**Table 11. Expected vs. actual frequency of the three-residue amyloidogenic motif LEF surrounded by amyloid breakers in protein sequences**

| Database* | + + LEF† | | LEF – –† | | – – LEF† | | – P LEF† | |
|---|---|---|---|---|---|---|---|---|
| | $O_{++LEF}$ § | $E_{++LEF}$ ¶ | $O_{LEF--}$ § | $E_{LEF--}$ ¶ | $O_{--LEF}$ § | $E_{--LEF}$ ¶ | $O_{-PLEF}$ § | $E_{-PLEF}$ ¶ |
| Nonredundant (NR) | 1,055 | 571.86 | 763 | 613.43 | 873 | 613.43 | 248 | 232.69 |
| TrEMBL (Tr) | 1,206 | 613.08 | 877 | 722.01 | 1,026 | 722.01 | 292 | 273.87 |
| Swiss-Prot and TrEMBL (SP +Tr) | 1,032 | 551.14 | 732 | 591.20 | 854 | 591.20 | 245 | 224.26 |
| Swiss-Prot (SP) | 117 | 97.932 | 145 | 105.05 | 175 | 105.05 | 56 | 39.848 |
| $P_{LEFXX}$ ($P_{XXLEF}$)‡ | $3.812e^{-6}$ | | $4.089e^{-6}$ | | $4.089e^{-6}$ | | $1.551e^{-6}$ | |
| Mean difference‖ | 393.947 | | 121.327 | | 224.077 | | 17.583 | |
| df‖ | 3 | | 3 | | 3 | | 3 | |
| *t* value‖ | 3.0858 | | 4.447 | | 4.2795 | | 14.563 | |
| *t* probability‖ | 0.05389 | | 0.02115 | | 0.02344 | | 0.00070 | |

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

†Only motifs where the difference between the number of observed and expected hits is statistically significant at a confidence level ≥5%.

‡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.

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