

## **Supplementary Information**

### **Evidence of the reduced abundance of proline *cis* conformation in protein poly-proline tracts**

Annika Urbanek<sup>1,#</sup>, Matija Popovic<sup>1,#</sup>, Carlos A. Elena-Real<sup>1,#</sup>, Anna Morató<sup>1</sup>, Alejandro Estaña<sup>1,2</sup>, Aurélie Fournet<sup>1</sup>, Frédéric Allemand<sup>1</sup>, Ana M. Gil<sup>3</sup>, Carlos Cativiela<sup>3</sup>, Juan Cortés<sup>2</sup>, Ana I. Jiménez<sup>3</sup>, Nathalie Sibille<sup>1</sup>, Pau Bernadó<sup>1,\*</sup>

<sup>1</sup> Centre de Biochimie Structurale (CBS), INSERM, CNRS, Université de Montpellier. 29, rue de Navacelles, 34090 Montpellier. France.

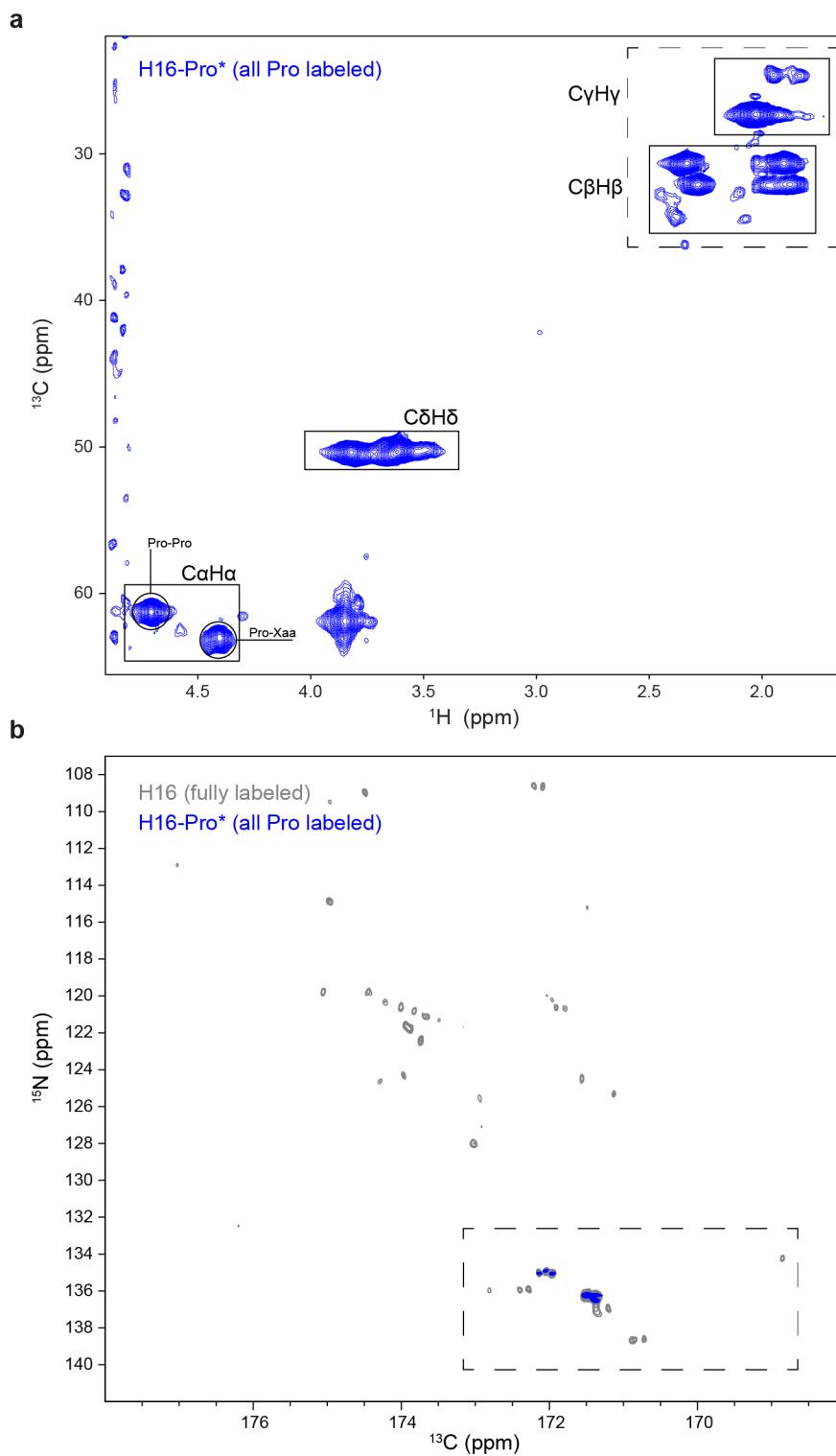
<sup>2</sup> LAAS-CNRS, Université de Toulouse, CNRS, 7 Avenue du Colonel Roche, 31400 Toulouse, France.

<sup>3</sup> Departamento de Química Orgánica, Instituto de Síntesis Química y Catalisis Homogénea (ISQCH), CSIC–Universidad de Zaragoza, 50009 Zaragoza, Spain.

# These authors contributed equally to this work

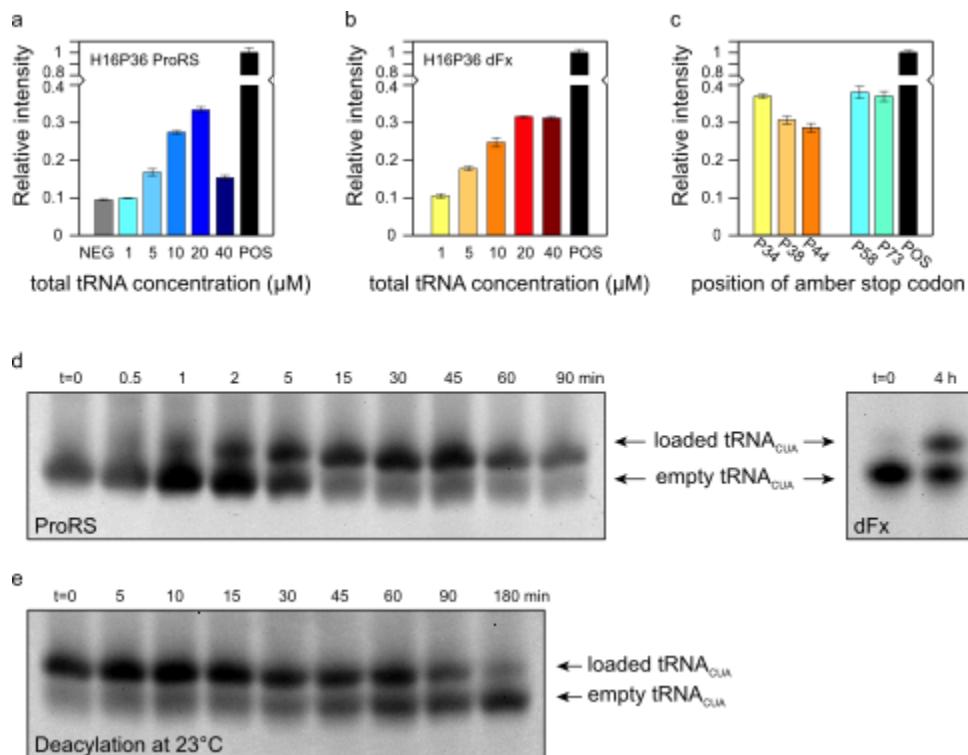
Corresponding Author: Pau Bernadó ([pau.bernardo@cbs.cnrs.fr](mailto:pau.bernardo@cbs.cnrs.fr))

**Figure S1**



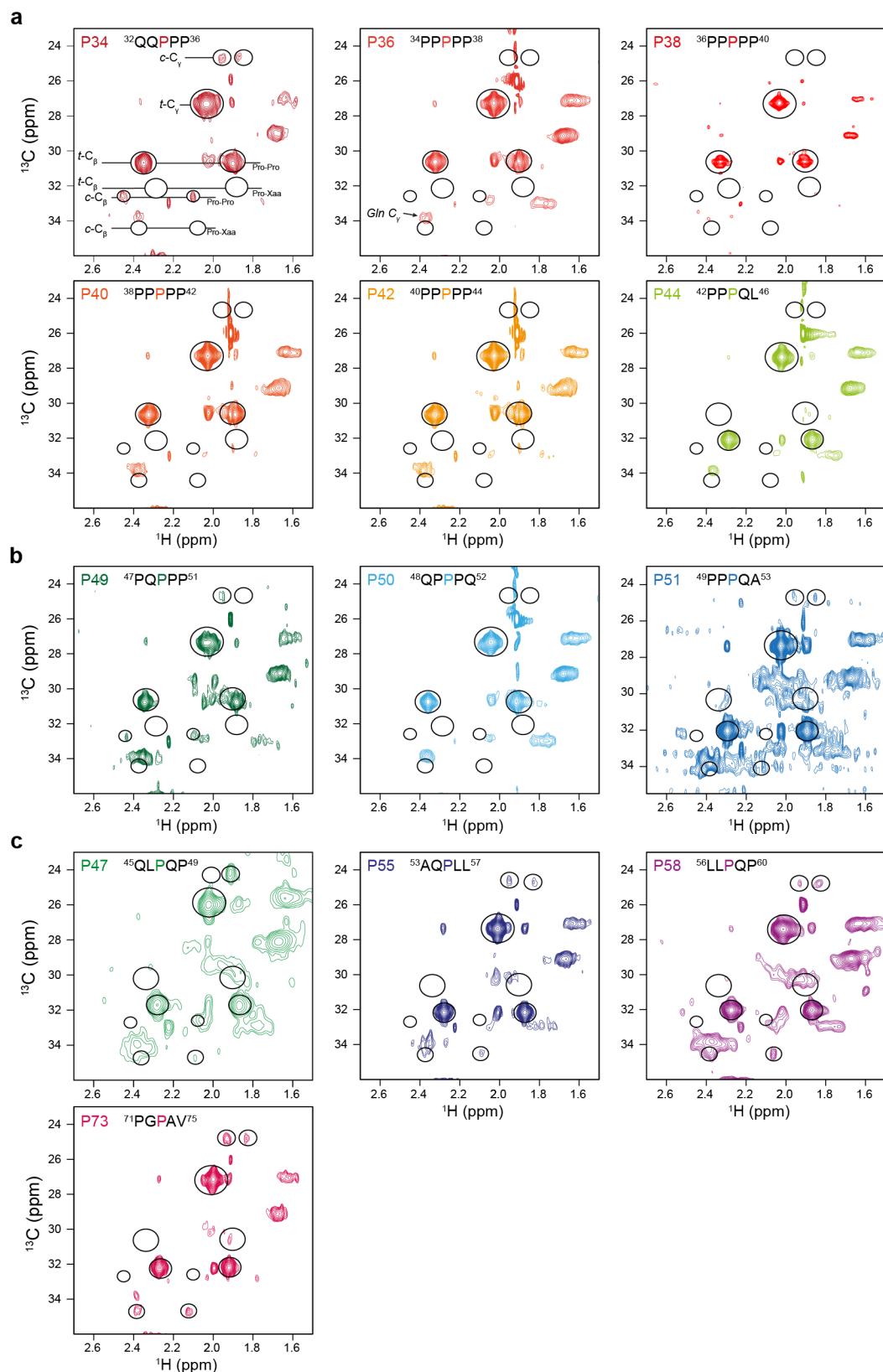
**Figure S1.**  $^{13}\text{C}$ -HSQC of H16-Pro\* and CON spectra. **(a)**  $^{13}\text{C}$ -HSQC of H16-Pro\* (only prolines are labeled, blue). Resonances of  $\text{Ca}$ ,  $\text{C}\beta$ ,  $\text{Cy}$  and  $\text{C}\delta$  are boxed. The two  $\text{Ca}$  peaks, corresponding to the two proline families (Pro-Pro and Pro-Xaa) are encircled. A zoom of the  $\text{C}\beta$  and  $\text{Cy}$  region (dashed box) is shown in Figure 2b in the main text. **(b)** CON overlay of fully labeled H16 (grey) and only proline labeled H16-Pro\*(blue). A zoom of the proline region (dashed box) is shown in Figure 2c in the main text.

**Figure S2**



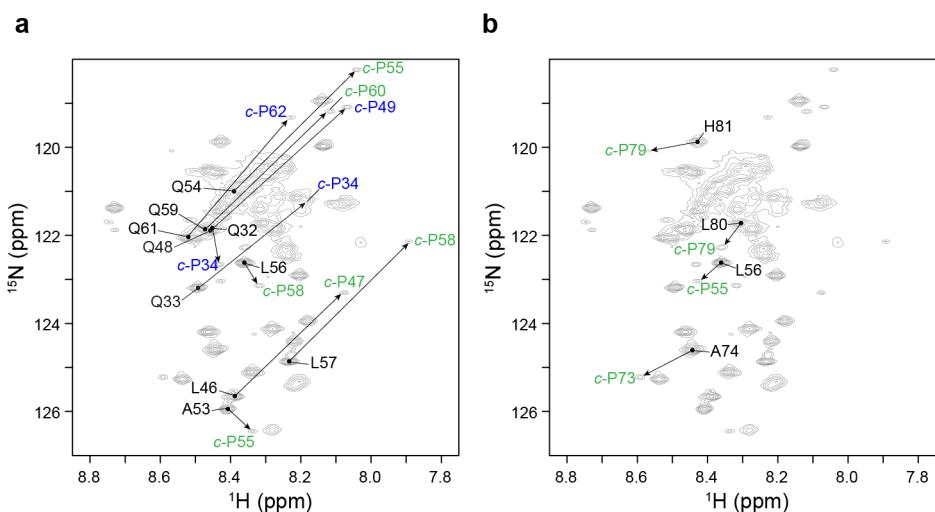
**Figure S2. Optimization of the SSIL approach for proline.** The optimal concentration of loaded tRNA<sub>CUA</sub> to obtain maximum amounts of full-length protein was determined using H16-sfGFP with an amber stop codon at position P36 (H16P36). **(a,b)** Titration of H16P36 with increasing amounts of loaded tRNA<sub>CUA</sub> (1, 5, 10, 20 or 40  $\mu\text{M}$  final concentration) prepared enzymatically (ProRS) and with flexizyme (dFx). NEG indicates a control reaction without added loaded tRNA<sub>CUA</sub> and POS corresponds to a control reaction with an H16-sfGFP plasmid without amber stop codon. Both preparations resulted in a similar concentration-dependent increase in sfGFP fluorescence, except that the intensity at 40  $\mu\text{M}$  loaded tRNA<sub>CUA</sub> dropped in case of ProRS and remained at a plateau for dFx. This might be explained by a combination of saturation of available elongation factor Tu with loaded tRNA<sub>CUA</sub> and simultaneous deacylation of free loaded tRNA<sub>CUA</sub> (see below). **(c)** A scan probing the suppression efficiency using 10  $\mu\text{M}$  ProRS-loaded tRNA<sub>CUA</sub>, which displayed the best ratio between yield and amount of tRNA<sub>CUA</sub>, showed no strong position-specific effects. A selection of different TAG positions in poly-P<sub>11</sub> (P34, P38 and P44) and isolated prolines (P58 and P73) is shown. All positions showed fluorescence intensities around 30 - 40% of the positive control (POS, H16-sfGFP without amber stop codon), indicating that the efficiency of the incorporation of the labeled proline was sequence-independent. **(d)** Gel of tRNA<sub>CUA</sub> loading time course using ProRS (left) and before/after loading with dFx for 4 hours (right). While enzymatically charged tRNA<sub>CUA</sub> was  $\geq 80\%$  loaded, tRNA<sub>CUA</sub> loaded by flexizyme was  $\leq 40\%$  loaded. **(e)** Time course following the deacylation of ProRS loaded tRNA<sub>CUA</sub> under similar conditions as used for cell-free expression (55 mM HEPES-KOH pH 7.5, 150 mM KCl, 23°C).

**Figure S3**



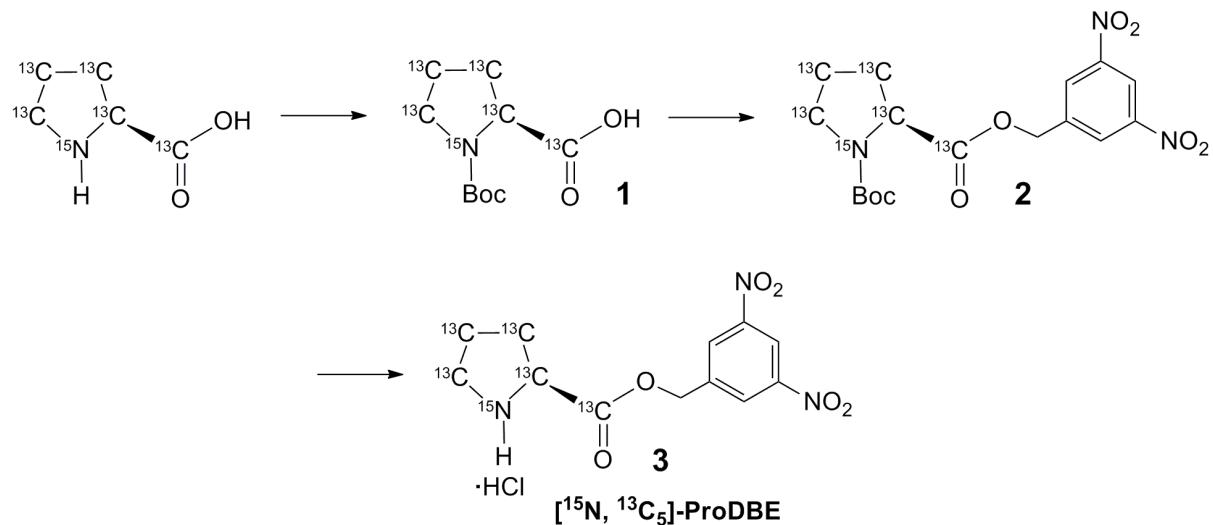
**Figure S3.**  $^{13}\text{C}$ -HSCQ spectra of all studied site-specifically labeled prolines. (a) prolines within poly-P<sub>11</sub>, (b) prolines in poly-P<sub>3</sub> and (c) isolated prolines. Circles indicate the positions of the C $\beta$  and Cy *trans* and *cis* peaks. An arrow in panel P36 indicates spectral contamination arising from natural abundance of  $^{13}\text{C}$  in glutamine Cy, which appear in all spectra.

**Figure S4**



**Figure S4.** <sup>15</sup>N-HSQC of H16. The assigned major peaks (*trans* conformation, black labels) of residues neighboring prolines are displayed. Arrows are pointing out the corresponding minor peaks (*cis* conformation, blue and green labels). *cis* peaks appearing upfield (a) and downfield (b) of the main *trans* peak are displayed separately.

**Figure S5**



**Figure S5. Synthesis of [<sup>15</sup>N, <sup>13</sup>C]-ProDBE:** L-[<sup>15</sup>N, <sup>13</sup>C<sub>5</sub>]-proline-3,5-dinitrobenzyl ester hydrochloride (**3**). Boc: *tert*-butoxycarbonyl; DBE: 3,5-dinitrobenzyl ester.

**Table S1. List of sample concentrations and NMR parameters.**

Proline	Protein concentration ( $\mu\text{M}$ )	Number of scans	Number of increments (TD in F1)	Magnetic field (MHz)	Loading method
34	11.4	928	96	800	dFx
36	25.0	1128	128	800	ProRS
38	23.0	768	120	900	dFx
40	28.2	706	128	800	ProRS
42	17.6	1656	128	800	ProRS
44	20.4	1144	128	800	ProRS
47	11.5	1024	128	800	dFx
49	6.7	1696	128	800	dFx
50	12.8	1152	128	800	ProRS
51	17.2	1412	128	800	ProRS
55	11.0	792	128	800	dFx
58	7.5	1024	128	800	dFx
73	9.4	512	128	800	ProRS

**Table S2. List of all 400 possible *Xaa-Pro-Yaa* tri-peptides (*Xaa* and *Yaa* = any amino acid) extracted from a coil database built from high-resolution crystallographic structures.** The total occurrence of each tripeptide (#total), as well as the number of prolines in *cis* (#cis), *trans* (#trans) conformation and the calculated percentage of prolines in *cis* configuration (%cis) in the respective tripeptide are listed. The four tripeptides with less than 100 instances (low confidence level) are highlighted in grey. Proline-centered sequences found in H16 are highlighted in yellow and (\*) indicates those for which we have experimentally measured the *cis/trans* either by SSIL samples or by 3D-NMR.

Sequence	#cis	#trans	#Total	%cis
ALA-PRO-ALA	184	3789	3973	4.63
ALA-PRO-ARG	136	2273	2409	5.65
ALA-PRO-ASN	134	2263	2397	5.59
ALA-PRO-ASP	136	3442	3578	3.80
ALA-PRO-CYS	29	427	456	6.36
ALA-PRO-GLN	173	1488	1661	10.40
ALA-PRO-GLU	207	3555	3762	5.50
ALA-PRO-GLY	356	6954	7310	4.87
ALA-PRO-HIS	83	1591	1674	4.96
ALA-PRO-ILE	64	1854	1918	3.34
ALA-PRO-LEU	344	3974	4318	7.97
ALA-PRO-LYS	94	1726	1820	5.16
ALA-PRO-MET	15	689	704	2.13
ALA-PRO-PHE	224	1533	1757	12.70
ALA-PRO-PRO	69	2699	2768	2.49
ALA-PRO-SER	156	3612	3768	4.14
ALA-PRO-THR	110	2107	2217	4.96
ALA-PRO-TRP	57	1019	1076	5.30
ALA-PRO-TYR	213	1586	1799	11.80
ALA-PRO-VAL	109	2598	2707	4.03
ARG-PRO-ALA	67	2917	2984	2.25
ARG-PRO-ARG	94	1118	1212	7.76
ARG-PRO-ASN	110	1177	1287	8.55
ARG-PRO-ASP	313	3199	3512	8.91
ARG-PRO-CYS	8	176	184	4.35
ARG-PRO-GLN	49	1176	1225	4.10
ARG-PRO-GLU	63	2940	3003	2.10
ARG-PRO-GLY	183	4116	4299	4.26
ARG-PRO-HIS	28	819	847	3.31
ARG-PRO-ILE	45	827	872	5.16
ARG-PRO-LEU	284	2395	2679	10.60
ARG-PRO-LYS	124	1026	1150	10.80
ARG-PRO-MET	26	424	450	5.78
ARG-PRO-PHE	213	1342	1555	13.70
ARG-PRO-PRO	167	1599	1766	9.46
ARG-PRO-SER	68	2139	2207	3.08
ARG-PRO-THR	108	1996	2104	5.13
ARG-PRO-TRP	29	632	661	4.39

Sequence	#cis	#trans	#Total	%cis
ARG-PRO-TYR	53	804	857	6.18
ARG-PRO-VAL	75	1542	1617	4.64
ASN-PRO-ALA	120	2717	2837	4.23
ASN-PRO-ARG	137	1637	1774	7.72
ASN-PRO-ASN	114	3357	3471	3.28
ASN-PRO-ASP	21	4629	4650	0.45
ASN-PRO-CYS	33	476	509	6.48
ASN-PRO-GLN	23	1822	1845	1.25
ASN-PRO-GLU	49	3600	3649	1.34
ASN-PRO-GLY	120	2670	2790	4.30
ASN-PRO-HIS	107	1168	1275	8.39
ASN-PRO-ILE	126	1044	1170	10.80
ASN-PRO-LEU	256	2243	2499	10.20
ASN-PRO-LYS	25	2344	2369	1.06
ASN-PRO-MET	23	719	742	3.10
ASN-PRO-PHE	102	1464	1566	6.51
ASN-PRO-PRO	280	2253	2533	11.10
ASN-PRO-SER	133	2874	3007	4.42
ASN-PRO-THR	298	2669	2967	10.00
ASN-PRO-TRP	292	489	781	37.40
ASN-PRO-TYR	312	1368	1680	18.60
ASN-PRO-VAL	221	1671	1892	11.70
ASP-PRO-ALA	80	5403	5483	1.46
ASP-PRO-ARG	41	3816	3857	1.06
ASP-PRO-ASN	113	4237	4350	2.60
ASP-PRO-ASP	58	3602	3660	1.58
ASP-PRO-CYS	3	423	426	0.70
ASP-PRO-GLN	57	2071	2128	2.68
ASP-PRO-GLU	119	3547	3666	3.25
ASP-PRO-GLY	139	1938	2077	6.69
ASP-PRO-HIS	18	1238	1256	1.43
ASP-PRO-ILE	138	1160	1298	10.60
ASP-PRO-LEU	88	2926	3014	2.92
ASP-PRO-LYS	47	2851	2898	1.62
ASP-PRO-MET	28	601	629	4.45
ASP-PRO-PHE	61	1388	1449	4.21
ASP-PRO-PRO	182	1813	1995	9.12
ASP-PRO-SER	67	6095	6162	1.09

<b>Sequence</b>	<b>#cis</b>	<b>#trans</b>	<b>#Total</b>	<b>%cis</b>
<b>ASP-PRO-THR</b>	32	3391	3423	0.93
<b>ASP-PRO-TRP</b>	43	543	586	7.34
<b>ASP-PRO-TYR</b>	141	1257	1398	10.10
<b>ASP-PRO-VAL</b>	195	2324	2519	7.74
<b>CYS-PRO-ALA</b>	27	1117	1144	2.36
<b>CYS-PRO-ARG</b>	21	663	684	3.07
<b>CYS-PRO-ASN</b>	19	702	721	2.64
<b>CYS-PRO-ASP</b>	32	1003	1035	3.09
<b>CYS-PRO-CYS</b>	3	84	87	3.45
<b>CYS-PRO-GLN</b>	5	242	247	2.02
<b>CYS-PRO-GLU</b>	6	951	957	0.63
<b>CYS-PRO-GLY</b>	61	951	1012	6.03
<b>CYS-PRO-HIS</b>	8	362	370	2.16
<b>CYS-PRO-ILE</b>	12	499	511	2.35
<b>CYS-PRO-LEU</b>	15	533	548	2.74
<b>CYS-PRO-LYS</b>	2	640	642	0.31
<b>CYS-PRO-MET</b>	0	278	278	0.00
<b>CYS-PRO-PHE</b>	58	502	560	10.40
<b>CYS-PRO-PRO</b>	148	658	806	18.40
<b>CYS-PRO-SER</b>	24	660	684	3.51
<b>CYS-PRO-THR</b>	7	634	641	1.09
<b>CYS-PRO-TRP</b>	4	272	276	1.45
<b>CYS-PRO-TYR</b>	22	433	455	4.84
<b>CYS-PRO-VAL</b>	11	659	670	1.64
<b>GLN-PRO-ALA</b>	126	1499	1625	7.75
<b>GLN-PRO-ARG</b>	55	763	818	6.72
<b>GLN-PRO-ASN</b>	36	976	1012	3.56
<b>GLN-PRO-ASP</b>	73	3423	3496	2.09
<b>GLN-PRO-CYS</b>	0	326	326	0.00
<b>GLN-PRO-GLN*</b>	77	871	948	8.12
<b>GLN-PRO-GLU</b>	36	2259	2295	1.57
<b>GLN-PRO-GLY</b>	87	3572	3659	2.38
<b>GLN-PRO-HIS</b>	14	874	888	1.58
<b>GLN-PRO-ILE</b>	29	1262	1291	2.25
<b>GLN-PRO-LEU*</b>	44	2500	2544	1.73
<b>GLN-PRO-LYS</b>	21	1278	1299	1.62
<b>GLN-PRO-MET</b>	51	289	340	15.00
<b>GLN-PRO-PHE</b>	78	961	1039	7.51
<b>GLN-PRO-PRO*</b>	100	1187	1287	7.77
<b>GLN-PRO-SER</b>	83	1907	1990	4.17
<b>GLN-PRO-THR</b>	58	1185	1243	4.67
<b>GLN-PRO-TRP</b>	14	548	562	2.49
<b>GLN-PRO-TYR</b>	50	562	612	8.17
<b>GLN-PRO-VAL</b>	94	2126	2220	4.23
<b>GLU-PRO-ALA</b>	194	1449	1643	11.80
<b>GLU-PRO-ARG</b>	281	1572	1853	15.20

<b>Sequence</b>	<b>#cis</b>	<b>#trans</b>	<b>#Total</b>	<b>%cis</b>
<b>GLU-PRO-ASN</b>	582	1308	1890	30.80
<b>GLU-PRO-ASP</b>	63	2341	2404	2.62
<b>GLU-PRO-CYS</b>	8	521	529	1.51
<b>GLU-PRO-GLN</b>	108	555	663	16.30
<b>GLU-PRO-GLU</b>	253	2597	2850	8.88
<b>GLU-PRO-GLY</b>	176	3378	3554	4.95
<b>GLU-PRO-HIS</b>	130	867	997	13.00
<b>GLU-PRO-ILE</b>	69	1664	1733	3.98
<b>GLU-PRO-LEU*</b>	161	2917	3078	5.23
<b>GLU-PRO-LYS</b>	111	1615	1726	6.43
<b>GLU-PRO-MET</b>	37	747	784	4.72
<b>GLU-PRO-PHE</b>	220	1250	1470	15.00
<b>GLU-PRO-PRO</b>	221	1956	2177	10.20
<b>GLU-PRO-SER</b>	373	2207	2580	14.50
<b>GLU-PRO-THR</b>	215	1502	1717	12.50
<b>GLU-PRO-TRP</b>	63	444	507	12.40
<b>GLU-PRO-TYR</b>	329	1204	1533	21.50
<b>GLU-PRO-VAL</b>	875	2012	2887	30.30
<b>GLY-PRO-ALA*</b>	111	2588	2699	4.11
<b>GLY-PRO-ARG</b>	179	2548	2727	6.56
<b>GLY-PRO-ASN</b>	178	3485	3663	4.86
<b>GLY-PRO-ASP</b>	203	2975	3178	6.39
<b>GLY-PRO-CYS</b>	65	498	563	11.50
<b>GLY-PRO-GLN</b>	88	1427	1515	5.81
<b>GLY-PRO-GLU</b>	77	2989	3066	2.51
<b>GLY-PRO-GLY</b>	310	4321	4631	6.69
<b>GLY-PRO-HIS</b>	106	1276	1382	7.67
<b>GLY-PRO-ILE</b>	96	1304	1400	6.86
<b>GLY-PRO-LEU</b>	285	3912	4197	6.79
<b>GLY-PRO-LYS</b>	87	2162	2249	3.87
<b>GLY-PRO-MET</b>	76	669	745	10.20
<b>GLY-PRO-PHE</b>	397	1155	1552	25.60
<b>GLY-PRO-PRO</b>	94	1581	1675	5.61
<b>GLY-PRO-SER</b>	187	2526	2713	6.89
<b>GLY-PRO-THR</b>	179	3103	3282	5.45
<b>GLY-PRO-TRP</b>	121	631	752	16.10
<b>GLY-PRO-TYR</b>	447	951	1398	32.00
<b>GLY-PRO-VAL</b>	157	2596	2753	5.70
<b>HIS-PRO-ALA</b>	111	1921	2032	5.46
<b>HIS-PRO-ARG</b>	30	963	993	3.02
<b>HIS-PRO-ASN</b>	16	1947	1963	0.81
<b>HIS-PRO-ASP</b>	72	2392	2464	2.92
<b>HIS-PRO-CYS</b>	8	209	217	3.69
<b>HIS-PRO-GLN</b>	47	665	712	6.60
<b>HIS-PRO-GLU</b>	26	2239	2265	1.15
<b>HIS-PRO-GLY</b>	47	3177	3224	1.46

Sequence	#cis	#trans	#Total	%cis
HIS-PRO-HIS	27	502	529	5.10
HIS-PRO-ILE	37	575	612	6.05
HIS-PRO-LEU	75	1662	1737	4.32
HIS-PRO-LYS	24	1289	1313	1.83
HIS-PRO-MET	46	865	911	5.05
HIS-PRO-PHE	73	1314	1387	5.26
HIS-PRO-PRO	193	989	1182	16.30
HIS-PRO-SER	433	1990	2423	17.90
HIS-PRO-THR	97	1343	1440	6.74
HIS-PRO-TRP	0	489	489	0.00
HIS-PRO-TYR	46	1170	1216	3.78
HIS-PRO-VAL	60	1109	1169	5.13
ILE-PRO-ALA	119	2945	3064	3.88
ILE-PRO-ARG	37	1353	1390	2.66
ILE-PRO-ASN	33	2023	2056	1.61
ILE-PRO-ASP	40	3243	3283	1.22
ILE-PRO-CYS	28	910	938	2.99
ILE-PRO-GLN	196	1617	1813	10.80
ILE-PRO-GLU	48	3138	3186	1.51
ILE-PRO-GLY	203	3379	3582	5.67
ILE-PRO-HIS	43	1200	1243	3.46
ILE-PRO-ILE	8	1340	1348	0.59
ILE-PRO-LEU	27	1923	1950	1.38
ILE-PRO-LYS	24	2222	2246	1.07
ILE-PRO-MET	5	617	622	0.80
ILE-PRO-PHE	14	1561	1575	0.89
ILE-PRO-PRO	56	2626	2682	2.09
ILE-PRO-SER	284	2058	2342	12.10
ILE-PRO-THR	119	1871	1990	5.98
ILE-PRO-TRP	101	645	746	13.50
ILE-PRO-TYR	47	1099	1146	4.10
ILE-PRO-VAL	51	1553	1604	3.18
LEU-PRO-ALA	177	5669	5846	3.03
LEU-PRO-ARG	94	2714	2808	3.35
LEU-PRO-ASN	50	3195	3245	1.54
LEU-PRO-ASP	56	6276	6332	0.88
LEU-PRO-CYS	41	553	594	6.90
LEU-PRO-GLN*	142	1910	2052	6.92
LEU-PRO-GLU	141	6134	6275	2.25
LEU-PRO-GLY	291	6735	7026	4.14
LEU-PRO-HIS	28	1336	1364	2.05
LEU-PRO-ILE	109	2362	2471	4.41
LEU-PRO-LEU	373	3592	3965	9.41
LEU-PRO-LYS	97	3547	3644	2.66
LEU-PRO-MET	39	721	760	5.13
LEU-PRO-PHE	192	2228	2420	7.93

Sequence	#cis	#trans	#Total	%cis
LEU-PRO-PRO	106	4539	4645	2.28
LEU-PRO-SER	127	5315	5442	2.33
LEU-PRO-THR	57	3721	3778	1.51
LEU-PRO-TRP	36	738	774	4.65
LEU-PRO-TYR	94	2098	2192	4.29
LEU-PRO-VAL	186	3462	3648	5.10
LYS-PRO-ALA	117	1977	2094	5.59
LYS-PRO-ARG	55	1194	1249	4.40
LYS-PRO-ASN	315	1619	1934	16.30
LYS-PRO-ASP	46	3039	3085	1.49
LYS-PRO-CYS	281	239	520	54.00
LYS-PRO-GLN	31	1294	1325	2.34
LYS-PRO-GLU	76	3742	3818	1.99
LYS-PRO-GLY	89	6721	6810	1.31
LYS-PRO-HIS	55	604	659	8.35
LYS-PRO-ILE	59	1185	1244	4.74
LYS-PRO-LEU	191	3006	3197	5.97
LYS-PRO-LYS	161	1871	2032	7.92
LYS-PRO-MET	46	931	977	4.71
LYS-PRO-PHE	326	1974	2300	14.20
LYS-PRO-PRO	54	1479	1533	3.52
LYS-PRO-SER	104	3120	3224	3.23
LYS-PRO-THR	33	1709	1742	1.89
LYS-PRO-TRP	36	728	764	4.71
LYS-PRO-TYR	94	945	1039	9.05
LYS-PRO-VAL	86	2779	2865	3.00
MET-PRO-ALA	15	1026	1041	1.44
MET-PRO-ARG	9	496	505	1.78
MET-PRO-ASN	19	720	739	2.57
MET-PRO-ASP	0	970	970	0.00
MET-PRO-CYS	0	95	95	0.00
MET-PRO-GLN	8	629	637	1.26
MET-PRO-GLU	85	728	813	10.50
MET-PRO-GLY	118	1400	1518	7.77
MET-PRO-HIS	42	450	492	8.54
MET-PRO-ILE	14	436	450	3.11
MET-PRO-LEU	9	785	794	1.13
MET-PRO-LYS	8	645	653	1.23
MET-PRO-MET	13	266	279	4.66
MET-PRO-PHE	20	435	455	4.40
MET-PRO-PRO	12	574	586	2.05
MET-PRO-SER	33	562	595	5.55
MET-PRO-THR	23	465	488	4.71
MET-PRO-TRP	1	935	936	0.11
MET-PRO-TYR	21	414	435	4.83
MET-PRO-VAL	11	564	575	1.91

<b>Sequence</b>	<b>#cis</b>	<b>#trans</b>	<b>#Total</b>	<b>%cis</b>
PHE-PRO-ALA	154	2578	2732	5.64
PHE-PRO-ARG	167	967	1134	14.70
PHE-PRO-ASN	258	1513	1771	14.60
PHE-PRO-ASP	142	2902	3044	4.66
PHE-PRO-CYS	18	184	202	8.91
PHE-PRO-GLN	117	1211	1328	8.81
PHE-PRO-GLU	1123	2453	3576	31.40
PHE-PRO-GLY	298	3777	4075	7.31
PHE-PRO-HIS	167	894	1061	15.70
PHE-PRO-ILE	62	1629	1691	3.67
PHE-PRO-LEU	221	1715	1936	11.40
PHE-PRO-LYS	55	1850	1905	2.89
PHE-PRO-MET	38	160	198	19.20
PHE-PRO-PHE	130	880	1010	12.90
PHE-PRO-PRO	211	1735	1946	10.80
PHE-PRO-SER	198	1863	2061	9.61
PHE-PRO-THR	84	1489	1573	5.34
PHE-PRO-TRP	34	436	470	7.23
PHE-PRO-TYR	62	809	871	7.12
PHE-PRO-VAL	128	1806	1934	6.62
PRO-PRO-ALA	154	2115	2269	6.79
PRO-PRO-ARG	72	1183	1255	5.74
PRO-PRO-ASN	65	1464	1529	4.25
PRO-PRO-ASP	78	2003	2081	3.75
PRO-PRO-CYS	188	289	477	39.40
PRO-PRO-GLN*	107	1498	1605	6.67
PRO-PRO-GLU	164	3497	3661	4.48
PRO-PRO-GLY	206	4761	4967	4.15
PRO-PRO-HIS	73	1312	1385	5.27
PRO-PRO-ILE	50	1213	1263	3.96
PRO-PRO-LEU	1060	2576	3636	29.20
PRO-PRO-LYS	90	1949	2039	4.41
PRO-PRO-MET	23	752	775	2.97
PRO-PRO-PHE	511	1195	1706	30.00
PRO-PRO-PRO*	25	3074	3099	0.81
PRO-PRO-SER	110	2720	2830	3.89
PRO-PRO-THR	198	1748	1946	10.20
PRO-PRO-TRP	29	450	479	6.05
PRO-PRO-TYR	167	1477	1644	10.20
PRO-PRO-VAL	78	1571	1649	4.73
SER-PRO-ALA	271	2238	2509	10.80
SER-PRO-ARG	99	1583	1682	5.89
SER-PRO-ASN	137	2421	2558	5.36
SER-PRO-ASP	129	4183	4312	2.99
SER-PRO-CYS	92	390	482	19.10
SER-PRO-GLN	214	1324	1538	13.90

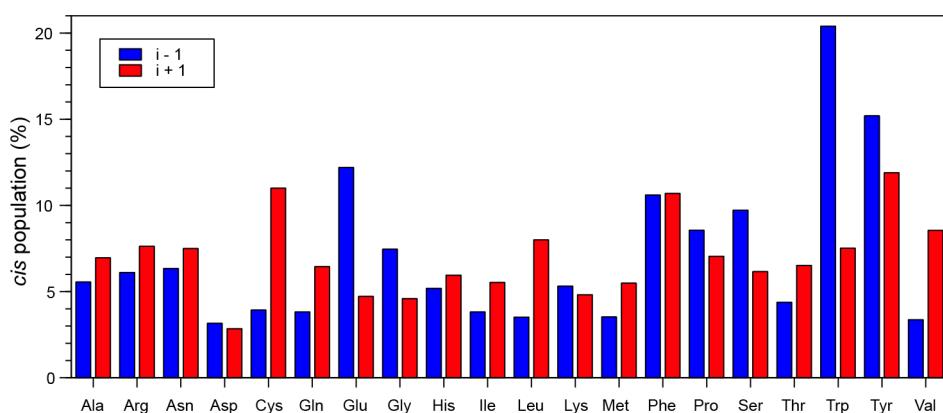
<b>Sequence</b>	<b>#cis</b>	<b>#trans</b>	<b>#Total</b>	<b>%cis</b>
SER-PRO-GLU	53	2837	2890	1.83
SER-PRO-GLY	139	2931	3070	4.53
SER-PRO-HIS	53	896	949	5.58
SER-PRO-ILE	213	1359	1572	13.50
SER-PRO-LEU	143	3239	3382	4.23
SER-PRO-LYS	95	1581	1676	5.67
SER-PRO-MET	43	655	698	6.16
SER-PRO-PHE	166	2835	3001	5.53
SER-PRO-PRO	425	1428	1853	22.90
SER-PRO-SER	112	3270	3382	3.31
SER-PRO-THR	606	1902	2508	24.20
SER-PRO-TRP	28	805	833	3.36
SER-PRO-TYR	467	1403	1870	25.00
SER-PRO-VAL	762	2157	2919	26.10
THR-PRO-ALA	649	2713	3362	19.30
THR-PRO-ARG	151	1671	1822	8.29
THR-PRO-ASN	275	1805	2080	13.20
THR-PRO-ASP	26	3882	3908	0.66
THR-PRO-CYS	19	551	570	3.33
THR-PRO-GLN	25	1146	1171	2.13
THR-PRO-GLU	65	3300	3365	1.93
THR-PRO-GLY	52	4884	4936	1.05
THR-PRO-HIS	32	1104	1136	2.82
THR-PRO-ILE	60	2198	2258	2.66
THR-PRO-LEU	114	3272	3386	3.37
THR-PRO-LYS	68	1764	1832	3.71
THR-PRO-MET	8	916	924	0.87
THR-PRO-PHE	92	1984	2076	4.43
THR-PRO-PRO	186	4206	4392	4.23
THR-PRO-SER	145	2883	3028	4.79
THR-PRO-THR	56	3070	3126	1.79
THR-PRO-TRP	26	1078	1104	2.36
THR-PRO-TYR	22	1547	1569	1.40
THR-PRO-VAL	102	3414	3516	2.90
TRP-PRO-ALA	115	692	807	14.30
TRP-PRO-ARG	266	208	474	56.10
TRP-PRO-ASN	157	394	551	28.50
TRP-PRO-ASP	80	960	1040	7.69
TRP-PRO-CYS	4	73	77	5.19
TRP-PRO-GLN	58	278	336	17.30
TRP-PRO-GLU	45	875	920	4.89
TRP-PRO-GLY	119	902	1021	11.70
TRP-PRO-HIS	33	153	186	17.70
TRP-PRO-ILE	109	147	256	42.60
TRP-PRO-LEU	104	360	464	22.40
TRP-PRO-LYS	44	224	268	16.40

<b>Sequence</b>	<b>#cis</b>	<b>#trans</b>	<b>#Total</b>	<b>%cis</b>
TRP-PRO-MET	4	22	26	15.40
TRP-PRO-PHE	101	833	934	10.80
TRP-PRO-PRO	216	241	457	47.30
TRP-PRO-SER	364	596	960	37.90
TRP-PRO-THR	84	663	747	11.20
TRP-PRO-TRP	21	177	198	10.60
TRP-PRO-TYR	54	143	197	27.40
TRP-PRO-VAL	203	557	760	26.70
TYR-PRO-ALA	818	1240	2058	39.70
TYR-PRO-ARG	461	945	1406	32.80
TYR-PRO-ASN	495	2008	2503	19.80
TYR-PRO-ASP	214	3147	3361	6.37
TYR-PRO-CYS	27	318	345	7.83
TYR-PRO-GLN	84	919	1003	8.37
TYR-PRO-GLU	97	1989	2086	4.65
TYR-PRO-GLY	323	3282	3605	8.96
TYR-PRO-HIS	85	643	728	11.70
TYR-PRO-ILE	62	793	855	7.25
TYR-PRO-LEU	241	1287	1528	15.80
TYR-PRO-LYS	484	1052	1536	31.50
TYR-PRO-MET	34	294	328	10.40
TYR-PRO-PHE	200	1132	1332	15.00
TYR-PRO-PRO	179	899	1078	16.60
TYR-PRO-SER	236	1454	1690	14.00

<b>Sequence</b>	<b>#cis</b>	<b>#trans</b>	<b>#Total</b>	<b>%cis</b>
TYR-PRO-THR	193	1439	1632	11.80
TYR-PRO-TRP	29	614	643	4.51
TYR-PRO-TYR	191	891	1082	17.70
TYR-PRO-VAL	68	971	1039	6.54
VAL-PRO-ALA	124	3295	3419	3.63
VAL-PRO-ARG	72	2086	2158	3.34
VAL-PRO-ASN	47	2287	2334	2.01
VAL-PRO-ASP	30	4242	4272	0.70
VAL-PRO-CYS	16	339	355	4.51
VAL-PRO-GLN	25	1561	1586	1.58
VAL-PRO-GLU	78	3368	3446	2.26
VAL-PRO-GLY	304	5442	5746	5.29
VAL-PRO-HIS	23	982	1005	2.29
VAL-PRO-ILE	47	1235	1282	3.67
VAL-PRO-LEU	113	2900	3013	3.75
VAL-PRO-LYS	34	1854	1888	1.80
VAL-PRO-MET	98	620	718	13.60
VAL-PRO-PHE	111	1595	1706	6.51
VAL-PRO-PRO	60	3782	3842	1.56
VAL-PRO-SER	89	2853	2942	3.03
VAL-PRO-THR	101	2072	2173	4.65
VAL-PRO-TRP	16	387	403	3.97
VAL-PRO-TYR	119	1734	1853	6.42
VAL-PRO-VAL	63	2337	2400	2.62

**Table S3.** Influence of the amino acids in position  $i-1$  and  $i+1$  on the percentage of *cis* conformations found in our large dataset of tripeptides. Numbers are directly derived by adding the number of #*cis*, #*trans* and #*total* from table S2. Amino acid specific influences on the *cis* population are displayed graphically in the appended plot.

Influence of aa in position $i-1$						Influence of aa in position $i+1$				
Sequence	# <i>cis</i>	# <i>trans</i>	#Total	% <i>cis</i>		Sequence	# <i>cis</i>	# <i>trans</i>	#Total	% <i>cis</i>
ALA-PRO-X	2893	49179	52072	5.56		X-PRO-ALA	3733	49888	53621	6.96
ARG-PRO-X	2107	32364	34471	6.11		X-PRO-ARG	2457	29753	32210	7.63
ASN-PRO-X	2792	41214	44006	6.34		X-PRO-ASN	3153	38901	42054	7.50
ASP-PRO-X	1650	50624	52274	3.16		X-PRO-ASP	1812	61853	63665	2.85
CYS-PRO-X	485	11843	12328	3.93		X-PRO-CYS	871	7081	7952	11.00
GLN-PRO-X	1126	28368	29494	3.82		X-PRO-GLN	1634	23704	25338	6.45
GLU-PRO-X	4469	32106	36575	12.20		X-PRO-GLU	2811	56738	59549	4.72
GLY-PRO-X	3443	42697	46140	7.46		X-PRO-GLY	3621	75291	78912	4.59
HIS-PRO-X	1468	26810	28278	5.19		X-PRO-HIS	1155	18271	19426	5.95
ILE-PRO-X	1483	37323	38806	3.82		X-PRO-ILE	1409	24086	25495	5.53
LEU-PRO-X	2436	66845	69281	3.52		X-PRO-LEU	4148	47717	51865	8.00
LYS-PRO-X	2255	40156	42411	5.32		X-PRO-LYS	1695	33490	35185	4.82
MET-PRO-X	461	12595	13056	3.53		X-PRO-MET	653	11235	11888	5.49
PHE-PRO-X	3667	30851	34518	10.60		X-PRO-PHE	3289	27561	30850	10.70
PRO-PRO-X	3448	36847	40295	8.56		X-PRO-PRO	2984	39318	42302	7.05
SER-PRO-X	4247	39437	43684	9.72		X-PRO-SER	3326	50704	54030	6.16
THR-PRO-X	2173	47388	49561	4.38		X-PRO-THR	2658	38079	40737	6.52
TRP-PRO-X	2181	8498	10679	20.40		X-PRO-TRP	980	12060	13040	7.52
TYR-PRO-X	4521	25317	29838	15.20		X-PRO-TYR	2951	21895	24846	11.90
VAL-PRO-X	1570	44971	46541	3.37		X-PRO-VAL	3535	37808	41343	8.55



This analysis suggests that aromatic residues and glutamic acid in position  $i-1$  strongly stabilize (> 10%) the *cis* conformation of prolines, in line with previous theoretical (Shen, Y. & Bax, A. *J. Biomol. NMR* 2010, **46**, 199–204; Pahlke, D. et al. *BMC Struct. Biol.* 2005, **5**, 8) and experimental studies (Mateos, B. et al. *J. Mol. Biol.* 2019, *in press*). The influence of position  $i+1$  seems to be less significant than position  $i-1$ , and only tyrosine, cysteine and phenylalanine seem to moderately privilege the *cis* conformation