPNAAAAAAAAALAAAAALAAAAAANPNN	-0.43	0.45
465	NNPNAAAAAAAAALAAAAALAAAAAAGPGG	-0.36	0.51
466	GGPGAAAAAAAAALAAAAALAAAAAANPNN	-0.49	0.38
467	DDPDAAAAAAAAALAAAAALAAAAAAGPGG	0.28	1.15
468	GGPGAAAAAAAAALAAAAALAAAAAAKPKK	-1.19	-0.32
469	DDPDAAAAAAAAALAAAAALAAAAAAKPKK	-0.41	0.46
470	NNPNAAAAAAAAALAAAAAANPNN	0.13	0.20
471	NNPNAAAAAAAAALAAAAAAGPGG	0.08	0.14
472	GGPGAAAAAAAAALAAAAAANPNN	0.02	0.09
473	NNPNAAAAAAAAALAAAAAANPNN	-0.32	0.05
474	NNPNAAAAAAAAALAAAAAAGPGG	-0.14	0.23
475	GGPGAAAAAAAAALAAAAAANPNN	-0.35	0.02

**TM helices from PDB structures**

ID	Sequence	$\Delta G_{app}$	$\Delta G_{app}^{pred}$	PDB ID	Position	Function	Species	
476	GGPGVMIGGSILAVILLIVMIGGPGG	-1.45	-0.76 (-0.76)	1pw4A	418-436	Glycerol-3-Phosphate transp.	Escherichia coli	
477	GGPGLWQIITICAAGAFISWALRGPGG	-1.08	-0.21 (-0.21)	leysL	93-111	Photosynth. Reaction Center	Thermochromatium tepidum	
478	GGPGASGGIILIIIAAILAMTMANGPGG	-0.97	-0.42 (-0.64)	lzcdA	12-30	Na+/H+ antiporter	Escherichia coli	
479	GGPGTSRLLLAGVALGIIICALMGP	-0.22	-0.22 (-0.22)	117vA	142-160	ABC transporter	Escherichia coli	
480	GGPGLGLFRLVRLRLLRLLRLLIIGPGG	-0.01	0.68 (0.68)	lorsC	94-112	KvAP potassium channel	Aeropyrum pernix	+
481	GGPGLGLFRLVRLRLLRLLRLLIIGPGG	0.60	0.92 (0.92)	lorsC	94-112	KvAP potassium channel	Aeropyrum pernix	+++
482	GGPGLGLFRLVRLRLLRLLRLLIIGPGG	0.56	0.37 (0.37)	lorsC	94-112	KvAP potassium channel	Aeropyrum pernix	++
483	GGPGALAIIDDLGAIILIIALFYTGPGG	0.61	0.55 (0.46)	lzcdA	158-176	Na+/H+ antiporter	Escherichia coli	
484	GGPGIAGYVGAATVGAANWFMYPGG	0.88	1.07 (1.07)	1wpgA	840-858	SR CA2+-ATPase	Oryctolagus cuniculus	*
485	GGPGIPVILAAALFANIQLWGLAGPGG	0.93	0.99 (0.99)	1rhzA	257-275	Translocase SecY subunit	Methanococcus jannaschii	
486	GGPGPEPMTMALSVLVTIEMCNAGPGG	1.28	3.93 (2.85)	1wpgA	894-912	SR CA2+-ATPase	Oryctolagus cuniculus	
487	GGPGMAIGGYVGAATVGAANWFMYPGG	1.43	1.80 (1.80)	1wpgA	838-856	SR CA2+-ATPase	Oryctolagus cuniculus	**
488	GGPGLIWAHGHPEVYILILPVFGPGG	1.48	2.79 (2.77)	1fftA	278-296	Ubiquinol oxidase	Escherichia coli	
489	GGPGIGGYVGAATVGAANWFMYPGG	1.71	1.42 (1.42)	1wpgA	840-858	SR CA2+-ATPase	Oryctolagus cuniculus	***

\*=essentially the same helix as \*\*\*, but mutated (G841A)  
\*\*=essentially the same helix as \*\*\*, but shifted 2 res.  
+=essentially the same helix as \*\*\*, but mutated (R104F, F105R, R107I, I108R)  
++=essentially the same helix as \*\*\*, but mutated (R107I, I108L, L109R)

**Single-pass TM protein from Swiss-Prot**

ID	Sequence	$\Delta G_{app}$	$\Delta G_{app}^{pred}$	Sprot ID	Position	Function	Species
490	GGPGLTGAGGFVLGLIICGVGIFGPGG	0.66	0.29 (0.29)	HB2P_HUMAN	227-245	HLA histocompat. antigen	Homo sapiens

**Multi-pass TM proteins from Swiss-Prot**

ID	Sequence	$\Delta G_{app}$	$\Delta G_{app}^{pred}$	Sprot ID	Position	Function	Species
491	GGPGLGFRIISMLRLWRLRRVSSGPGG	1.04	3.40 (1.81)	KAT1_ARATH	162-180	KAT1 potassium channel	Arabidopsis thaliana
492	GGPGLAILRVIRLVRVFRIFKLSGPGG	1.23	2.21 (1.90)	KCNAS_DROME	358-376	Shaker potassium channel	Drosophila melanogaster

**Soluble proteins from Swiss-Prot**

ID	Sequence	$\Delta G_{app}$	$\Delta G_{app}^{pred}$	Sprot ID	Position	Function	Species
493	GGPGTAAGGGAICAIAVMITIVMGPGG	0.59	0.62 (0.62)	LYS_BPP1	5-23	Lysozyme	Bacteriophage P1
494	GGPGIFISPVSIISMALMSLGTGPGG	1.25	1.68 (1.68)	CBG_HUMAN	63-81	Corticosteroid-binding glob.	Homo sapiens
495	GGPGQRVIVVAGVAGLVAAKVLGPGG	1.39	1.78 (1.54)	OXLA_HUMAN	60-78	L-amino-acid oxidase	Homo sapiens
496	GGPGMAGAAAAGAVVGLGGYMLGPGG	1.85	2.07 (2.07)	Q53YK7	112-130	Prion protein	Homo sapiens
497	GGPGFVMATATAGLLYGAYAVTGPGG	2.04	2.04 (2.04)	PEX14_YEAST	98-115	Peroxisomal membrane prot.	Saccharomyces cerevisiae

$\Delta G_{app}^{pred}$  for optimal subsequence is given in parenthesis

**Specification of which constructs are included in the figures**

Fig. 2 All of #1-357 with  $\Delta G_{app}$  between -1.0 and +1.5, plus #184, #185 and #192 were used in the training resulting in the profiles

Fig. 3 #358-386

Fig. S1A	C	#333-335
	D	#81-92
	E	#230-236
	F	#321-322
	G	#194-204
	H	#219-225
	I	#339-343
	K	#44-56
	3L	#1-7
	M	#345-349
	N	#155-163
	P	#124-136
	Q	#249-255
	R	#184-192
	S	#100-114
	T	#240-247
	V	#351-355
	W	#261-276
	Y	#296-311

Fig. S1B	2F	#323-331
	2G	#205-211
	2K	#66-73
	2L	#18-27
	2N	#166-175
	2P	#137-146
	2S	#115-121
	2W	#277-284
	2Y	#312-318

Fig. S2 Black dots: #1-357  
Red dots: #476-497  
*Note: construct #43 is outside the limits of the plot*

Fig. S4A #44-56, #397-416

Fig. S4B #18-27, #387-396

Fig. S5A #423-457

Fig. S5B #458-469

**Supplementary table S2:**  
**Optimized model parameter values**

*Position specific  $\Delta G_{app}^{aa(i)}$  matrix (kcal/mol):*

i:	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9
aa:																			
A	0.02	0.03	0.04	0.06	0.07	0.09	0.10	0.12	0.12	0.13	0.12	0.12	0.10	0.09	0.07	0.06	0.04	0.03	0.02
C	0.00	0.00	0.00	0.00	-0.01	-0.02	-0.03	-0.05	-0.07	-0.08	-0.07	-0.05	-0.03	-0.02	-0.01	0.00	0.00	0.00	0.00
D	0.44	0.59	0.77	0.96	1.17	1.36	1.54	1.67	1.76	1.79	1.76	1.67	1.54	1.36	1.17	0.96	0.77	0.59	0.44
E	0.69	0.80	0.92	1.03	1.14	1.23	1.31	1.37	1.41	1.42	1.41	1.37	1.31	1.23	1.14	1.03	0.92	0.80	0.69
F	-0.25	-0.26	-0.26	-0.27	-0.27	-0.27	-0.27	-0.28	-0.28	-0.28	-0.28	-0.28	-0.27	-0.27	-0.27	-0.27	-0.26	-0.26	-0.25
G	0.33	0.36	0.38	0.41	0.43	0.45	0.46	0.47	0.48	0.48	0.48	0.47	0.46	0.45	0.43	0.41	0.38	0.36	0.33
H	0.63	0.72	0.81	0.90	0.98	1.06	1.12	1.16	1.19	1.20	1.19	1.16	1.12	1.06	0.98	0.90	0.81	0.72	0.63
I	-0.11	-0.14	-0.19	-0.24	-0.29	-0.34	-0.39	-0.43	-0.45	-0.46	-0.45	-0.43	-0.39	-0.34	-0.29	-0.24	-0.19	-0.14	-0.11
K	0.32	0.46	0.63	0.84	1.07	1.30	1.52	1.69	1.81	1.85	1.81	1.69	1.52	1.30	1.07	0.84	0.63	0.46	0.32
L	-0.35	-0.37	-0.38	-0.39	-0.40	-0.41	-0.42	-0.42	-0.43	-0.43	-0.43	-0.42	-0.42	-0.41	-0.40	-0.39	-0.38	-0.37	-0.35
M	0.00	0.00	0.00	0.00	-0.01	-0.02	-0.03	-0.05	-0.07	-0.08	-0.07	-0.05	-0.03	-0.02	-0.01	0.00	0.00	0.00	0.00
N	0.63	0.73	0.84	0.95	1.05	1.14	1.22	1.28	1.31	1.33	1.31	1.28	1.22	1.14	1.05	0.95	0.84	0.73	0.63
P	0.48	0.57	0.66	0.76	0.84	0.92	0.99	1.04	1.08	1.09	1.08	1.04	0.99	0.92	0.84	0.76	0.66	0.57	0.48
Q	0.54	0.65	0.77	0.89	1.01	1.11	1.21	1.28	1.32	1.33	1.32	1.28	1.21	1.11	1.01	0.89	0.77	0.65	0.54
R	0.03	0.06	0.13	0.26	0.46	0.73	1.04	1.34	1.57	1.65	1.57	1.34	1.04	0.73	0.46	0.26	0.13	0.06	0.03
S	0.37	0.43	0.48	0.53	0.58	0.62	0.65	0.68	0.70	0.70	0.70	0.68	0.65	0.62	0.58	0.53	0.48	0.43	0.37
T	0.04	0.07	0.11	0.17	0.24	0.32	0.40	0.46	0.51	0.53	0.51	0.46	0.40	0.32	0.24	0.17	0.11	0.07	0.04
V	0.00	0.00	0.00	-0.01	-0.02	-0.05	-0.10	-0.17	-0.22	-0.24	-0.22	-0.17	-0.10	-0.05	-0.02	-0.01	0.00	0.00	0.00
W	-0.24	-0.35	-0.42	-0.39	-0.27	-0.10	0.07	0.18	0.25	0.27	0.25	0.18	0.07	-0.10	-0.27	-0.39	-0.42	-0.35	-0.24
Y	-0.11	-0.19	-0.20	-0.10	0.08	0.27	0.44	0.54	0.60	0.62	0.60	0.54	0.44	0.27	0.08	-0.10	-0.20	-0.19	-0.11

Gauss function parameters (Eqn 5 and 6):

Amino acid:	$a_0$	$a_1$	$a_2$	$a_3$	$a_4$
A	1.27E-01	2.15E-02			
C	-7.65E-02	9.94E-02			
D	1.79E+00	1.73E-02			
E	1.42E+00	8.94E-03			
F	-2.77E-01	1.03E-03			
G	4.81E-01	4.72E-03			
H	1.20E+00	8.01E-03			
I	-4.60E-01	1.81E-02			
K	1.85E+00	2.18E-02			
L	-4.28E-01	2.38E-03			
M	-7.75E-02	9.84E-02			
N	1.33E+00	9.24E-03			
P	1.09E+00	1.01E-02			
Q	1.33E+00	1.12E-02			
R	1.65E+00	5.12E-02			
S	7.02E-01	7.77E-03			
T	5.27E-01	3.12E-02			
V	-2.45E-01	9.79E-02			
W	2.91E-01	1.89E-02	-5.48E-01	9.30E-02	6.47E+00
Y	6.28E-01	1.04E-02	-5.74E-01	9.48E-02	6.92E+00

Hydrophobic moment weight (Eqn 1):

$c_0$	2.70E-01
-------	----------

Length parameters (Eqn 8):

$c_1$	9.29E+00
$c_2$	-6.45E-01
$c_3$	8.22E-03