

**Table 10. Expected vs. actual frequency of the three-residue amyloidogenic motif LNF surrounded by amyloid breakers in protein databases**

Motif*	$P_{LNFXX}$ ( $P_{XXLNF}$ ) <sup>†</sup>	Nonredundant protein sequences (NR)*		TrEMBL (Tr)*		Swiss-Prot and TrEMBL (SP + Tr)*		Swiss-Prot (SP)*	
		$O_{LNFXX}$ ( $O_{XXLNF}$ ) <sup>‡</sup>	$E_{LNFXX}$ ( $E_{XXLNF}$ ) <sup>§</sup>	$O_{LNFXX}$ ( $O_{XXLNF}$ ) <sup>‡</sup>	$E_{LNFXX}$ ( $E_{XXLNF}$ ) <sup>§</sup>	$O_{LNFXX}$ ( $O_{XXLNF}$ ) <sup>‡</sup>	$E_{LNFXX}$ ( $E_{XXLNF}$ ) <sup>§</sup>	$O_{LNFXX}$ ( $O_{XXLNF}$ ) <sup>‡</sup>	$E_{LNFXX}$ ( $E_{XXLNF}$ ) <sup>§</sup>
LNF ++	2.219e <sup>-6</sup>	463	714.73	521	841.69	426	694.77	78	86.87
++ LNF	2.219e <sup>-6</sup>	<b>804</b>	<b>714.73</b>	811	841.69	645	694.77	<b>106</b>	<b>86.87</b>
LNF --	2.380e <sup>-6</sup>	545	766.68	722	902.87	535	745.27	95	93.18
-- LNF	2.380e <sup>-6</sup>	504	766.68	608	902.87	496	745.27	<b>115</b>	<b>93.18</b>
P LNF	7.784e <sup>-6</sup>	1,523	2,507.0	1,739	2,952.3	1,443	2,437	242	304.70
LNF P	7.784e <sup>-6</sup>	<b>12,795</b>	<b>2,507.0</b>	<b>13,885</b>	<b>2,952.3</b>	<b>12,739</b>	<b>2,437</b>	254	304.70
P + LNF	8.718e <sup>-7</sup>	140	253.05	166	298.01	136	245.99	25	30.76
LNF + P	8.718e <sup>-7</sup>	169	253.05	217	298.01	156	245.99	25	30.76
P – LNF	9.029e <sup>-7</sup>	143	290.82	162	342.48	136	282.70	28	35.35
LNF – P	9.029e <sup>-7</sup>	145	290.82	188	342.48	145	282.70	34	35.35
+ P LNF	8.718e <sup>-7</sup>	145	253.05	194	298.01	145	245.99	20	30.76
LNF P +	8.718e <sup>-7</sup>	156	253.05	197	298.01	156	245.99	33	30.76
– P LNF	9.029e <sup>-7</sup>	166	290.82	193	342.48	162	282.70	18	35.35
LNF P –	9.029e <sup>-7</sup>	212	290.82	261	342.48	208	282.70	21	35.35

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

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

<sup>‡</sup> $O_{LNFXX}$  ( $O_{XXLNF}$ ) is the number of hits of an amyloid breaker motif that have been found on the database.

<sup>§</sup> $E_{LNFXX}$  ( $E_{XXLNF}$ ) 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 LNF hits found in database [ $O_{LNF}$  (NR) = 56,977;  $O_{LNF}$  (Tr) = 67,098;  $O_{LNF}$  (SP + Tr) = 55,386;  $O_{LNF}$  (SP) = 6,925].

Data are in boldface if  $O > E$ .

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