Table 8. Expected vs. actual frequency of the three-residue amyloidogenic motif LEF surrounded by amyloid breakers in protein databases

Motif*	P_{LEFXX} (P_{XXLEF})[†]	Nonredundant protein sequences (NR)*		TrEMBL (Tr)*		Swiss-Prot and TrEMBL (SP + Tr)*		Swiss-Prot (SP)*	
		O_{LEFXX} (O_{XXLEF})[‡]	E_{LEFXX} (E_{XXLEF})[§]	O_{LEFXX} (O_{XXLEF})[‡]	E_{LEFXX} (E_{XXLEF})[§]	O_{LEFXX} (O_{XXLEF})[‡]	E_{LEFXX} (E_{XXLEF})[§]	O_{LEFXX} (O_{XXLEF})[‡]	E_{LEFXX} (E_{XXLEF})[§]
LEF + +	$3.812e^{-6}$	520	571.86	633	673.08	515	551.14	**108**	**97.93**
+ + LEF	$3.812e^{-6}$	**1,055**	**571.86**	**1,206**	**673.08**	**1,032**	**551.14**	**117**	**97.93**
LEF − −	$4.089e^{-6}$	**763**	**613.43**	**877**	**722.01**	**732**	**591.20**	**145**	**105.05**
− − LEF	$4.089e^{-6}$	**873**	**613.43**	**1026**	**722.01**	**854**	**591.20**	**175**	**105.05**
P LEF	$1.337e^{-5}$	1755	2005.9	2093	2360.9	1716	1933.2	346	343.51
LEF P	$1.337e^{-5}$	1631	2005.9	1867	2360.9	1520	1933.2	270	343.51
P + LEF	$1.350e^{-6}$	138	202.47	153	238.31	116	195.14	24	34.67
LEF + P	$1.350e^{-6}$	198	202.47	225	238.31	197	195.14	41	34.67
P − LEF	$1.551e^{-6}$	203	232.69	224	273.87	201	224.26	46	39.85
LEF − P	$1.551e^{-6}$	185	232.69	235	273.87	185	224.26	46	39.85
+ P LEF	$1.350e^{-6}$	171	202.47	199	238.31	166	195.14	28	34.67
LEF P +	$1.350e^{-6}$	136	202.47	158	238.31	134	195.14	24	34.67
− P LEF	$1.551e^{-6}$	**248**	**232.69**	**292**	**273.87**	**245**	**224.26**	**56**	**39.85**
LEF P −	$1.551e^{-6}$	237	232.69	276	273.87	231	224.26	35	39.85
Mean difference[¶]		226.897		252.508		226.300		31.037	
df[¶]		3		3		3		4	
t value[¶]		2.2932		2.2914		2.3047		2.8344	
t probability[¶]		0.1053		0.1058		0.1045		0.04713	

*Motif database scanning has been carried out by using PATTINPROT (1).

[†]Motif probabilities have been calculated as the product of the expected individual amino acid frequencies (see *Methods* in the main text).

[‡]O_{LEFXX} (O_{XXLEF}) is the number of hits of an amyloid breaker motif that have been found on the database.

[§]E_{LEFXX} (E_{XXLEF}) is the expected number of hits of a given amyloid breaker motif based on the average composition of proteins (see *Methods* in the main text) and the total number of LEF hits found in database [O_{LEF} (NR) = 45,588; O_{LEF} (Tr) = 53,657; O_{LEF} (SP + Tr) = 43,936; O_{LEF} (SP) = 7,807].

[¶]Student's test of significance. Only data where O > E have been compared (boldface data).

1. Combet, C., Blanchet, C., Geourjon, C. & Deleage, G. (2000) *Trends Biochem. Sci.* **25,** 147-150.