

**Table 2: Comparison of the experimental cytosolic sample with the complete predicted *E. coli* proteome with respect to the number of predicted transmembrane segments (TMS), cellular localization from the PSORT-database and experimental localization data (EXP). Shown is the amount of unique proteins and the relation to the measured number of molecules in the cell.**

Attribute <sup>a</sup>	<i>E. coli</i> complete		Experimental cytosolic dataset		
	Proteins <sup>b</sup>	% Proteins <sup>c</sup>	Proteins <sup>b</sup>	% Proteins <sup>c</sup>	% Abundance <sup>d</sup>
TMS = 0	3202	75.66	940	89.5	97.6
TMS = 1	265	6.26	50	4.8	1.7
TMS = 2	117	2.76	10	1.0	0.2
TMS = 3	54	1.28	7	0.7	0.1
TMS = 4	82	1.94	7	0.7	6.2E-02
TMS = 5	61	1.44	5	0.5	2.9E-02
TMS = 6	81	1.91	5	0.5	4.0E-02
TMS = 7	30	0.71	1	0.1	1.1E-02
TMS = 8	52	1.23	3	0.3	2.6E-02
PSORT = Cytoplasmic (C)	1574	36.51	554	52.8	65.3
PSORT = CytoplasmicMembrane (CM)	851	19.74	93	8.9	1.2
PSORT = Periplasmic (P)	142	3.29	61	5.8	1.6
PSORT = OuterMembrane (OM)	91	2.11	25	2.4	2.3
PSORT = Extracellular (E)	20	0.46	0	0.0	
PSORT = Unknown (U)	1577	36.58	288	27.4	29.0
PSORT = Unknown (multiple sites) (UM)	56	1.30	14	1.3	0.4
PSORT = C   CM   U   UM	4058	94.13	949	90.4	95.9
PSORT = C   U	3054	71.21	842	80.2	94.3
TMS = 0 & PSORT = C	1253	29.21	548	52.2	65.1
TMS = 0 & PSORT = C   CM	1903	44.37	580	55.3	65.7
TMS = 0 & PSORT = C   CM   U	3111	72.53	843	80.3	94.3
TMS <= 1 & PSORT = C	1335	31.13	553	52.7	65.3
TMS <= 1 & PSORT = C   CM	2033	47.40	592	56.4	65.8
TMS <= 1 & PSORT = C   CM   U	3334	77.73	877	83.5	94.8
TMS <= 1 & PSORT = C   U	2636	61.46	838	79.8	94.3
EXP = C	370	18.57	279	26.6	63.0
EXP = IM	76	3.82	46	4.4	4.7
EXP = OM	62	3.11	40	3.8	2.1
EXP = P	60	3.01	43	4.1	1.7
TMS <= 1 & EXP = C	281	6.55	279	26.6	63.0
TMS <= 1 & EXP = IM	62	1.45	42	4.0	4.6
TMS <= 1 & EXP = OM	44	1.03	36	3.4	2.0
TMS <= 1 & EXP = P	48	1.12	43	4.1	1.7
TMS <= 1 & (PSORT = C U   EXP = C)	2655	61.90	853	81.2	94.6
(TMS <= 1 & PSORT = C U)   EXP = C	2680	62.49	853	81.2	94.6

<sup>a</sup> Annotated attributes of the proteins depicted as logical statements. An ampersand (&) indicates that both conditions must be fulfilled ('and'), a vertical line (|) indicates 'or'. The following abbreviations are used:

TMS – number of predicted transmembrane segments

PSORT – localization annotation from the PSORT database (C Cytoplasmic, CM Cytoplasmic Membrane, E Extracellular, OM Outer Membrane, P Periplasmic, U Unknown, UM Unknown – this protein may have multiple localization sites)

EXP – experimental localization data from [71] (C Cytoplasmic, IM Inner membrane, OM Outer Membrane, P Periplasmic)

<sup>b</sup> Number of unique proteins with the given attributes annotated

<sup>c</sup> Percentage of the unique proteins relative to the sum of unique proteins in the predicted *E. coli* proteome or in the experimental cytosolic sample, respectively

<sup>d</sup> Percentage of the actual number of protein copies found in the experimental sample, i.e. fraction of the total protein copy number sum.