Table 9. Expected vs. actual frequency of the three-residue amyloidogenic motif WEF surrounded by amyloid breakers in protein databases

Motif*	P_{WEFXX} $(P_{XXWEF})^{\dagger}$	Nonredundant protein sequences (NR)*		TrEMBL (Tr)*		Swiss-Prot and TrEMBL (SP + Tr)*		Swiss-Prot (SP)*	
		O_{WEFXX} $(O_{XXWEF})^{\ddagger}$	E_{WEFXX} $(E_{XXWEF})^{\S}$	O_{WEFXX} $(O_{XXWEF})^{\ddagger}$	E_{WEFXX} $(E_{XXWEF})^{\S}$	O_{WEFXX} $(O_{XXWEF})^{\ddagger}$	E_{WEFXX} $(E_{XXWEF})^{\S}$	O_{WEFXX} $(O_{XXWEF})^{\ddagger}$	E_{WEFXX} $(E_{XXWEF})^{\S}$
WEF + +	$4.236e^{-7}$	45	90.69	35	74.70	45	86.39	10	11.69
+ + WEF	$4.236e^{-7}$	89	90.69	65	74.70	88	86.39	**23**	**11.69**
WEF − −	$4.544e^{-7}$	**110**	**97.29**	82	80.13	92	92.67	10	12.54
− − WEF	$4.544e^{-7}$	79	97.29	68	80.13	77	92.67	9	12.54
P WEF	$1.486e^{-6}$	248	318.12	196	262.02	247	303.03	**51**	**41.01**
WEF P	$1.486e^{-6}$	**551**	**318.12**	**440**	**262.02**	**537**	**303.03**	97	41.01
P + WEF	$1.664e^{-7}$	**59**	**35.63**	**49**	**29.35**	**58**	**33.94**	9	4.59
WEF + P	$1.664e^{-7}$	20	35.63	13	29.35	20	33.94	7	4.59
P − WEF	$1.724e^{-7}$	**1,124**	**36.90**	967	30.39	991	35.15	24	4.76
WEF − P	$1.724e^{-7}$	**76**	**36.90**	**65**	**30.39**	**76**	**35.15**	**11**	**4.76**
+ P WEF	$1.664e^{-7}$	34	35.63	27	29.35	34	33.94	7	4.59
WEF P +	$1.664e^{-7}$	**160**	**35.63**	**117**	**29.35**	**153**	**33.94**	36	4.59
− P WEF	$1.724e^{-7}$	17	36.90	12	30.39	16	35.15	4	4.76
WEF P −	$1.724e^{-7}$	**51**	**36.90**	**41**	**30.39**	**51**	**35.15**	**10**	**4.76**
Mean difference¶		219.090		211.184		231.607		14.865	
df¶		6		5		5		9	
t value¶		1.482		1.4339		1.5577		2.7618	
t probability¶		0.1889		0.2110		0.1800		0.0221	

*Motif database scanning has been carried by out 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_{WEFXX} (O_{XXWEF}) is the number of hits of an amyloid breaker motif that have been found on the database.

[§]E_{WEFXX} (E_{XXWEF}) 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 WEF hits found in database [O_{WEF} (NR) = 7,230; O_{WEF} (Tr) = 5,955; O_{WEF} (SP + Tr) = 6,887; O_{WEF} (SP) = 932].

[¶]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.