Text S1: Supplemental Methods and Tables File

Supplemental Methods

Changes to Protein Location Annotations

Protein sequences and predicted cellular locations were from PSORTdb (exceptions *E. coli* – Echobase; *S. glossinidius, B. aphidicola* Sg, and *H. cicadicola* Dsem - NCBI). The location of several proteins in PSORTdb were annotated as ‘Unknown’ or given the wrong location altogether. For instance, some known ribosomal proteins were annotated to locations other than the cytoplasms (81). To refine the protein localization data, all known ribosomal proteins were changed to cytoplasmic. We also searched the literature for known extracellular proteins and reannotated their locations, as appropriate, when supported by experimental evidence. Those changes are as follows: CsgA,B and FimH for *S. typhimurium* LT2; MshA,B,O for *C. psychrerythraea* 34H (101); MshA,B,O and CsgA,B for *S. oneidensis* MR-1; CiaB for *W. succinogenes* DSM 1740 (9); AcuG, ComC, and ComP for *Acinetobacter sp.* ADP1 (BD413) (30, 51, 78); AprX and PopB for *P. aeruginosa* PAO1 (25, 35); TcpA, HylA (5), MshA,B (101), Hcp, and VgrG (16) for *V. cholera* El Tor; and HfaA for *C. cresentus* CB15 (99) were all changed to extracellular. In Gram-negative bacteria, PilA, FlgD,E,K,L (55), FlgM (40), FliK (61), and flagellins were changed to extracellular. For *F. johnsoniae* UW101, SprB (44) and homologs of the curli subunits CsgA,B were changed to extracellular. For *C. hutchinsonii* ATCC 33406, SprB was identified by blasting with FjoH_0979 of *F. johnsoniae* UW101. CHU_2225 (Uniprot Q11SX6_CYTH3) was identified as a SprB homolog (E value of 1e-104) and was changed to extracellular.

Electron Microscopy

For flagella, Type 1 pili, and curli TEMs, BW25113 was grown to late log in Lysogeny Broth (LB: 10g Tryptone, 5g Yeast extract, 10g NaCl per liter) shaking at 250 rpm, stationary in LB at 37°C for 16 h, or on YESCA plates (10g Casamino acids, 1g Yeast extract, 20g Agar per liter) for 48h at 26°C respectively. All samples were resuspended in PBS, incubated with formvar-coated copper grids (Ernest F Fullam, Inc), briefly washed with H_20, and stained 2 minutes in 2% uranyl acetate. Grids were viewed using a Phillips CM10 microscope.

Calculating Percent Savings for a Specific Amino Acid Bias in a Given Protein

Total ATPs saved in a protein:

Protein chain length × (Organism's mean Average Synthetic Cost – Protein Average Synthetic Cost)

ATPs saved per amino acid:

(Number of a specific amino acid in a protein – (Organism's mean % of specific amino acid × Protein chain length)) × (Average cost of all amino acids – Cost of a specific amino acid)

Percent Savings of a specific amino acid:

(ATPs saved per amino acid / Total ATPs saved) × 100

Type III Effectors

Type III effectors were culled from a recent survey of plant and animal pathogens (7) and are listed in Table S18B (Tab E in Data Set S1). For *P. syringae pv. tomato* str. DC3000 Type III effectors (also called Hops – Hrp outer protein) we used the Hop database, the T3SS helper database, and the Discontinued Hops database at the "*P. syringae* Hop Identification and Nomenclature Home Page" (http://pseudomonas-syringae.org/pst_func_gen2.htm). We included the following proteins: PSPTO_4101 (*hopAK1),* PSPTO_1022 (*hopAM1-1),* and PSPTO_5061 (*hopAN1)* which have plant phenotypes (31, 43, 86); PSPTO_1378 (*hrpH*) which was shown to be a T3SS substrate (75), and PSPTO_3292,3293,0907 which are translocated into plant cells (86). Confirmed Hop proteins were considered to be extracellular in later analysis. Only one copy of Hcp (*P. aeruginosa* PAO1*)* and HopAM1 (*P. syringae tomato*) were used in data analysis. See Table S18A in Data Set S1, Tab D for a complete table of *P. syringae* Hops and Hop helpers. Fig. 4A, S6, and Table S19 compare the ASC of Type III effectors to other cellular proteins.

Serine Proteases

Extracellular serine proteases were those annotated as “secreted” in Uniprot (http://www.uniprot.org). Cellular serine proteases were serine proteases not annotated as “secreted.” 5,784 bacterial serine proteases were downloaded from Uniprot using the keywords ”taxonomy:2 keyword:720” on August 18, 2009. Bacterial extracellular serine proteases (191 total) were downloaded using “taxonomy:2 keyword:720 keyword:964”. Taxonomy 2 = “Bacteria”; keyword 720 = “Serine protease”; and keyword 964 = “Secreted”.

Autotransporters

The autotransporters used were those described in Table 1 of Henderson et al 2004 (38). Uniprot (http://www.uniprot.org) was used to select autotransporters which had defined secreted α-domains. The ASC of both the α-domain and their respective β-domains were calculated using the cost values for chemoheterotrophic bacteria (2, 36).

Graphing of Protein Average Synthetic Cost (ASC)

To vertically spread the cost data within a location, protein length was used. The value $0.2\times\mathrm{LOG}_{10}$ (Length) was added to a specific integer of each location. For example cytoplasmic proteins were $1+0.2\times\mathrm{LOG}_{10}$ (Length), and inner membranes were $2+0.2\times\mathrm{LOG}_{10}$ (Length). These numbers were graphed vs. proteins costs to generate the graphs in Fig. 1A, 3B, 4, S1, S2, S5, S6.

Protein Abundance Values

Abundance values from Masuda et al (57) and Ishihama et al (42). Protein sequences and location were from Echobase; abundance values were merged with Echobase data using GenProtEC ECK and Uniprot numbers respectively.

Prediction of Secondary Structure and Intrinsic Disorder

The secondary structure of cytoplasmic, periplasmic, and extracellular proteins from *E. coli* K-12 and *P. syringae* were predicted using PSIPRED version 2.61 (58). Intrinsically disordered regions were predicted with DISOPRED2 version 2.4 (107) using a 2% false positive rate. Both programs were downloaded from (http://bioinfadmin.cs.ucl.ac.uk/downloads/).

Comparison of Amino Acid Composition of Flagellar Subunits

Extracellular assembly or structural components of *E. coli* flagella were individually blasted against 216 *Gammaproteobacteria*. Many flagella genes have similar regions, so sequences with E values below 10^{-10} were removed except for FlgG ($<10^{-20}$) and FliK ($<10^{-5}$) which have high identity to FlgG (52) and low ortholog identity respectively (108). Sequences whose lengths were 50% (+/-) their respective *E. coli* counterpart were removed. The remaining homologs were used in our analysis.

Table S1: Synthetic cost of amino acids.

									Gibbs Free Energy (ΔG_r)		Bacteria					Yeast	
				Codons		Molecular			Hydrothermal	Seawater	Chemoheterophs		Photoautotrophs*	Li et al*	Craig&Weber*	Respiratory	Fermentative
			Enzymes in *E. coli*	#	%GC	Mass	Carbons	Nitrogens	(kJ mol^{-1}) 100°C	(kJ mol^{-1}) 18°C	Cost (~P)	Savings (~P)	Cost (~P)	Cost (~P)	Cost (ATPs)	Cost (~P)	Cost (~P)
Alanine	Ala	A	1	4	83.33	71.079	3	1	-12.12	113.66	11.7	15.66	11.7	11.7	12.5	14.5	2
Arginine	Arg	R	10	6	72.22	156.188	6	4	197.52	409.46	27.3	0.06	27.3	**38.3**	18.5	20.5	13
Asparagine	Asn	N	1	2	16.67	114.104	4	2	83.53	201.56	14.7	12.66	14.7	**12.7**	4	18.5	6
Aspartic Acid	Asp	D	1	2	50.00	115.089	4	1	32.78	146.74	12.7	14.66	12.7	**14.7**	1	15.5	3
Cysteine	Cys	C	9	2	50.00	103.145	3	1	60.24	224.67	24.7	2.66	24.7	24.7	24.5	26.5	13
Glutamic Acid	Glu	E	1	2	50.00	129.116	5	1	-1.43	172.13	15.3	12.06	15.3	**34.3**	8.5	9.5	2
Glutamine	Gln	Q	2	2	50.00	128.131	5	2	44.03	223.36	16.3	11.06	16.3	**15.3**	9.5	10.5	3
Glycine	Gly	G	4	4	83.33	57.052	2	1	14.89	80.49	11.7	15.66	11.7	11.7	14.5	14.5	1
Histidine	His	H	1	2	50.00	137.141	5	3	154.48	350.52	38.3	-10.94	**40.3**	**32.3**	33	29	5
Isoleucine	Ile	I	11	3	11.11	113.16	6	1	-96.4	213.93	32.3	-4.94	32.3	**30.3**	20	38	14
Leucine	Leu	L	7	6	38.89	113.16	6	1	-105.53	205.03	27.3	0.06	27.3	27.3	33	37	4
Lysine	Lys	K	10	2	16.67	128.17	6	2	-28.33	258.56	30.3	-2.94	30.3	**27.3**	18.5	36	12
Methionine	Met	M	9	1	33.33	131.199	5	1	-174.71	113.22	34.3	-6.94	34.3	**20.3**	18.5	36.5	24
Phenylalanine	Phe	F	9	2	16.67	147.177	9	1	-114.54	303.64	52	-24.64	**54**	52	63	61	10
Proline	Pro	P	4	4	83.33	97.117	5	1	-38.75	192.83	20.3	7.06	20.3	**16.3**	12.5	14.5	7
Serine	Ser	S	3	6	50.00	87.078	3	1	69.47	173.73	11.7	15.66	11.7	11.7	15	14.5	1
Threonine	Thr	T	6	4	50.00	101.105	4	1	53.51	216.5	18.7	8.66	18.7	18.7	6	21.5	9
Tryptophan	Trp	W	12	1	66.67	186.213	11	2	-38.99	431.17	74.3	-46.94	**76.3**	74.3	78.5	75.5	14
Tyrosine	Tyr	Y	9	2	16.67	163.176	9	1	-59.53	334.2	50	-22.64	**52**	50	56.5	59	8
Valine	Val	V	4	4	50.00	99.133	5	1	-70.12	178	23.3	4.06	23.3	23.3	25	29	4
Source			(21)						(6)	(6)	(2, 36)		(36)	(50)	(21)	(105)	(105)
Average aa			5.7	3.05		118.887	5.3	1.45	-1.5	227.17	27.36		27.76	27.36	23.625	29.075	7.75

* Numbers in **BOLD** differ from chemoheterotrophic bacteria.

Notes: Values in Akashi and Gojobori 2002 (2) are based on the average of overall synthetic cost under growth on three substrates: glucose, acetate, and malate. Cost varied slightly with different carbon sources but were highly correlated (glucose vs acetate r^2=0.968, glucose vs malate r^2=0.998)(2). Bragg and Wagner(17) use 16 ~P (respiratory) and 1 ~P (fermentative) for Lys instead of 36 ~P and 12~P from Wagner(105); Lys production in yeast uses α-ketoglutarate instead of oxaloacetate(17). Average cost (equal amino acids) for respiratory and fermentative growth using these numbers would be 28.075 and 7.2 respectively.

Li et al 2009 (50) has the same values for eukaryotes as Wagner(105) and sources Heizer et al(36) values for prokaryotes. However the values listed in their supplementary section do not match with Heizer et al(36): Arg is listed at 38.3 not 27.3, Asp and Asn values were inverted, Glu is listed 34.3 not 15.3, Gln is listed 15.3 not 16.3, His is listed 32.3 not 38.3, Ile is listed 30.3 not 32.3, Lys is listed 27.3 not 30.3, Met is listed 20.3 not 34.3, and Pro is listed 16.3 not 20.3. No explanation was given in the manuscript as to why the alternative values were used (50).

Table S2: Correlation between different amino acid synthetic costs.

	Spearman		Kendall			Linear Correlation (Pearson)	
vs Chemoheterotrophic	**R**	**p value**	**Tau**	**p value**	**Gamma**	**R**	**p value**
Number of Enyzmes *E. coli*	0.6974	6.31E-04	0.5736	4.06E-04	0.6145	0.6920	7.24E-04
Number of Codons	-0.4565	0.0431	-0.3427	0.0346	-0.4058	-0.4309	0.0579
Number of Carbons	0.8429	3.08E-06	0.7176	9.71E-06	0.7875	0.9125	2.07E-08
Number of Nitrogens	0.2536	0.2806	0.1930	0.2341	0.2826	0.1871	0.4295
Average Codon %GC	-0.4438	0.0500	-0.3536	0.0293	-0.4000	-0.2855	0.2223
Average Codon %GC Weighted*	-0.4211	0.0645	-0.2979	0.0663	-0.3011	-0.3026	0.1947
Molecular Mass	0.7502	1.39E-04	0.6027	2.03E-04	0.6108	0.8030	2.02E-05
Surface Seawater ΔG(kJ mol-1) 18°C	0.7149	3.97E-04	0.5639	5.09E-04	0.5699	0.7480	1.49E-04
Hydrothermal Sol. ΔG (kJ mol-1) 100°C	-0.4588	0.0419	-0.2979	0.0663	-0.3011	-0.3260	0.1607
Photoautotrophic		<0.000001				0.9996	0, t=156
Yeast respiratory	0.8737	4.87E-07	0.7147	1.05E-05	0.7363	0.9540	7.37E-11
Yeast fermentative	0.7724	6.59E-05	0.6183	1.38E-04	0.6354	0.5414	0.0137
Yeast respiratory Bragg	0.8427	3.12E-06	0.6720	3.43E-05	0.6923	0.9360	1.35E-09
Yeast fermentative Bragg	0.6788	0.0010	0.5460	7.64E-04	0.5642	0.5155	0.0200
Li et al Prokaryotic	0.8615	1.06E-06	0.7566	3.10E-06	0.7730	0.9256	5.08E-09
Craig&Weber *E. coli*	0.8090	1.57E-05	0.6307	1.01E-04	0.6464	0.9337	1.84E-09
vs Yeast respiratory							
Number of Enyzmes *E. coli*	0.7223	3.23E-04	0.5525	6.60E-04	0.5976	0.7328	2.38E-04
Number of Codons	-0.3207	0.1680	-0.2575	0.1124	-0.3088	-0.3748	0.1035
Number of Carbons	0.7103	4.49E-04	0.5913	2.67E-04	0.6561	0.8809	2.96E-07
Number of Nitrogens	0.0417	0.8616	0.0292	0.8569	0.0435	-0.0059	0.9805
Average Codon %GC	-0.5558	0.0109	-0.4436	0.0063	-0.5067	-0.4183	0.0664
Molecular Mass	0.5272	0.0169	0.3871	0.0170	0.3956	0.6895	7.71E-04
Surface Seawater ΔG(kJ mol-1) 18°C	0.5459	0.0128	0.4129	0.0109	0.4208	0.6131	0.0040
Hydrothermal Sol. ΔG (kJ mol-1) 100°C	-0.5595	0.0103	-0.3164	0.0511	-0.3224	-0.4797	0.0323
Photoautotrophic	0.8737	4.87E-07	0.7147	1.05E-05	0.7363	0.9523	1.02E-10
Yeast fermentative	0.7106	4.46E-04	0.5203	0.0013	0.5393	0.5140	0.0204
Yeast respiratory Bragg	0.9682	2.85E-12	0.9000	2.89E-08	0.9344	0.9702	1.56E-12
Yeast fermentative Bragg	0.6167	0.0038	0.4469	0.0059	0.4659	0.4693	0.0368
Li et al Prokaryotic	0.6778	1.02E-03	0.5547	6.28E-04	0.5714	0.8590	1.24E-06
Craig&Weber *E. coli*	0.8063	1.76E-05	0.6541	5.53E-05	0.6760	0.9286	3.54E-09
vs Molecular Mass							
Number of Enyzmes *E. coli*	0.3811	0.0973	0.3068	0.0586	0.3254	0.5097	0.0217
Number of Codons	-0.6168	0.0038	-0.5225	0.0013	-0.6115	-0.4507	0.0461
Number of Carbons	0.7633	9.03E-05	0.6364	8.74E-05	0.6914	0.8669	7.59E-07
Number of Nitrogens	0.5111	0.0213	0.4265	0.0086	0.6129	0.4740	0.0347
Average Codon %GC	-0.3603	0.1187	-0.2823	0.0819	-0.3158	-0.3606	0.1184
Average Codon %GC Weighted*	-0.4438	0.0500	-0.3325	0.0404	-0.3333	-0.3769	0.1014
Surface Seawater ΔG(kJ mol-1) 18°C	0.6702	0.0012	0.5330	0.0010	0.5344	0.8352	4.60E-06
Hydrothermal Sol. ΔG (kJ mol-1) 100°C	-0.1121	0.6381	-0.0792	0.6256	-0.0794	-0.0357	0.8813
Photoautotrophic	0.7502	1.39E-04	0.6027	2.03E-04	0.6108	0.8024	2.08E-05
Yeast fermentative	0.5300	0.0162	0.3807	0.0189	0.3880	0.4854	0.0300
Yeast respiratory Bragg	0.5030	0.0238	0.3656	0.0242	0.3736	0.6667	0.0013
Yeast fermentative Bragg	0.4664	0.0382	0.3289	0.0426	0.3370	0.4488	0.0472
Li et al Prokaryotic	0.8060	1.79E-05	0.6774	2.97E-05	0.6865	0.8434	3.00E-06
Craig&Weber *E. coli*	0.4414	0.0514	0.3102	0.0559	0.3152	0.6568	0.0017
vs Average Codon %GC							
Number of Enyzmes *E. coli*	-0.2779	0.2355	-0.2526	0.1195	-0.3043	-0.3195	0.1698
Number of Codons	0.3559	0.1236	0.2104	0.1947	0.2846	0.3803	0.0981
Average Codon %GC Weighted*	0.9375	1.11E-09	0.8504	1.58E-07	0.9477	0.9840	6.18E-15
Number of Carbons	-0.4215	0.0642	-0.3365	0.0381	-0.4179	-0.3484	0.1323
Number of Nitrogens	0.0489	0.8377	0.0429	0.7914	0.0789	0.1211	0.6111
vs Number of Codons							
Number of Enyzmes *E. coli*	-0.1029	.6660	-0.0601	0.7109	-0.0758	-0.0982	0.6805

* Based on codon bias of W3110: http://www.kazusa.or.jp/codon/

Table S4: Localization of proteins with extreme average amino acid costs in *E. coli*.

	Overall Number	Overall Percent	Top 100	Top 200	Top 300	Bottom 100	Bottom 200	Bottom 300	Increase fold Top 100	Increase fold Bottom 100
Cytoplasmic	2859	65.9589	36	95	167	19	38	68	0.546	0.288
IM	972	22.4324	12	34	43	78	156	223	0.535	3.477
Periplasmic	337	7.8006	24	37	50	2	3	4	3.077	0.256
OM	149	3.43872	17	22	28	1	3	5	4.944	0.291
Extracellular	16	0.36925	11	12	12	0	0	0	29.789	0
Total	4333		100	200	300	100	200	300		

Table S5A: Statistics comparing ASCs of proteins in different locations of *E. coli*.

	Mann Whitney U		Kolmogorov–Smirnov	
vs Cytoplasm	**Z**	**p value**	**Max Dif**	**p value**
IM	23.196	4.98E-119	-0.418	9.82E-112
Periplasm	6.029	1.65E-09	0.156	6.71E-07
OM	3.713	2.05E-04	0.144	4.82E-03
Extracellular	5.983	2.19E-09	0.674	3.57E-07
vs All				
Cytoplasm	6.223	4.89E-10	0.084	5.46E-11
IM	19.529	6.17E-85	-0.336	5.73E-79
Periplasm	8.242	1.69E-16	0.225	2.26E-14
OM	5.149	2.62E-07	0.174	2.66E-04
Extracellular	6.001	1.96E-09	0.663	5.70E-07
vs IM				
Periplasm	17.023	5.56E-65	0.549	6.61E-67
OM	11.598	4.21E-31	0.489	7.70E-28
Extracellular	6.447	1.14E-10	0.776	2.74E-09
vs OM				
Periplasm	0.146	0.8839	0.107	0.1803
Extracellular	4.868	1.13E-06	-0.567	8.89E-05

Student *t* Test*

vs Cytoplasm	**T value**	**95% CI**	**99.9% CI**	**Deg Free**	**p value**
IM	22.178	1.962	3.298	1314	7.87E-93
Periplasm	6.378	1.966	3.314	414	4.79E-10
OM	3.818	1.975	1.975	159	1.93E-04
Extracellular	8.461	2.131	4.073	15	4.29E-07

* Assumes Unequal Variance. Confidence Intervals are 2 tailed
Mann Whitney U test results used instead of Student *t* test
Most locations fail normality (even taking Square root of data)

Table S5B: Statistics comparing ASCs of proteins in different locations of *S. cerevisiae.*

	Mann Whitney U		Kolmogorov–Smirnov	
vs Cytoplasm	**Z**	**p value**	**Max Dif**	**p value**
Extracellular	3.080	0.0021	-0.271	6.09E-04
Cell Wall	5.780	7.46E-09	-0.432	9.93E-08
Cytoskeleton	5.654	1.56E-08	-0.196	7.05E-07
Bud	3.345	0.0008	-0.211	4.66E-06
Nucleus	3.609	0.0003	-0.056	9.77E-04
PM	9.011	2.04E-19	0.410	1.92E-26
ER	13.623	2.94E-42	0.311	1.33E-39
Golgi	7.736	1.02E-14	0.281	6.95E-11
ER or Golgi Unique**	11.700	1.28E-31	0.338	1.08E-29
Integral/Endo Membranes	16.560	1.35E-61	0.648	7.49E-61
Cell Periphery	2.780	0.0054	0.216	1.07E-08
Trans Vesicles	5.295	1.19E-07	0.241	2.21E-07
Mitochondria	11.708	1.16E-31	0.175	5.03E-21
Peroxisome	5.031	4.88E-07	0.378	5.19E-07
Endosome	0.857	0.3912	0.125	0.3150
Vacoule	9.486	2.41E-21	0.329	3.85E-25
Microsome	2.113	0.0346	0.732	4.09E-03
Lipid Particles	4.712	2.45E-06	0.504	1.15E-06
Punctate Comp.	2.261	0.0237	0.128	0.0221
Ambigious	3.368	0.0008	0.145	1.73E-04
ER or Golgi Only***	13.398	6.25E-41	0.513	1.42E-34
Extracellular or Cell Wall	4.080	4.51E-05	-0.286	1.93E-05
Cyto & Nuc*	1.127	0.2598	-0.017	0.7410
All*	14.862	5.85E-50	0.159	8.70E-45
vs All*				
Extracellular	4.846	1.26E-06	-0.294	0.0001
Cell Wall	6.898	5.27E-12	-0.470	3.57E-09
Cytoskeleton	9.542	1.41E-21	-0.299	4.40E-16
Bud	6.606	3.94E-11	-0.279	1.36E-10
Cytoplasm	14.862	5.85E-50	-0.159	8.70E-45
Nucleus	16.904	4.19E-64	-0.183	6.98E-48
PM	4.390	1.13E-05	0.255	7.21E-11
ER	6.641	3.11E-11	0.156	2.18E-11
Golgi	3.324	0.0009	0.144	0.0030
ER or Golgi Unique**	5.950	2.68E-09	0.182	1.82E-09
Integral/Endo Membranes	12.337	5.71E-35	0.493	1.66E-36
Cell Periphery	1.701	0.0889	-0.126	0.0023
Trans Vesicles	1.031	0.3024	0.104	0.092486
Mitochondria	1.170	0.2420	0.084	4.60E-06
Peroxisome	2.031	0.0423	0.264	0.0011
Endosome	1.529	0.1262	-0.093	0.6795
Vacoule	3.988	6.66E-05	0.181	2.84E-08
Microsome	1.059	0.2897	0.584	0.0387
Lipid Particles	2.566	0.0103	0.363	0.0011
Punctate Comp.	1.559	0.1191	-0.084	0.2713
Ambigious	1.628	0.1035	-0.061	0.3430
ER or Golgi Only***	9.218	3.02E-20	0.357	1.77E-17
Extracellular or Cell Wall	6.026	1.68E-09	-0.320	8.49E-07
Cyto & Nuc*	17.351	1.93E-67	-0.159	1.70E-53

Table S5B: Statistics comparing ASCs of proteins in different locations of *S. cerevisiae* (continued).

vs Cyto & Nuc*	**Z**	**p value**	**Max Dif**	**p value**
Extracellular	2.9331	0.0034	-0.2616	0.0010
Cell Wall	5.6425	1.68E-08	-0.4221	1.94E-07
Cytoskeleton	5.2603	1.44E-07	-0.1812	4.79E-06
Bud	3.0696	0.0021	-0.1976	2.01E-05
Cytoplasm	1.1269	0.2598	0.0169	0.7412
Nucleus	2.7344	0.0062	-0.0388	0.0329
PM	9.2554	2.13E-20	0.4125	4.14E-27
ER	14.2357	5.50E-46	0.3127	9.94E-42
Golgi	8.0091	1.16E-15	0.2868	1.77E-11
ER or Golgi Unique**	12.1017	1.04E-33	0.3390	1.57E-30
Integral/Endo Membranes	16.7279	8.21E-63	0.6476	1.63E-61
Cell Periphery	3.0619	0.0022	0.2190	4.43E-09
Trans Vesicles	5.5631	2.65E-08	0.2486	6.98E-08
Mitochondria	12.7576	2.83E-37	0.1906	1.99E-26
Peroxisome	5.1641	2.42E-07	0.3945	1.27E-07
Endosome	1.0502	0.2936	0.1283	0.2846
Vacoule	9.8263	8.68E-23	0.3340	2.17E-26
Microsome	2.1375	0.0326	0.7372	0.0037
Lipid Particles	4.7690	1.85E-06	0.5112	7.58E-07
Punctate Comp.	2.5557	0.0106	0.1258	0.0255
Ambigious	3.7317	0.0002	0.1513	6.27E-05
ER or Golgi Only***	13.5775	5.45E-42	0.5142	4.10E-35
Extracellular or Cell Wall	3.9124	9.14E-05	-0.2767	3.84E-05
All*	17.3513	1.93E-67	0.1593	1.70E-53

Student *t* Assumes Unequal Variance. Confidence Intervals are 2 tailed

vs Cytoplasm	**T value**	**95% CI**	**99.9% CI**	**Deg Free**	**p value**
Extracellular	3.1178	2.0049	3.4800	54	0.0029
Cell Wall	5.7357	2.0154	3.5258	44	8.24E-07
Cytoskeleton	5.6785	1.9700	3.3320	237	3.96E-08
Bud	2.4981	1.9750	3.3528	159	0.0135
Nucleus	3.6015	1.9605	3.2926	4592	0.0003
PM	7.9820	1.9720	3.3401	199	1.12E-13
ER	12.8933	1.9636	3.3055	654	4.72E-34
Golgi	7.9141	1.9741	3.3490	169	3.13E-13
ER or Golgi Unique**	11.3575	1.9666	3.3178	360	9.33E-26
Integral/Endo Membranes	18.3235	1.9728	3.3436	186	1.59E-43
Cell Periphery	2.0926	1.9702	3.3330	232	0.0375
Trans Vesicles	5.1461	1.9758	3.3561	151	8.16E-07
Mitochondria	12.3166	1.9612	3.2958	1859	1.45E-33
Peroxisome	5.4403	2.0057	3.4838	53	1.38E-06
Endosome	1.2543	2.0017	3.4663	58	0.2148
Vacoule	8.5034	1.9675	3.3218	314	7.60E-16
Microsome	2.8618	2.7764	8.6103	4	0.0458
Lipid Particles	4.9165	2.0555	3.7066	26	4.19E-05
Punctate Comp.	2.3600	1.9762	3.3579	147	0.0196
Ambigious	3.6335	1.9690	3.3279	263	0.0003
ER or Golgi Only***	13.7257	1.9747	3.3516	162	6.17E-29
Extracellular or Cell Wall	4.0728	1.9939	3.4329	71	0.0001
Cyto & Nuc*	1.1471	1.9603	3.2921	6186	0.2514
All*	18.9590	1.9603	3.2918	7445	2.55E-78

* Duplicates removed or none present

** Duplicates, Cytoplasmic, Nuclear removed

*** Only ER, Golgi, or both

Table S5C: Statistics comparing ASC of proteins in different locations in *M tuberculosis*.

	Mann Whitney U		Kolmogorov–Smirnov	
vs Cytoplasm	**Z**	**p value**	**Max Dif**	**p value**
Cytoplasmic Membrane	10.7167	8.49E-27	-0.2290	6.74E-22
Unknown	8.9557	3.37E-19	0.1446	1.93E-13
Extracellular	12.3710	3.75E-35	0.5584	8.58E-37
Cell Wall	2.4495	0.0143	0.4744	0.0365
Extracellular and Cell Wall	12.5622	3.41E-36	0.5546	3.44E-38
vs All				
Cytoplasm	2.0838	0.0372	-0.0488	0.0028
Cytoplasmic Membrane	11.7610	6.20E-32	-0.2285	3.28E-24
Unknown	7.5395	4.72E-14	0.1152	3.32E-10
Extracellular	11.8675	1.75E-32	0.5126	5.74E-32
Cell Wall	2.2392	0.0251	0.4366	0.0670
Extracellular and Cell Wall	12.0425	2.12E-33	0.5093	2.94E-33
vs Cytoplasmic Membrane				
Unknown	14.5359	7.18E-48	0.3347	3.10E-38
Extracellular	13.4196	4.64E-41	0.6219	1.81E-39
Cell Wall	2.9964	0.0027	0.4716	0.0397
Extracellular and Cell Wall	13.6098	3.50E-42	0.6174	1.19E-40

Student *t* Test*

vs Cytoplasm	**T value**	**95% CI**	**99.9% CI**	**Deg Free**	**p value**
Cytoplasmic Membrane	10.5911	1.9631	3.3035	754	1.57E-24
Unknown	8.9146	1.9613	3.2958	1841	1.15E-18
Extracellular	12.2829	1.9768	3.3604	142	4.42E-24
Cell Wall	2.0153	2.3646	5.4079	7	0.0837
Extracellular and Cell Wall	12.3553	1.9759	3.3566	150	1.24E-24

* Assumes Unequal Variance. Confidence Intervals are 2 tailed

Mann Whitney U test results used in text instead of Student *t* test

Most locations fail normality

Table S6: Statistics comparing protein ASCs in different sublocations in *E. coli*.

	Mann Whitney U Test		Kolmogorov–Smirnov	
vs Cytoplasm	**Z**	**p value**	**Max Dif**	**p value**
Integral Membrane Protein	27.6379	3.90E-168	-0.5100	7.80E-140
Membrane anchored	0.6490	0.5164	-0.1178	0.0287
Periplasmic w N-term mem anchor	0.3666	0.7139	-0.3830	0.2696
Inner Membrane Lipoprotein	4.1532	3.28E-05	0.5465	0.0002
Membrane associated	1.7771	0.0756	0.2870	0.3271
Membrane Lipoprotein*	1.1243	0.2609	0.2708	0.6113
OM ß-barrel protein	0.3050	0.7604	-0.1226	0.3914
OM Lipoprotein	4.8920	9.98E-07	0.2395	3.51E-05
Cell surface appendage	5.6057	2.07E-08	0.7376	1.40E-06
Secreted and released	2.2686	0.0233	0.3138	0.7407
Extracellular**	5.9832	2.19E-09	0.6739	3.57E-07
vs All				
Integral Membrane Protein	24.0127	2.05E-127	-0.4287	1.32E-106
Membrane anchored	2.1015	0.0356	0.1275	0.0127
Periplasmic w N-term mem anchor	0.0509	0.9594	-0.3453	0.3899
Inner Membrane Lipoprotein	4.2749	1.91E-05	0.5735	9.75E-05
Membrane associated	2.1025	0.0355	0.3283	0.1873
Membrane Lipoprotein*	1.3172	0.1878	0.2730	0.6005
OM ß-barrel protein	0.7611	0.4466	0.0973	0.6821
OM Lipoprotein	5.8900	3.86E-09	0.2483	1.36E-05
Cell surface appendage	5.5665	2.60E-08	0.7417	1.17E-06
Secreted and released	2.3687	0.0179	0.3570	0.583488
Extracellular**	6.0012	1.96E-09	0.6630	5.70E-07
vs Integral Membrane Protein				
Membrane anchored	10.8649	1.69E-27	0.4402	5.02E-23
Periplasmic w N-term mem anchor	2.6038	0.0092	0.3674	0.3187
Inner Membrane Lipoprotein	6.0709	1.27E-09	0.8172	4.64E-09
Membrane associated	4.6997	2.61E-06	0.7301	1.85E-05
Membrane Lipoprotein*	2.9816	0.0029	0.4839	0.0514
Periplasmic	20.1056	6.59E-90	0.6431	2.86E-86
OM ß-barrel protein	7.3597	1.84E-13	0.5222	1.24E-12
OM Lipoprotein	12.4994	5.22E-37	0.6116	5.52E-29
Cell surface appendage	5.9058	3.51E-09	0.8948	2.19E-09
Secreted and released	3.2553	0.0011	0.6689	0.0296
Extracellular**	6.7169	1.86E-11	0.8564	3.58E-11

Table S6: Statistics comparing protein ASCs in different sublocations in *E. coli* (continued).

vs OM Lipoprotein	**Z**	**p value**	**Max Dif**	**p value**
Periplasmic	1.5790	0.1143	-0.1245	0.1832
Inner Membrane Lipoprotein	1.6951	0.0901	0.3467	0.0817
Membrane Lipoprotein*	0.1966	0.8442	-0.3125	0.4666
vs Periplasmic				
Periplasmic w N-term mem anchor	1.1713	0.2415	-0.4956	0.0735
Inner Membrane Lipoprotein	2.8413	0.0045	0.4298	0.0092
Membrane Lipoprotein*	0.2054	0.8372	-0.2861	0.5512

Student *t* test

with Unequal Variance. CI 2 tailed

vs Cytoplasm	**T value**	**95% CI**	**99.9% CI**	**Deg Free**	**p value**
Integral Membrane Protein	28.8969	1.9622	3.2996	1072	3.01E-136
Membrane anchored	0.3593	1.9745	3.3508	164	0.7198
Periplasmic w N-term mem anchor	0.3731	2.5706	6.8688	5	0.7244
Inner Membrane Lipoprotein	5.6518	2.1604	4.2208	13	7.91E-05
Membrane associated	2.6739	2.2622	4.7809	9	0.0255
Membrane Lipoprotein*	0.5475	2.4469	5.9588	6	0.6038
OM ß-barrel protein	0.2808	2.0057	3.4838	53	0.7799
OM Lipoprotein	5.2490	1.9840	3.3905	100	8.61E-07
Cell surface appendage	9.6144	2.2010	4.4370	11	1.09E-06
Secreted and released	2.2569	3.1824	12.9240	3	0.1092
Extracellular**	8.4609	2.1315	4.0728	15	4.29E-07
vs OM Lipoprotein					
Periplasmic	1.8387	1.9767	3.3599	143	0.0680
Inner Membrane Lipoprotein	2.1090	2.0639	3.7454	24	0.0456
Membrane Lipoprotein*	0.6253	2.3646	5.4079	7	0.5516
vs Periplasmic					
Periplasmic w N-term mem anchor	1.4189	2.5706	6.8688	5	0.2152
Inner Membrane Lipoprotein	3.5501	2.1315	4.0728	15	0.0029
Membrane Lipoprotein*	0.1825	2.4469	5.9588	6	0.8612

*Difficulty determining which membrane for seven “Membrane Lipoproteins”: VisC, YfbK, YiaF, YifL, YihN, YjbH, and YraM. These proteins were included in “Inner Membrane” for Fig. 1A, 3B, S1,S2, S5, S6, and Tables S3-6.

**Subdivided into “Secreted and released” and “Cell surface appendage”

Table S7: Amino acid percent of proteins in different compartments of *E. coli*.

	Mean ATP per aa	%A	%R	%N	%D	%C	%E	%Q	%G	%H	%I	%L	%K	%M	%F	%P	%S	%T	%W	%Y	%V	Total Amino Acids	Number
Overall Mean	23.555	9.37	5.60	3.87	5.03	1.29	5.83	4.42	7.12	2.33	6.10	10.59	4.70	2.96	3.91	4.30	5.81	5.36	1.52	2.79	7.09	312.46	4333
Overall Median	23.398	9.23	5.45	3.68	5.15	1.08	5.90	4.24	7.06	2.17	5.87	10.37	4.27	2.78	3.62	4.24	5.61	5.18	1.32	2.61	6.99	275	4333
Overall STDEVP	1.5038	2.90	2.43	1.79	2.04	1.24	2.59	2.06	2.50	1.50	2.27	3.20	2.44	1.33	2.06	1.75	2.07	1.95	1.20	1.47	2.28	209.216	4333
Extracellular*	20.652	**12.11**	**2.53**	**6.91**	5.81	**0.57**	**3.03**	**6.47**	8.04	**0.70**	4.83	8.45	4.43	**1.71**	**2.56**	**2.74**	**9.11**	**9.86**	**0.63**	2.42	7.09	280.438	16
OM	22.852	9.04	4.59	**5.53**	5.66	1.14	**4.16**	5.16	8.20	**1.44**	4.97	8.64	4.88	2.79	3.23	4.16	**7.40**	**6.81**	1.51	**3.99**	6.69	334.946	149
Periplasm	22.854	10.10	4.24	**5.05**	5.75	1.07	4.69	5.10	7.02	**1.71**	5.06	8.84	**6.15**	2.64	3.44	4.58	6.65	6.44	1.44	3.09	6.96	309.418	337
IM	24.653	9.93	4.56	3.22	**3.29**	0.99	**3.61**	3.54	7.80	**1.67**	7.48	13.18	3.56	3.53	**5.32**	4.13	6.04	5.25	**2.16**	2.72	8.04	357.546	972
Cytoplasm	23.317	9.10	6.18	3.86	5.50	1.44	6.82	4.59	6.84	2.68	5.82	10.03	4.90	2.82	3.54	4.34	5.53	5.17	1.32	2.72	6.80	296.498	2859
Sublocations																							
Cell surface appendage	20.474	**12.23**	**2.21**	**7.95**	5.26	**0.49**	**2.28**	**6.33**	**9.25**	**0.68**	4.91	8.20	**3.49**	**1.62**	**2.72**	**2.18**	**9.35**	**10.50**	**0.56**	2.67	7.11	299.167	12
Secreted and released	21.187	**11.72**	**3.52**	3.78	**7.44**	**0.78**	5.27	**6.91**	**4.41**	**0.75**	4.61	9.19	**7.24**	**1.98**	**2.07**	4.43	**8.40**	**7.93**	**0.85**	**1.69**	7.01	**224.25**	4
OM Lipoprotein	22.564	9.58	4.63	4.78	5.07	1.58	**4.09**	**5.61**	7.47	**1.41**	5.20	9.26	5.30	3.26	**2.75**	4.62	**7.33**	6.65	1.20	3.08	7.11	**216.46**	96
OM ß-barrel protein	23.374	8.05	4.52	**6.88**	**6.72**	**0.36**	**4.27**	4.34	**9.52**	**1.49**	**4.55**	**7.53**	4.12	**1.95**	4.09	3.31	**7.53**	**7.12**	**2.08**	**5.64**	5.94	**549.57**	53
Integral Membrane Protein	25.01	9.98	4.29	3.00	**2.84**	0.99	**3.14**	**3.11**	8.04	**1.64**	**7.94**	**13.83**	**3.11**	**3.72**	**5.78**	4.05	6.09	5.19	**2.31**	2.74	8.21	369.584	777
Membrane anchored	23.374	9.69	5.90	3.98	4.95	**0.94**	5.45	5.09	6.77	1.79	5.88	10.79	5.23	2.82	3.56	4.28	5.81	5.49	1.67	2.52	7.39	284.576	158
Peri. w N-term anchor	23.48	9.06	5.22	4.06	5.74	**1.83**	6.53	5.48	7.00	**3.19**	5.23	8.29	**6.33**	3.21	2.95	**5.78**	5.05	4.20	1.73	3.07	6.06	271.333	6
IM Lipoprotein	21.987	9.47	**4.04**	4.76	**6.37**	**0.83**	6.14	**6.03**	7.11	**1.21**	**4.44**	9.65	**6.61**	**1.91**	**2.57**	4.38	6.44	6.32	**0.79**	3.09	7.85	**485.5**	14
Membrane associated	22.708	11.10	5.65	4.17	5.03	1.02	6.13	5.22	7.08	2.19	4.92	9.79	5.17	3.25	**2.68**	4.60	5.65	5.06	1.46	2.46	7.36	390.5	10
Membrane Lipoprotein	22.97	9.77	**3.91**	4.62	5.27	**0.92**	**3.66**	**6.78**	7.54	**1.15**	**4.29**	10.19	5.31	2.40	4.46	**5.94**	6.24	6.28	1.27	3.35	6.64	**439.29**	7
Cost Savings (ATP)		15.66	0.06	12.66	14.66	2.66	12.06	11.06	15.66	-10.9	-4.94	0.06	-2.94	-6.94	-24.6	7.06	15.66	8.66	-46.9	-22.6	4.06		

Notes: **Bold** indicates >25% change in amino acid composition from the overall mean.
Underlined indicates >50% change in amino acid composition from the overall mean.

Amino acid percents are on a per protein basis. The percent amino acid composition of each protein in that location were averaged; they are unweighted for chain length. See Table S22A for weighted percents (% of all amino acids) of proteins in nonmembrane locations of *E. coli*.

* We also looked at Extracellular proteins cost removing the SEC signal peptide when present (except FlgD for which a SEC signal is predicted; however, FlgD is secreted by the flagella system(76)). The mean changed to 20.433 ATPs per amino acid.

Table S8: Percent of amino acids in curli specific genes.

Standard Competition Ranking

	% Gly	Rank Gly	% Ser	Rank Ser	%GS	Rank GS	% Asn	Rank Asn	% Gln	Rank Gln	%NQ	Rank NQ	%GSNQ	Rank GSNQ
Mean	7.12		5.81		12.93		3.87		4.42		8.29		21.22	
Median	7.06		5.61		12.84		3.68		4.24		8.16		21.01	
STDEVP	2.50		2.07		3.09		1.79		2.06		2.77		3.84	
CsgA	**19.21**	6	**9.27***	227	**28.48**	6	**10.60**	**22**	7.28**	354	**17.89**	**18**	**46.36**	**2**
CsgB	9.93	492	8.61	361	**18.54**	165	**8.61**	**87**	**9.93**	**55**	**18.54**	**12**	**37.08**	**8**
CsgC	3.64	4028	**14.55**	12	**18.18**	198	3.64	2218	**10.91**	**33**	**14.55**	**102**	**32.73**	**47**
CsgD	4.17	3884	8.33	420	12.50	2377	**8.80**	**75**	4.17	2214	12.96	223	25.46	438
CsgE	4.65	3693	5.43	2358	10.08	3669	4.65	1160	3.88	2483	8.53	1907	18.61	3415
CsgF	8.70	1077	7.97	536	16.67	419	**11.59**	**13**	**7.97**	**218**	**19.57**	**7**	**36.23**	**13**
CsgG	6.86	2323	7.22	846	14.08	1405	5.78	531	5.78	940	11.55	460	25.63	411
$>P_{95}$	11.11		9.33		18.04		7.14		7.97		13.03		27.78	
$>P_{90}$	10.11		8.33		16.60		6.06		7.01		11.64		25.49	

* 94.7th Percentile

** 91.8th Percentile

Note: Bold ≥95th percentile

Fractional Ranking

	Rank G	Rank S	Rank GS	Rank N	Rank Q	Rank NQ	Rank GSNQ
CsgA	6	227	6	22	354	18	2
CsgB	492	361	165	87	55	12	8
CsgC	4029.5(4)	12	201.5(8)	2223(11)	34(3)	102	47
CsgD	3889.5(12)	427.5(16)	2392(31)	75.5(2)	2224(21)	224(3)	438
CsgE	3694.5(4)	2359.5(4)	3669.5(2)	1168.5(18)	2484.5(4)	1909.5(6)	3417.5(6)
CsgF	1080.5(8)	536	423.5(10)	13	218	7	13
CsgG	2323	847(3)	1405.5(2)	531	940	460	411

Note: Numbers in parenthesis are number of proteins with same % amino acid

Table S9: Percentiles of flagella, curli, and Type 1 pili proteins.

Percentile	99.9999	99.999	99.99	99.9	99	95	90	80	70
Avg E	18.0171	18.0362	18.2268	18.9156	20.4295	21.4267	21.8936	22.4035	22.7847
Savings	5.5378	5.5187	5.3281	4.6393	3.1254	2.1282	1.6613	1.1514	0.7703

Flagella

Percentile	Avg Cost	Rank	Name	Length	Location	Role	Savings per aa
99.885	19.1548	6	FliC	498	Cell surface appendage	Structural	4.4001
99.793	19.3949	10	FlgM	97	Secreted&released(40, 48)	Regulatory	4.1601
99.677	19.6541	15	FlgD	231	Cell surface appendage	Assembly	3.9008
99.608	19.7397	18	FliK	375	Secreted&released(61)	Assembly	3.8152
99.539	19.8705	21	FlgF	251	Periplasmic	Structural	3.6844
99.493	19.9291	23	FliD	468	Cell surface appendage	Structural	3.6258
99.308	20.1981	31	FlgI	365	Periplasmic	Structural	3.3568
98.985	20.4642	45	FlgC	134	Periplasmic	Structural	3.0907
98.962	20.4713	46	FlgK	547	Cell surface appendage	Structural	3.0836
98.846	20.5953	51	FlgE	402	Cell surface appendage	Structural	2.9596
98.639	20.6565	60	FlgG	260	Periplasmic	Structural	2.8984
98.5	20.6683	66	FliE	104	Cytoplasmic *	Structural	2.8866
98.015	20.8473	87	FlgL	317	Cell surface appendage	Structural	2.7076
97.831	20.9435	95	FlgB	138	Periplasmic	Structural	2.6114
96.03	21.2669	173	FliO	121	Membrane anchored	Export	2.2880
95.868	21.2995	180	FliF	552	IM Protein	Structural	2.2555
95.476	21.3480	197	FlgA	219	Periplasmic	Chaperone	2.2070
95.407	21.3540	200	FliN	137	Cytoplasmic	Structural	2.2009
95.361	21.3680	202	FliH	228	Cytoplasmic	Export	2.1869
94.368	21.5162	245	FliS	136	Cytoplasmic	Chaperone	2.0387
93.883	21.5544	266	FliG	331	Membrane associated	Structural	2.0005
93.583	21.5961	279	FlgH	232	OM Lipoprotein	Structural	1.9588
87.166	22.0514	557	FliI	457	Cytoplasmic	Export	1.5035
83.796	22.2189	703	FlgJ	313	Periplasmic(67)	Assembly	1.3361
82.018	22.3130	780	MotB	308	Membrane anchored	Structural	1.2419
81.072	22.3541	821	FliY	266	Periplasmic	?	1.2008
79.225	22.4391	901	FlgN	138	Cytoplasmic	Chaperone	1.1158
78.048	22.4958	952	FliA	239	Cytoplasmic	Regulatory	1.0591
70.868	22.7515	1263	FlhE	130	Periplasmic	?	0.8034
66.898	22.8840	1435	FlhA	692	IM Protein	Export	0.6709
66.205	22.9020	1465	FliJ	147	Cytoplasmic	Export	0.6529
59.65	23.0984	1749	FliT	121	Cytoplasmic	Chaperone	0.4566
59.28	23.1126	1765	FliM	334	Cytoplasmic	Structural	0.4423
57.11	23.1824	1859	MotA	295	IM Protein	Structural	0.3725
53.948	23.2740	1996	FlhD	119	Cytoplasmic	Regulatory	0.2810
45.568	23.5490	2359	FlhC	192	Cytoplasmic	Regulatory	0.0060
41.736	23.6766	2525	FliL	154	OM Lipoprotein	?	-0.1217
40.859	23.7039	2563	FlhB	382	IM Protein	Export	-0.1490
17.544	24.7843	3573	FliQ	89	IM Protein	Export	-1.2294
17.314	24.8102	3583	FliP	245	IM Protein	Export	-1.2553
13.805	25.0825	3735	FliZ	183	Cytoplasmic	Regulatory	-1.5276
11.866	25.2406	3819	FliR	261	IM Protein	Export	-1.6857

77.64364	21.90035	969.5		266.8571		**AVERAGE**	1.6546
93.733	21.57525	272.5		242		**MEDIAN**	1.5876
27.17407	1.598852	1177.18		143.2881		**STDEV**	1.5989

Table S9: Percentiles of flagella, curli, and Type 1 pili proteins (continued).

Curli

Percentile	Avg Cost	Rank	Name	Length	Location	Role	Savings per aa
99.816	19.3887	9	CsgA	151	Cell surface appendage	Major	4.1662
98.685	20.6517	58	CsgB	151	Cell surface appendage	Minor	2.9033
96.538	21.2036	151	CsgC	110	Periplasmic	?	2.3513
24.146	24.3593	3287	CsgD	216	Cytoplasmic	Regulatory	-0.8044
11.981	25.2326	3814	CsgE	129	Periplasmic	Chaperone	-1.6776
88.874	21.9725	483	CsgF	138	Periplasmic**	Chaperone	1.5824
68.006	22.8498	1387	CsgG	277	OM Lipoprotein	export	0.7051
69.72086	22.23688	1312.714		167.4286		**AVERAGE**	1.3180
88.874	21.97246	483		151		**MEDIAN**	1.5824
37.06872	2.066261	1605.809		58.5174		**STDEV**	2.0663

Type 1 pili

Percentile	Avg Cost	Rank	Name	Length	Location	Role	Savings per aa
99.77	19.5247	11	FimA	182	Cell surface appendage	Major	4.0302
10.942	25.3265	3859	FimB	200	Cytoplasmic	regulatory	-1.7716
70.291	22.7739	1288	FimC	241	Periplasmic	Chaperone	0.7811
72.115	22.7075	1209	FimD	878	OM ß-barrel protein	export	0.8474
30.772	24.0742	3000	FimE	198	Cytoplasmic	Regulatory	-0.5193
84.419	22.1852	676	FimF	176	Cell surface appendage	Adaptor	1.3697
97.554	20.9916	107	FimG	167	Cell surface appendage	Adaptor	2.5633
82.503	22.2920	759	FimH	300	Cell surface appendage	Tip	1.2629
57.941	23.1536	1823	FimI	179	Periplasmic	?	0.4013
29.156	24.1405	3070	FimZ	210	Cytoplasmic	Regulatory	-0.5856
63.5463	22.71698	1580.2		273.1		**AVERAGE**	0.7487
71.203	22.74069	1248.5		199		**MEDIAN**	0.8321
30.61971	1.650011	1326.442		216.1437		**STDEV**	1.6500

*FliE interacts with FlgB and is likely periplasmic (55, 62).

**CsgF has been shown to be cell surface exposed (73).

Table S10: Cost savings and amino acid composition or extracellular proteins in *E. coli*.

	Cost vs Average aa		15.66	0.06	12.66	14.66	2.66	12.06	11.06	15.66	-10.9	-4.94	0.06	-2.94	-6.94	-24.6	7.06	15.66	8.66	-46.9	-22.6	4.06
		ATP/aa	**%A**	**%R**	**%N**	**%D**	**%C**	**%E**	**%Q**	**%G**	**%H**	**%I**	**%L**	**%K**	**%M**	**%F**	**%P**	**%S**	**%T**	**%W**	**%Y**	**%V**
Overall	**MEAN**	23.555	9.37	5.60	3.87	5.03	1.29	5.83	4.42	7.12	2.33	6.10	10.59	4.70	2.96	3.91	4.30	5.81	5.36	1.52	2.79	7.09
	MEDIAN	23.398	9.23	5.45	3.68	5.15	1.08	5.90	4.24	7.06	2.17	5.87	10.37	4.27	2.78	3.62	4.24	5.61	5.18	1.32	2.61	6.99
	STDEVP	1.504	2.90	2.43	1.79	2.04	1.24	2.59	2.06	2.50	1.50	2.27	3.20	2.44	1.33	2.06	1.75	2.07	1.95	1.20	1.47	2.28
Extracellular	**MEAN**	20.652	12.11	2.535	6.911	5.805	0.566	3.0305	6.474	8.0422	0.698	4.832	8.45	4.427	1.714	2.558	2.744	9.112	9.857	0.629	2.422	7.087

		ATP/aa	**A**	**R**	**N**	**D**	**C**	**E**	**Q**	**G**	**H**	**I**	**L**	**K**	**M**	**F**	**P**	**S**	**T**	**W**	**Y**	**V**
ATP savings per amino Acid	**FliC**	19.155	11.85	2.21	9.64	7.83	0	2.81	5.42	8.84	0	5.62	7.43	5.02	0.80	1.00	1.20	8.63	13.05	0	2.01	6.63
	CsgA	19.389	10.60	1.32	10.60	5.30	0	1.32	7.28	19.21	1.99	3.31	5.96	2.65	1.32	2.65	1.32	9.27	5.96	0.66	2.65	6.62
	FimA	19.525	19.78	1.65	6.04	4.40	1.10	1.65	5.49	8.79	1.10	3.30	8.24	2.75	0.55	3.85	1.10	7.69	12.09	0	1.10	9.34
	FlgM	19.395	11.34	5.15	5.15	7.22	0	5.15	6.19	3.09	0	5.15	9.28	6.19	3.09	0	5.15	11.34	11.34	0	0	5.15
	FlgD	19.654	7.79	1.30	6.49	4.76	0	3.46	6.49	10.82	0.87	6.06	9.52	1.73	1.30	1.73	3.03	10.39	15.58	0.43	0.87	7.36
	FliK	19.740	12.53	1.87	2.40	6.93	0	5.07	9.60	4.53	1.87	3.73	11.20	3.20	1.07	1.87	8.80	8.27	10.67	0.27	0	6.13
	FliD	19.929	11.54	1.50	6.62	7.26	0	1.50	4.91	6.41	0.21	6.20	9.62	5.56	1.28	1.71	0.64	14.32	14.10	0.43	2.14	4.06
	FlgK	20.471	12.07	2.93	8.59	6.76	0	2.93	8.04	7.86	0.18	4.94	8.59	3.29	1.65	3.29	1.65	8.78	8.78	0.37	2.74	6.58
	FlgE	20.595	10.20	1.49	11.94	5.22	0	1.99	6.22	9.70	0.25	4.23	7.46	2.99	2.24	4.23	2.74	8.71	10.70	0.50	2.49	6.72
	CsgB	20.652	15.23	3.31	8.61	3.31	0	1.99	9.93	9.93	0.00	7.95	6.62	5.30	3.31	1.99	0.66	8.61	5.30	0	3.97	3.97
	FlgL	20.847	11.67	3.47	5.36	6.94	0	5.05	8.52	5.99	0.32	4.10	8.83	4.73	3.47	2.52	1.58	10.41	8.20	0.63	2.84	5.36
	FimG	20.992	13.77	1.80	4.79	5.39	1.80	1.20	5.99	7.78	1.20	4.19	7.78	3.59	1.20	1.80	1.80	9.58	12.57	1.20	3.59	8.98

		Savings/aa	**A**	**R**	**N**	**D**	**C**	**E**	**Q**	**G**	**H**	**I**	**L**	**K**	**M**	**F**	**P**	**S**	**T**	**W**	**Y**	**V**
Percent of total savings for each amino Acid	**FliC**	4.40	8.80	-0.05	**16.59**	9.33	-0.78	-8.26	2.52	6.10	5.79	0.53	-0.04	-0.22	3.40	**16.30**	-4.97	**10.06**	**15.14**	**16.22**	4.02	-0.43
	CsgA	4.17	4.59	-0.06	**20.43**	0.94	-0.83	-13.03	7.61	**45.42**	0.90	3.30	-0.07	1.45	2.72	7.48	-5.04	**13.02**	1.25	9.67	0.76	-0.45
	FimA	4.03	**40.43**	-0.06	6.82	-2.31	-0.13	-12.50	2.95	6.49	3.34	3.43	-0.03	1.42	4.15	0.42	-5.61	7.32	**14.46**	**17.71**	9.49	2.27
	FlgM	4.16	7.40	-0.01	3.90	7.70	-0.83	-1.95	4.70	-15.17	6.13	1.12	-0.02	-1.05	-0.23	**23.18**	1.45	**20.82**	**12.45**	**17.16**	15.18	-1.89
	FlgD	3.90	-6.35	-0.07	8.50	-1.01	-0.88	-7.31	5.88	14.86	4.10	0.05	-0.02	2.24	2.95	**13.79**	-2.30	**18.39**	**22.70**	**13.09**	11.16	0.28
	FliK	3.82	**12.96**	-0.06	-4.89	7.31	-0.90	-2.40	**15.02**	-10.63	1.33	3.06	0.01	1.15	3.44	**13.22**	8.33	**10.09**	12.04	**15.43**	16.55	-1.02
	FliD	3.63	9.34	-0.07	9.60	9.03	-0.95	-14.41	1.51	-3.07	6.38	-0.13	-0.02	-0.70	3.21	**14.98**	-7.13	**36.75**	**20.88**	**14.16**	4.07	-3.39
	FlgK	3.08	**13.67**	-0.05	**19.37**	8.24	-1.12	-11.35	**13.00**	3.75	7.62	1.86	-0.04	1.34	2.95	4.98	-6.08	**15.07**	9.59	**17.59**	0.34	-0.67
	FlgE	2.96	4.36	-0.08	**34.50**	0.95	-1.16	-15.63	6.73	**13.65**	7.69	3.12	-0.06	1.70	1.69	-2.62	-3.73	**15.34**	15.61	**16.23**	2.30	-0.51
	CsgB	2.90	**31.59**	-0.05	**20.65**	-8.69	-1.19	-15.95	**21.01**	**15.17**	8.78	-3.15	-0.08	-0.61	-0.85	**16.36**	-8.85	**15.11**	-0.19	**24.59**	-9.24	-4.36
	FlgL	2.71	**13.29**	-0.05	6.96	**10.33**	-1.27	-3.47	**16.74**	-6.53	8.14	3.64	-0.04	-0.04	-1.31	**12.65**	-7.10	**26.62**	9.09	**15.43**	-0.42	-2.59
	FimG	2.56	**26.87**	-0.09	4.53	2.04	0.52	-21.78	6.77	4.05	4.83	3.67	-0.07	1.27	4.76	**20.36**	-6.90	**23.05**	**24.37**	5.92	-7.10	3.00
	Average	>1.5ATPs	**13.91**	-0.06	**12.25**	3.65	-0.79	-10.67	8.70	6.17	5.42	1.71	-0.04	0.66	2.24	**11.76**	-3.99	**17.64**	**13.12**	**15.27**	3.92	-0.81
		ALL	**17.60**	-0.09	**14.79**	6.47	-0.57	-11.87	8.58	-1.23	9.30	2.78	-0.06	-1.79	4.54	**10.88**	-5.89	**19.54**	**14.10**	**15.21**	-3.30	1.10

Bold: Greater than 10% of total savings due to composition of that amino acid

Percent of Total Cost Savings for *E. coli* Major Extracellular Fiber Subunits

	Excluding Aromatics				**Aromatics**		
	Top 2	**Savings**	**>10% Savings**	**Savings**	**FW**	**FYW**	**FYWH**
CsgA	GN	65.85	GNS	78.86	17.16	17.91	18.81
FimA	AT	54.89	AT	54.89	18.13	27.62	30.96
FliC	NT	31.72	NTS	41.78	32.52	36.54	42.33

Table S11: Flagella protein location vs average synthetic cost (ASC).

	Student *t* test (assuming heteroscedastic)						Two tailed		One tailed	
	Num.	Mean	Variance	Degrees Freedom	T crit. value	Pooled Variance	P value	Critical Value*	P value	Critical Value*
vs Cytoplasm	**13**	**22.532**	**1.364**							
IM	10	23.104	2.264	17	0.994	1.750	0.33405	2.110	0.16703	1.740
Periplasm	9	21.201	1.059	19	2.821	1.242	**0.01091**	2.093	**0.00546**	1.729
OM	2	22.636	2.164	1	0.096	1.426	0.93899	12.706	0.46949	6.314
Extracellular	8	19.973	0.366	19	6.590	0.997	**2.624E-06**	2.093	**1.312E-06**	1.729
vs Cyto & IM	**23**	**22.780**	**1.754**							
Cytoplasm	13	22.532	1.364	28	0.584	1.617	0.56359	2.048	0.28179	1.701
IM	10	23.104	2.264	15	0.588	1.902	0.56529	2.131	0.28265	1.753
Periplasm	9	21.201	1.059	19	3.587	1.569	**0.00196**	2.093	**0.00098**	1.729
OM	2	22.636	2.164	1	0.134	1.772	0.91528	12.706	0.45764	6.314
Extracellular	8	19.973	0.366	26	8.036	1.419	**1.626E-08**	2.056	**8.128E-09**	1.706
vs all flagella	**42**	**21.900**	**2.556**							
Cytoplasm	13	22.532	1.364	27	1.550	2.286	0.13273	2.052	0.06636	1.703
IM	10	23.104	2.264	14	2.246	2.504	**0.04139**	2.145	**0.02069**	1.761
Periplasm	9	21.201	1.059	18	1.656	2.312	0.11504	2.101	0.05752	1.734
OM	2	22.636	2.164	1	0.688	2.547	0.61616	12.706	0.30808	6.314
Extracellular	8	19.973	0.366	29	5.902	2.237	**2.085E-06**	2.045	**1.043E-06**	1.699

BOLD = Statistically Significant. **P <0.05**

* 5% Confidence Interval

Table S12: Comparing flagella ASC/economic rank with subunit stoichiometry or chain length.

With FliC included (26 total)

Comparing Avg AA Cost & Number of Subunits
Linear regression
$y = 17322 - 761.2x$ R= 0.31913

Spearman R	-0.25497
Rank Difference Squares Sum	3662
t-test value for hypothesis r = 0	-1.29179
p-level	0.208733
Kendall Tau	-0.12521
Inversions Count	177
Z	0.896972
p-level	0.369734
Gamma	-0.12739
Pearson Correlation Coefficient	-0.31913

Comparing Economic Rank & Number of Subunits
Linear regression
$y = 1178.2 - 0.4517x$ R= 0.14222

Spearman R	-0.25497
Rank Difference Squares Sum	3662
t-test value for hypothesis r = 0	-1.29179
p-level	0.208733
Kendall Tau	-0.12521
Inversions Count	177
Z	0.896972
p-level	0.369734
Gamma	-0.12739
Pearson Correlation Coefficient	-0.14222

Without FliC (25 total)

Comparing Avg AA Cost & Number of Subunits
Linear regression
$y = 84.851 - 2.8219x$ R= 0.13516

Spearman R	-0.16159
Rank Difference Squares Sum	3012
t-test value for hypothesis r = 0	-0.78528
p-level	0.440307
Kendall Tau	-0.05094
Inversions Count	152
Z	0.356929
p-level	0.721145
Gamma	-0.0519
Pearson Correlation Coefficient	-0.13516

Comparing Economic Rank & Number of Subunits
Linear regression
$y = 28.262 - 0.0055992x$ R= 0.21034

Spearman R	-0.16159
Rank Difference Squares Sum	3012
t-test value for hypothesis r = 0	-0.78528
p-level	0.440307
Kendall Tau	-0.05094
Inversions Count	152
Z	0.356929
p-level	0.721145
Gamma	-0.0519
Pearson Correlation Coefficient	-0.21034

Comparing Avg AA cost & Chain Length (all 42 proteins)
Linear regression

Pearson Correlation Coefficient	-0.21355
Spearman R	-0.18857
Rank Difference Squares Sum	14667
t-test value for hypothesis r = 0	-1.21443
p-level	0.231703
Kendall Tau	-0.12674
Inversions Count	484
Z	1.18265
p-level	0.236948
Gamma	-0.12689
Pearson Correlation Coefficient	-0.21355

Table S13: Comparison of protein abundance and average synthetic cost (ASC) or mean amino acid mass in *E. coli*.

						Number per cell vs. Avg Cost				Number per cell vs. Avg Mass			
						Spearman		Pearson		Spearman		Pearson	
	Location	Proteins	Avg Number per cell	Avg Cost	Avg Mass	R	P value	R	P value	R	P value	R	P value
Whole Cell Lysates **Masuda et al 2009(57)**	Overall	1270	1362.67	22.98	110.58	**-0.1263**	**6.31E-06**	**-0.0717**	**0.0105**	**-0.0750**	**0.0075**	-0.0512	0.0684
	Overall without Ribosomal Proteins	1217	529.21	22.98	110.55	**-0.1349**	**2.33E-06**	**-0.0691**	**0.0159**	**-0.0995**	**0.0005**	**-0.0578**	**0.0437**
	Cytoplasmic	1022	1625.28	22.95	110.81	**-0.1078**	**0.0006**	**-0.0902**	**0.0039**	**-0.0773**	**0.0135**	**-0.0632**	**0.0435**
	Cytoplasmic without Ribosomal Proteins	969	592.86	22.95	110.78	**-0.1187**	**0.0002**	**-0.0978**	**0.0023**	**-0.1019**	**0.0015**	**-0.0709**	**0.0273**
	Inner membrane	134	348.65	23.51	109.80	0.0038	0.9656	0.1262	0.1461	-0.1584	0.0676	0.0066	0.9392
	Integral Membrane Protein	81	371.32	24.27	109.84	0.1829	0.1022	**0.2394**	**0.0313**	-0.1156	0.3042	**0.0061**	**0.9570**
	Inner membrane lipoprotein	3	162.75	21.25	107.65	-	-	-0.8752	0.3215	-	-	-0.8093	0.3997
	Membrane anchored	43	337.14	22.44	110.14	-0.0854	0.5860	-0.0880	0.5747	-0.1093	0.4854	0.0015	0.9926
	Membrane associated	3	213.32	22.65	108.07	-	-	0.5438	0.6340	-	-	-0.0805	0.9487
	Membrane Lipoprotein	3	328.33	21.77	107.64	-	-	0.9649	0.1691	-	-	0.4960	0.6696
	Periplasmic w N-term memb anchor	1	32.03	22.16	111.15	-	-	-	-	-	-	-	-
	Periplasmic	63	147.88	22.76	109.65	**-0.3890**	**0.0016**	**-0.3009**	**0.0166**	**-0.4100**	**0.0008**	**-0.2913**	**0.0205**
	Outer membrane	51	265.22	22.63	109.27	-0.2168	0.1264	0.0290	0.8397	-0.2627	0.0626	-0.0437	0.7609
	Outer membrane Lipoprotein	31	140.62	22.11	108.35	**-0.5761**	**0.0007**	**-0.4159**	**0.0200**	**-0.4994**	**0.0042**	**-0.3591**	**0.0472**
	Outer Membrane ß-barrel protein	20	458.36	23.43	110.68	0.1579	0.5061	-0.0841	0.7244	-0.0692	0.7720	-0.2258	0.3385
Membrane Enriched Fraction **Masuda et al 2009(57)**	Overall	961	1385.17	23.18	110.53	**-0.2868**	**1.08E-19**	**-0.0741**	**0.0216**	-0.0196	0.5440	-0.0506	0.1171
	Cytoplasmic	579	2265.89	22.95	110.87	**-0.1250**	**0.0026**	**-0.1035**	**0.0127**	**-0.1381**	**0.0009**	-0.0793	0.0566
	Inner membrane	294	7.47	23.80	110.22	-0.1115	0.0563	**-0.1212**	**0.0378**	-0.2400	3.20E-05	-0.0904	0.1221
	Integral Membrane Protein	208	1.54	24.30	110.23	**0.1511**	**0.0293**	**0.1558**	**0.0246**	**-0.2283**	**0.0009**	**-0.1823**	**0.0084**
	Inner membrane lipoprotein	3	162.75	21.25	107.65	-	-	-0.8752	0.3215	-	-	-0.8093	0.3997
	Membrane anchored	74	9.31	22.67	110.44	-0.1596	0.1744	0.0930	0.4305	-0.1601	0.1731	-0.0227	0.8478
	Membrane associated	3	213.32	22.65	108.07	-	-	0.5438	0.6340	-	-	-0.0805	0.9487
	Membrane Lipoprotein	3	17.67	21.77	107.64	-	-	**-0.9998**	**0.0138**	-	-	-0.2296	0.8525
	Periplasmic w N-term memb anchor	3	1.58	23.12	111.58	-	-	0.7273	0.4816	-	-	0.9170	0.2612
	Periplasmic	33	174.87	22.66	109.78	**-0.4170**	**0.0158**	**-0.3849**	**0.0270**	**-0.3652**	**0.0367**	**-0.3454**	**0.0490**
	Outer membrane	54	207.98	22.57	109.14	0.0094	0.9461	0.0632	0.6496	0.0271	0.8457	0.0328	0.8136
	Outer membrane Lipoprotein	33	96.59	22.10	108.42	-0.2361	0.1860	**-0.3545**	**0.0429**	-0.1899	0.2898	-0.2224	0.2135
	Outer Membrane ß-barrel protein	21	383.00	23.29	110.26	0.2519	0.2706	-0.0215	0.9262	0.2286	0.3190	-0.0383	0.8689
	Cell Surface Appendage	1	1.45	19.52	99.41	-	-	-	-	-	-	-	-

BOLD: $P < 0.05$

Notes: Echobase sequence and location data were merged with abundance values of Masuda et al (57) using ECK numbers. GatA had two abundance values in whole cell lysates. InsH and yi81 have multiple Echobase number but only one ECK number; so ASC and Mass values were averaged.

Table S13: Comparison of protein abundance and average synthetic cost (ASC) or mean amino acid mass in *E. coli* (continued)

	Location	Proteins	Avg Number per cell	Avg Cost	Avg Mass	Number per cell vs. Avg Cost				Number per cell vs. Avg Mass			
						Spearman		Pearson		Spearman		Pearson	
						R	P value	R	P value	R	P value	R	P value
Cytosolic Fraction Ishihama et al 2008(42)	Overall	1091	201883.19	22.95	110.54	**-0.1606**	**9.69E-08**	**0.0653**	**0.0309**	**-0.0696**	**0.0215**	0.0502	0.0973
	Overall without Ribosomal Proteins	1038	3685.60	22.95	110.50	**-0.1747**	**1.45E-08**	**-0.0844**	**0.0065**	**-0.1035**	**0.0008**	-0.0388	0.2113
	Cytoplasmic	874	251724.34	22.91	110.75	**-0.1248**	**0.0002**	**0.0849**	**0.0120**	**-0.0663**	**0.0499**	0.0530	0.1174
	Cytoplasmic without Ribosomal Proteins	821	4353.39	22.91	110.72	**-0.1440**	**3.44E-05**	**-0.0980**	**0.0049**	**-0.0995**	**0.0043**	-0.0494	0.1572
	Inner membrane	99	428.06	23.53	109.91	-0.0805	0.4285	-0.0532	0.6008	**-0.2200**	**0.0287**	**-0.2605**	**0.0092**
	Integral Membrane Protein	59	365.79	24.37	110.34	0.2052	0.1190	0.0922	0.4872	-0.0968	0.4657	-0.0782	0.5561
	Inner membrane lipoprotein	4	507.25	21.60	107.92	-	-	-0.3656	0.6344	-	-	-0.5773	0.4227
	Membrane anchored	29	457.27	22.29	109.69	-0.3169	0.0940	-0.1789	0.3532	-0.2479	0.1948	-0.2518	0.1876
	Membrane associated	4	541.25	23.05	109.46	-	-	-0.2223	0.7777	-	-	-0.3074	0.6926
	Membrane Lipoprotein	2	384.00	21.62	108.03	-	-	-	-	-	-	-	-
	Periplasmic w N-term mem anchor	1	147.00	23.44	111.77	-	-	-	-	-	-	-	-
	Periplasmic	77	1322.96	22.84	109.59	0.0642	0.0505	-0.1758	0.1262	-0.2120	0.0642	-0.1067	0.3557
	Outer membrane	41	2616.32	22.61	109.27	-0.2947	0.0615	-0.0260	0.8719	**-0.3497**	**0.0250**	-0.1445	0.3675
	Outer membrane Lipoprotein	21	1432.67	21.95	108.14	**-0.5651**	**0.0076**	**-0.4722**	**0.0307**	**-0.4462**	**0.0426**	-0.2605	0.2542
	Outer Membrane ß-barrel protein	20	3859.16	23.30	110.47	0.1173	0.6224	-0.0602	0.8010	-0.2165	0.3591	-0.3991	0.0813

Note: 12 proteins exluded (plasmid origins: P07676, P06717,Q00191,P22997,P62593,P27190,P02978,P00642,Q52280,Q57154 - Not in K-12: P15484,P11257)

BOLD: P < 0.05

Notes: Echobase sequence and location data was merged with the abundance values of Ishihama et al (42) using Uniprot numbers. We excluded 12 proteins that did not have Echobase information. Ten had plasmid origins: P07676, P06717, Q00191, P22997, P62593, P27190, P02978, P00642, Q52280, Q57154, and two were not found K-12: P15484 and P11257.

Table S14: Comparison of protein chain length and average synthetic cost (ASC) or mean amino acid mass in *E. coli.*

					Length vs. Avg Cost				Length vs. Avg Mass			
			Resp.		Pearson		Spearman		Pearson		Spearman	
Location	**Number**	**Avg Length**	**Avg Cost**	**Avg Mass**	**R**	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**
Overall - *E. coli*	4333	312.46	23.55	111.01	**-0.0992**	**5.92E-11**	**-0.0503**	**0.0009**	**-0.1225**	**6.00E-16**	**-0.1738**	**0, t=-11.62**
Cytoplasmic	2859	296.50	23.32	111.61	**-0.1329**	**9.86E-13**	**-0.1263**	**1.23E-11**	**-0.1414**	**3.08E-14**	**-0.1949**	**0, t=-10.62**
Inner membrane	972	357.55	24.65	109.99	**-0.3051**	**0, t=-9.98**	**-0.2850**	**1.08E-19**	**-0.1115**	**0.0005**	**-0.1469**	**4.22E-06**
Integral Membrane Protein	777	369.58	25.01	109.78	**-0.4555**	**0, t=-14.25**	**-0.4232**	**0, t=-13.00**	**-0.0961**	**0.0074**	**-0.1275**	**0.0004**
Inner membrane lipoprotein	14	485.50	21.99	109.27	0.1662	0.5701	-0.2489	0.3908	0.0715	0.8080	-0.0419	0.8870
Membrane anchored	158	284.58	23.37	110.99	**-0.2592**	**0.0010**	**-0.2887**	**0.0002**	-0.1158	0.1472	-0.1588	0.0463
Membrane associated	10	390.50	22.71	109.97	0.5009	0.1403	0.4681	0.1725	0.0841	0.8172	0.0790	0.8282
Membrane Lipoprotein	7	439.29	22.97	109.51	0.0601	0.8982	0.2500	0.5887	0.5479	0.2030	0.6071	0.1482
Periplasmic w N-term memb anchor	6	271.33	23.48	111.99	-0.4745	0.3416	-0.2571	0.6228	-0.4790	0.3364	-0.3143	0.5441
Periplasmic	337	309.42	22.85	109.72	0.1030	0.0590	**0.1835**	**0.0007**	0.1189	0.0291	0.1389	0.0107
Outer membrane	149	334.95	22.85	109.62	-0.0217	0.7929	0.1089	0.1862	0.0678	0.4111	0.1858	0.0233
Outer membrane Lipoprotein	96	216.46	22.56	109.01	-0.0435	0.6739	0.0324	0.7543	0.0413	0.6892	0.1490	0.1474
Outer Membrane ß-barrel protein	53	549.57	23.37	110.73	**-0.4286**	**0.0014**	-0.3440	0.0117	-0.2493	0.0718	-0.1236	0.3779
Extracellular	16	280.44	20.65	104.90	-0.1676	0.5351	-0.1118	0.6800	0.0041	0.9880	-0.1472	0.5865
Yeast All	6732	449.90	26.59	113.20	**-0.1853**	**0, t=-15.47**	**-0.2173**	**0, t=-18.26**	**-0.0428**	**0.0004**	**-0.0952**	**4.83E-15**
Yeast Cytoplasmic	2845	498.02	25.86	112.86	0.0164	0.3824	0.0119	0.5261	0.0247	0.1872	-0.0078	0.6783
Yeast Extracellullar	54	479.54	24.88	107.16	-0.1672	0.2268	0.0081	0.9535	-0.0316	0.8203	0.0959	0.4905
Yeast Cell Wall	44	493.25	24.04	105.27	-0.0991	0.5220	0.0394	0.7998	0.0073	0.9627	0.2272	0.1381
***M. tuberculosis* All**	3927	339.23	22.15	107.14	**-0.1116**	**2.36E-12**	-0.0171	0.2848	**-0.1952**	**0, t=-12.47**	**-0.1437**	**1.08E-19**
***M. tuberculosis* Cytoplasmic**	2111	348.93	22.25	108.56	-0.0439	0.0438	-0.0154	0.4793	**-0.1451**	**2.09E-11**	**-0.1841**	**1.53E-17**
***M. tuberculosis* Extracellullar**	140	513.19	19.45	96.00	**-0.4957**	**4.73E-10**	**-0.4958**	**4.72E-10**	**-0.5585**	**7.51E-13**	**-0.5670**	**2.81E-13**
***M. tuberculosis* Cell Wall**	8	552.38	20.97	103.19	-0.4829	0.2255	-0.0952	0.8225	-0.4522	0.2606	-0.0952	0.8225
***P. syringae* All**	5608	329.66	23.05	110.08	**-0.0967**	**3.98E-13**	**-0.0595**	**8.16E-06**	**-0.0867**	**7.67E-11**	**-0.1284**	**0, t=-9.69**
***P. syringae* Cytoplasmic**	1973	317.70	22.96	110.98	**-0.1083**	**1.42E-06**	**-0.0913**	**4.92E-05**	**-0.1407**	**3.41E-10**	**-0.1934**	**4.34E-18**
***P. syringae* Extracellular**	84	433.73	21.58	107.30	-0.0855	0.4394	-0.1262	0.2525	-0.0902	0.4147	-0.1653	0.1330
***P. syringae* HOPs & helpers**	51	437.31	21.48	107.68	-0.0635	0.6577	-0.2346	0.0975	-0.0445	0.7567	-0.2175	0.1253

BOLD: P < 0.01

Table S14: Comparison of protein chain length and average synthetic cost (ASC) or mean amino acid mass in *E. coli* (continued).

***E. coli* proteins**					**Length vs. Avg Cost**				**Length vs. Avg Mass**			
					Pearson		**Spearman**		**Pearson**		**Spearman**	
Protein Length	**Number**	**Avg Length**	**Avg Cost**	**Avg Mass**	**R**	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**
All proteins	4333	312.46	23.55	111.01	**-0.0992**	**5.92E-11**	**-0.0503**	**0.0009**	**-0.1225**	**6.00E-16**	**-0.1738**	**0, t=-11.62**
≤25 amino acids	7	17.57	30.50	121.29	-0.7388	0.0578	-0.8469	0.0162	-0.5140	0.2379	-0.3784	0.4026
≤50 amino acids	65	37.82	25.61	113.55	**-0.4416**	**2.31E-04**	**-0.2536**	**0.0416**	-0.2284	0.0672	-0.1155	0.3595
≤75 amino acids	240	57.76	24.46	113.03	**-0.3644**	**6.01E-09**	**-0.2314**	**0.0003**	-0.1140	0.0781	-0.0710	0.2732
≤100 amino acids	478	72.91	23.95	112.32	**-0.3524**	**2.01E-15**	**-0.2698**	**2.05E-09**	**-0.1645**	**3.04E-04**	**-0.1442**	**0.0016**
≤125 amino acids	718	86.09	23.88	112.20	**-0.2215**	**2.00E-09**	**-0.1504**	**5.20E-05**	**-0.1172**	**0.001658**	**-0.1124**	**0.0026**
≤150 amino acids	954	98.88	23.81	112.07	**-0.1888**	**4.17E-09**	**-0.1285**	**6.90E-05**	**-0.1104**	**6.37E-04**	**-0.1131**	**0.0005**
≤175 amino acids	1187	111.28	23.75	111.97	**-0.1697**	**4.00E-09**	**-0.1150**	**7.15E-05**	**-0.1092**	**1.63E-04**	**-0.1129**	**9.68E-05**
≤200 amino acids	1438	124.62	23.71	111.82	**-0.1574**	**1.95E-09**	**-0.1098**	**2.99E-05**	**-0.1296**	**8.24E-07**	**-0.1361**	**2.22E-07**
≤300 amino acids	2403	175.21	23.64	111.50	**-0.1033**	**3.90E-07**	**-0.0549**	**0.0071**	**-0.1451**	**8.83E-13**	**-0.1504**	**1.26E-13**
≤500 amino acids	3759	251.50	23.59	111.01	**-0.0674**	**3.53E-05**	-0.0318	0.0513	**-0.2223**	**0, t=-13.97**	**-0.2296**	**0, t=-14.46**
≤1000 amino acids	4276	300.17	23.56	111.01	**-0.0794**	**2.03E-07**	**-0.0412**	**0.0070**	**-0.1320**	**4.45E-18**	**-0.1765**	**0, t=-11.72**
≥50 amino acids	4272	316.39	23.53	110.97	**-0.0774**	**4.13E-07**	-0.0331	0.0306	**-0.1129**	**1.35E-13**	**-0.1648**	**0, t=-10.92**
≥100 amino acids	3861	341.79	23.50	110.84	**-0.0706**	**1.13E-05**	-0.0191	0.2360	**-0.0830**	**2.42E-07**	**-0.1303**	**4.37E-16**
≥150 amino acids	3388	372.17	23.49	110.71	**-0.0669**	**9.69E-05**	-0.0157	0.3594	**-0.0465**	**0.0068**	**-0.0937**	**4.65E-08**
≥200 amino acids	2902	405.27	23.48	110.61	**-0.0767**	**3.50E-05**	-0.0222	0.2311	-0.0132	0.4765	**-0.0595**	**0.0013**
≥300 amino acids	1938	482.59	23.45	110.39	**-0.0848**	**1.87E-04**	0.0087	0.7017	**0.0695**	**0.0022**	**0.0699**	**0.0021**
≥500 amino acids	581	709.12	23.35	110.97	**-0.2317**	**1.60E-08**	**-0.1908**	**3.64E-06**	-0.0355	0.3935	0.0479	0.2488
≥1000 amino acids	57	1234.70	22.82	110.50	**-0.3158**	**0.016716**	-0.3325	0.0115	-0.1590	0.2375	-0.0148	0.9131

BOLD: P < 0.01

Table S15: Comparison of the ASCs of extracellular vs intracellular fiber and fiber-associated proteins in *E. coli*.

	Intracellular	**Extracellular**
Fiber proteins:	RecA, MreB, ZapB (YiiU) RodZ (YfgA), FtsZ, MinD	FliC, FlgE, FimA, CsgA PapA* (P04127), FlaK**
Fiber-associated proteins‡:	*FtsA, ZipA, ZapA*, FtsE FtsX, FtsK, FtsQ, FtsL FtsB, FtsW, FtsI, FtsN *MinC, MinE*, DicB, *MreC* MreD	*FliD, FlgK, FlgL, FlgD*, *FliK* *CsgB*, *FimF*, FimG, FimH, PapE (P08407)*, PapF (P08408)* PapG (P13720) *, *PapK* (P62532)*

Note: Direct fiber-associated proteins (contact major fiber protein) are italicized above

	Mean	Student *t* Value	P Value†	Degrees of Freedom	Cost Diff (ATP)	U-Test Z	U-Test P Value
vs Intracellular Fibers	21.575						
Extracellular Fibers	19.844	5.410	0.000428	9	1.731	2.8823	0.00395
Extracellular Fibers without PapA	19.696	5.912	0.000357	8	1.879	2.7386	0.00617
Extracellular Fibers without FlaK	19.850	4.713	0.002175	7	1.725	2.7386	0.00617
Extracellular Fibers without PapA, FlaK	19.666	5.087	0.003811	5	1.909	2.5584	0.01052
vs Intracellular Fiber-associated	22.944						
Extracelluar Fiber-associated	21.534	2.352	0.026514	26	1.410	2.532	0.01134
Extracellular Fiber-associated without Pap Proteins	20.782	4.265	0.000269	24	2.162	3.4224	0.00062
vs Intracellular Direct Fiber-associated	22.031						
Extracelluar Direct Fiber-associated	20.752	2.843	0.014807	12	1.280	1.8074	0.0707
Extracelluar Direct Fiber-associated without PapK	20.536	3.512	0.004863	11	1.495	2.2857	0.02227
vs Cytoplasmic Proteins	23.555						
Intracellular Fibers	21.575	9.952	0.000175	5	1.980	3.5948	0.00032
Extracellular Fibers	19.844	14.685	2.65E-05	5	3.711	4.1923	2.76E-05
Extracellular Fibers wihout PapA	19.696	15.435	0.000103	4	3.859	3.8367	0.00012
Extracellular Fibers without FlaK	19.850	11.991	0.000277	4	3.705	3.8256	0.00013
Extracellular Fibers wihout PapA, FlaK	19.666	12.157	0.001198	3	3.889	3.4321	0.0006
Intracellular Fiber-associated	22.944	1.523	0.147252	16	0.611	2.4607	0.01387
Extracellular Fiber-associated	21.534	4.523	0.000698	12	2.021	4.366	1.27E-05
Intracellular Direct Fiber-associated	22.031	5.334	0.003104	5	1.524	2.9116	0.0036
Extracelluar Direct Fiber-associated	20.752	8.025	8.93E-05	7	2.803	4.4242	9.68E-06

† Student *t* test: Assumes Unequal Variance. Confidence Intervals are 2 tailed

* Not in K-12

** Average ASC of Flagellin phase 2 variable forms in Table S18D (Tab G in Data Set S1). Not in K-12

‡ Intracellular fiber associated proteins located in the cytoplasm or inner membrane compartments.

Table S16: Calculations of cost savings and fitness.

Description	Value calculated/used	Literature values and Notes
Average Doubling Time	30 minutes	30 minutes(1)
Number of H^+ per ATP	4 (92, 104)	4 (92, 104)
Number of H^+ per flagellar rotation	1240 (60)	800 (59, 65), ~1000 (22), 1200 (23, 71), 1240±240 (60), 1303 (106),1170-1270 (106),
Flagella Rotation rates*.**	100Hz (45, 71)	10-20 (88),13.4 (12),100 (45, 53, 59, 71), 170 (88), 200(88), 250 (22),60-270 (60, 88),300 (14, 23, 66, 88)
Number of H^+/s (100Hz, 1flag.)	124,000	-
Number of H^+/s (100Hz, 5flag.)	620,000	-
Growth disadvantage flagella	-	2% (71)
Flagellar rotation cost**	-	0.1% total energy growth conditions (71)
Growth of flagella	-	10000 aa/s [49, 50]
Respiratory chain total	20 million H^+/s (71)	20 million H^+/s (71)
Total H^+ per doubling (30 min.)	36 billion	or 9 billion ATPs if all H^+ are converted to ATPs
F_0F_1-ATPases H^+/s	-	3 million (71), 8 million(71)
ATP molecules in cell	-	0.5 – 3 million (1), 0.3-1.4 billion(32)
ATP Cost per cell***	20 billion	6-11 billion (theoretical) (64), 12-20 billion (experimental) (64), 20-60 billion (2, 94), 55 billion(1)
Cost per aa (% each aa is same)	27.36	-
Average Cost per aa (MG1655)	23.555	-
FliC Cost per aa	19.155	-
FimA Cost per aa	19.525	-
Flagella: FliC		
Number of Flagella per cell	5	0-15(8 typical) (70, 71),5-10 (71),10 (1, 23), 9.9±2.9 log & 4.5±2.1 in stationary (49), 6.6±1.4 (46),6±2 (82), 2.5 to 4.4 (41), 4 (45)
Subunits FliC per flagella	20000 (45, 71)	20000 (20, 22, 45, 55, 71), 5340/helix turn (13, 34), 30000 (110), 33000 (3000x11) (1)
Savings per aa vs average	4.4001	-
Length of FliC	498 (63)	-
Savings per FliC	2191.2 ATPs	-
One Gly to Trp substitution in FliC	6,260,000 ATPs	using 5 flagella
One Gly to Trp substitution in FliC	0.0313%(10^{-4}) of total cell cost	using 5 flagella and 20 billion ATPs per cell
Total savings per flagellum	43824894 ATP	20000 FliC subunits
Saving for 5 flagella	219124470 ATP	-
Saving for 5 flagella	876497883 H^+	$4H^+$/ATP
Time Swimming (100Hz ,1 flag.)	1412 seconds	23.5 min: 1 flag. savings, constantly spinning
% Savings Total Respiratory Chain	2.43%	using 30 minute doubling, 5 flagella
% Savings ATP Cost per cell	1.096%	using 5 flagella and 20 billion ATPs per cell
Type 1 Pili: FimA		
Number of Pili per cell	300	200-500(15, 45),100-1000 (71), 100-300(1, 70, 71), 100-200(54), 171-200 (90), 100-500 (103)
Number of FimA per pili	1000	1000 (15, 45), 3000 (45), 500 to 3000 (18, 33, 74), >>100(47), ~2800 (33) (19.31nm*27 & 2um)
Savings per aa vs average	4.0302	-
Length of FimA	182 (63)	-
Savings per FimA	733.5 ATPs	-
Total savings for Type 1 pili	220048920 ATP	-
Total savings for Type 1 pili	880548129 H^+	$4H^+$/ATP
Time Swimming (100Hz ,5 flag.)	1420 seconds	23.7 minutes: all 5 flagella spinning
% Savings Total Respiratory Chain	2.45%	30 minute doubling
% Savings ATP Cost per cell	1.100%	using 300 fimbriae and 20 billion ATPs per cell

Table S16: Calculations of cost savings and fitness (continued).

Typical Aerobic Gram Negative Organism

Extracellular protein aa cost (50% GC)	21.443 ATPs/aa	from equation of line fitting Fig. 5B
Total protein aa cost (50%GC)	23.492 ATPs/aa	from equation of line fitting Fig. 5B
Savings per Extracellular protein	2.0495 ATPs/aa	8.274% reduction in cost vs Total proteins
Total proteins/cell	2.35x10^6(71)	3.6x10^6 (1)
Extracellular proteins/cell	5.0x10^5	1x10^6(1); 100875 (5 flagella); 300900 (300 Type I Pili); 5x10^5 S-layer proteins per cell(87)
Extracellular protein savings	1.537%	Assumes all amino acids are recovered. 300aa/protein, 2.0495ATPs/aa, 2x10^{10} ATPs/cell, 5x10^5 extracellular proteins per cell
Extracellular protein synthetic cost	16.08%	Assumes all amino acids are lost. 300aa/protein, 21.443 ATPs/aa, 2x10^{10} ATPs/cell, 5x10^5 extracellular proteins per cell
Extracellular protein synthetic cost if aa biases equaled cellular proteins	17.62%	Assumes all amino acids are lost. 300aa/protein, 23.492 ATPs/aa, 2x10^{10} ATPs/cell, 5x10^5 extracellular proteins per cell

Direct Competition

Assumes savings directly affect doubling times

Generations	**Fold Increase**	**% Cheaper Organism**
100	2.922	74.503
200	8.539	89.516
250	14.596	93.588
500	213.043	99.533
1000	45387.3	99.998

Notes: Most values are for *E. coli* K-12 or *S. typhimurium* except in a few instances (*Streptococcus* (53, 60))

*Many values are maximum rates and depend on temperature(14, 60) and other factors such as PMF (28) and pH (59, 65). Values can differ significantly depending on how they are measured (bundle, individual fiber, tethered cell, bead on polyhook, electrorotation of tethered cells) (91)(88). Above 100-175 Hz(19) torque decreases significantly (efficiency plummets) and approaches zero at ~300Hz (effect is greater at lower temperatures and lower pH) (11, 14, 28, 65). For free-swimming cells the motor saturation point is about 100Hz (MacNab in (72)).

** Source of flagella power is PMF. Under aerobic respiration flagella rotation is saturation as it is a small percent of the total protons pumped. Under anaerobic glycolysis, ATP hydrolysis becomes the major source of flagellar power and swimming speed is reduced (71).

*** Values vary considerably and depend on conditions including growth rate, aerobic vs anaerobic conditions, availability and type of carbon source (glucose vs acetate), etc. Generally aerobic growth and richer media conditions require less ATPs/cell. A good discussion is listed in source (64) (references cited therein are (2, 8, 10, 26, 27, 37, 72, 77, 93-96, 102)). The calculation requires multiplying the cells per gram (dry cell weight) times YATP$_{max}$ (Number of ATP per gram of cells). Cell weight used in (64) is 0.28pg per cell (40 minutes growth on glucose(71) = 3.6 trillion cells per gram); however, cell weight varies with growth conditions (71) and changes would affect values listed in the table. Theoretical values are smaller than experimental values because the former exclude some cell maintenance cost including component turnover, transport cost, and membrane polarization (64).

Table S17A: Comparisons of the mean ASCs of proteins in locations of *E. coli* using different cost values.

		Akashi & Gojobori		Craig & Weber		Molecular Mass	
	Number	Mean	% of Overall	Mean	% of Overall	Mean	% of Overall
Overall	4333	23.5549		20.3139		111.0052	
Cytoplasm	2859	23.3171	98.99	19.8168	97.55	111.6093	100.54
Inner Membrane	972	24.6529	104.66	22.2954	109.75	109.9862	99.08
Integral Membrane Protein	777	25.0102	106.18	22.9046	112.75	109.7845	98.90
Membrane anchored	158	23.3743	99.23	20.0457	98.68	110.9876	99.98
Periplasmic N-term mem anchor	6	23.4800	99.68	19.6510	96.74	111.9889	100.89
Inner Membrane Lipoprotein	14	21.9869	93.34	18.3978	90.57	109.2690	98.44
Membrane associated	10	22.7075	96.40	19.2283	94.66	109.9723	99.07
Membrane Lipoprotein*	7	22.9700	97.52	19.8973	97.95	109.5073	98.65
Periplasm	337	22.8538	97.02	19.3281	95.15	109.7200	98.84
Outer Membrane	149	22.8521	97.02	19.4946	95.97	109.6250	98.76
OM ß-barrel protein	53	23.3737	99.23	20.2872	99.87	110.7320	99.75
OM Lipoprotein	96	22.5642	95.79	19.0570	93.81	109.0138	98.21
Cell surface appendage	12	20.4738	86.92	17.1651	84.50	103.8034	93.51
Secreted and released	4	21.1865	89.95	17.1547	84.45	108.1931	97.47
Extracellular	16	20.6520	87.68	17.1625	84.49	104.9008	94.50

		Cabons per aa		Nitrogens per aa		Sulfur per aa*		Mean ΔG per aa	
	Number	Mean	% Overall	Mean	% Overall	Mean	%Overall	Mean	% Overall
Overall	4333	4.9663		1.3596		0.0425		206.0956	
Cytoplasm	2859	4.9530	99.73	1.3857	101.92	0.0425	100.00	207.2800	100.57
Inner Membrane	972	5.0646	101.98	1.2950	95.25	0.0452	106.30	204.6274	99.29
Periplasm	337	4.8686	98.03	1.3387	98.46	0.0371	87.33	202.4783	98.24
Outer Membrane	149	4.8464	97.59	1.3373	98.36	0.0394	92.57	202.7336	98.37
Extracellular	16	4.5437	91.49	1.2744	93.73	0.0228	53.63	191.1559	92.75

*U-test of S/aa: Extracellular vs Cytoplasmic, P=1.47E-6. For Sec- (Table S7) %Overall =42.8%, P=1.33E-9.

	Mean cost	Spearman			Kendall				Pearson		
vs Akashi & Gojobori	23.555	R	t-test	p value	Tau	Z	p value	Gamma	R	t-test	p value
Craig & Weber	20.314	0.90471	139.756	→0	190.576	18809.6	→0	0.73907	0.91756	151.879	→0
Molecular Mass	111.005	0.57107	45.7812	→0	89.0564	8789.73	→0	0.42936	0.59229	48.3777	→0
Carbons per Amino Acid	4.966	0.90249	137.894	→0	279.165	27553.2	→0	0.91605	0.91605	150.311	→0
Nitrogens per Amino Acid	1.360	0.05849	3.8556	0.00012				0.05176	0.03388	2.23099	0.02573
Seawater ΔG(kJ /mol) 18^{o}C	206.096	0.62187	52.2601	→0	95.4152	9417.33	→0	0.46002	0.636	54.2393	→0
vs Molecular Mass	111.005										
Seawater ΔG(kJ /mol) 18^{o}C	206.096	0.90002	135.896	→0				0.733	0.90347	138.706	→0
Carbons per Amino Acid	4.966	0.784	83.1171	→0	201.118	19850	→0	0.61254	0.80177	88.2911	→0
Nitrogens per Amino Acid	1.360	0.65254	56.6727	→0				0.4741	0.64499	55.5455	→0

Table S17A: Comparisons of the mean ASCs of proteins in locations of *E. coli* using different cost values (continued).

Mann Whitney U Test	Akashi & Gojobori		Craig & Weber		Molecular Mass	
	Z	p value	Z	p value	Z	p value
vs Cytoplasm						
Inner Membrane	23.196	4.98E-119	29.497	3.18E-191	13.025	8.85E-39
Integral Membrane Protein	27.638	3.90E-168	34.510	5.65E-261	13.352	1.16E-40
Membrane anchored	0.649	0.5164	0.013	0.9897	2.258	0.0240
Periplasmic w N-term mem anchor	0.367	0.7139	0.245	0.8068	0.277	0.7820
Inner Membrane Lipoprotein	4.153	3.28E-05	3.795	0.0001	2.908	0.0036
Membrane associated	1.777	0.0755	1.286	0.1985	1.831	0.0670
Membrane Lipoprotein*	1.124	0.2609	0.470	0.6384	1.950	0.0512
Periplasm	6.029	1.65E-09	5.698	1.21E-08	9.772	1.48E-22
Outer Membrane	3.713	2.05E-04	2.510	0.0121	5.975	2.30E-09
OM ß-barrel protein	0.305	0.7604	2.374	0.0176	1.843	0.0653
OM Lipoprotein	4.892	9.98E-07	4.906	9.31E-07	6.151	7.72E-10
Cell surface appendage	5.606	2.07E-08	4.883	1.05E-06	5.690	1.27E-08
Secreted and released	2.269	0.0233	2.264	0.0236	1.980	0.0477
Extracellular	5.983	2.19E-09	5.355	8.55E-08	5.912	3.38E-09
vs All						
Cytoplasm	6.223	4.89E-10	8.631	6.09E-18	7.263	3.78E-13
Inner Membrane	19.530	6.17E-85	24.304	1.77E-130	8.729	2.58E-18
Integral Membrane Protein	24.013	2.05E-127	29.427	2.51E-190	9.405	5.20E-21
Membrane anchored	2.102	0.0356	2.156	0.0311	0.204	0.8381
Periplasmic w N-term mem anchor	0.051	0.9594	0.735	0.4621	0.807	0.4198
Inner Membrane Lipoprotein	4.275	1.91E-05	4.034	5.49E-05	2.215	0.0268
Membrane associated	2.103	0.0355	1.777	0.0756	1.181	0.2374
Membrane Lipoprotein*	1.317	0.1878	0.836	0.4030	1.410	0.1587
Periplasm	8.243	1.69E-16	8.648	5.25E-18	6.727	1.74E-11
Outer Membrane	5.149	2.62E-07	4.599	4.25E-06	3.921	8.83E-05
OM ß-barrel protein	0.761	0.4466	0.633	0.5269	0.479	0.6316
OM Lipoprotein	5.890	3.86E-09	6.231	4.62E-10	4.559	5.14E-06
Cell surface appendage	5.567	2.60E-08	4.958	7.12E-07	5.496	3.89E-08
Secreted and released	2.369	0.0179	2.393	0.0167	1.643	0.1004
Extracellular	6.001	1.96E-09	5.487	4.10E-08	5.578	2.44E-08

Table S17B: Comparisons of the mean ASCs of proteins in locations of *S. cerevisiae* using different cost values.

		Wagner Respiratory		Wagner Fermentative		Bragg & Wagner Respiratory		Bragg & Wagner Fermentative		Amino Acid Molecular Mass	
	Number	Mean	% of Overall	Mean	% of Overall	Mean	% of Overall	Mean	% of Overall	Mean	% of Overall
Overall	6732	26.59		6.61		25.14		5.81		113.20	
Cytoplasm	2845	25.86	97.25	6.46	97.68	24.32	96.73	5.61	96.51	112.86	99.70
PM	186	27.05	101.70	6.54	98.99	25.94	103.17	5.94	102.11	111.66	98.64
Integral/Endo Memb.	172	28.46	107.04	6.69	101.22	27.42	109.04	6.12	105.19	112.81	99.65
Cell Wall	44	24.04	90.39	5.64	85.31	23.21	92.30	5.18	89.17	105.27	93.00
Extracellular	54	24.88	93.54	5.92	89.56	24.00	95.46	5.44	93.57	107.16	94.66
Extracellular & Cell Wall	98	24.50	92.13	5.79	87.65	23.65	94.04	5.32	91.59	106.31	93.92

	Mean	Spearman			Kendall				Linear (Pearson)		
vs Yeast respiratory (Wagner)	26.5937	**R**	**t-test**	**p value**	**Tau**	**Z**	**p value**	**Gamma**	**R**	**t-test**	**p value**
Yeast respiratory (Bragg&Wagner)	25.145	0.959	278.2	→0	772.5	95050	→0	0.83524	0.966	308.5	→0
Yeast fermentative (Wagner)	6.61024	0.677	75.37	→0	380.3	46794	→0	0.49749	0.69	78.23	→0
Yeast fermentative (Bragg&Wagner)	5.81345	0.693	78.9	→0				0.51175	0.713	83.42	→0
Molecular Mass	113.198	0.483	45.31	→0				0.34333	0.547	53.6	→0
vs Yeast fermentative (Wagner)											
Yeast respiratory (Bragg&Wagner)		0.586	59.32	→0	237.7	29252	→0	0.42314	0.604	62.11	→0
Yeast fermentative (Bragg&Wagner)		0.802	110.1	→0	520.4	64025	→0	0.62318	0.855	135.1	→0
Molecular Mass		0.59	60.02	→0				0.42631	0.647	69.67	→0

Table S17B: Comparisons of the mean ASCs of protein in locations of *S. cerevisiae* using different cost values (continued).

Mann Whitney U Test	**Wagner Respiratory**		**Wagner Fermentative**		**Bragg & Wagner Respiratory**		**Bragg & Wagner Fermentative**		**Amino Acid Molecular Mass**	
vs Overall	**Z**	**p value**	**Z**	**p value**	**Z**	**p value**	**Z**	**p value**	**Z**	**p value**
Cytoplasm	14.8616	5.85E-50	12.7545	2.94E-37	15.4240	1.13E-53	15.0629	2.84E-51	5.6203	1.91E-08
PM	4.3898	1.13E-05	0.2869	0.7741849	6.0996	1.06E-09	4.9427	7.71E-07	7.0659	1.60E-12
Integral/Endo Membranes	12.3372	5.71E-35	3.8335	0.0001263	13.2731	3.32E-40	9.4490	3.42E-21	2.2301	0.0257411
Cell Wall	6.8980	5.27E-12	8.5229	1.56E-17	5.3273	9.97E-08	6.1732	6.69E-10	8.6931	3.53E-18
Extracellular	4.8461	1.26E-06	7.1035	1.22E-12	3.1973	0.001387	12.6762	8.01E-37	7.6434	2.11E-14
Extracellular & Cell Wall	8.1895	2.62E-16	10.9442	7.08E-28	5.9212	3.20E-09	10.1865	2.28E-24	11.4573	2.16E-30
vs Cytoplasm										
PM	9.0114	2.04E-19	3.7753	0.0001598	10.9088	1.05E-27	9.9169	3.52E-23	5.6918	1.26E-08
Integral/Endo Membranes	16.5602	1.35E-61	8.1420	3.89E-16	17.4758	2.19E-68	14.3009	2.16E-46	0.7752	0.4382106
Cell Wall	5.7802	7.46E-09	7.9861	1.39E-15	3.7201	0.0001991	4.8510	1.23E-06	8.5915	8.59E-18
Extracellular	3.0802	0.0020684	6.0747	1.24E-09	1.0085	0.3132149	1.8219	0.0684758	7.2476	4.24E-13
Extracellular & Cell Wall	6.1067	1.02E-09	9.7773	1.41E-22	3.2127	0.0013148	4.5628	5.05E-06	11.0433	2.36E-28

Table S17C: Comparisons of the mean ASCs of proteins in locations of *M. tuberculosis* H37Rv using different cost values.

		Akashi & Gojobori		Craig & Weber		Molecular Mass	
	Number	**Mean**	**% of Overall**	**Mean**	**% of Overall**	**Mean**	**% of Overall**
Overall	3927	22.1469		19.3088		107.1443	
Cytoplasm	2111	22.2479	100.46	19.1672	99.27	108.5627	101.32
Cytoplasmic Membrane	598	22.9850	103.78	20.8788	108.13	106.1764	99.10
Unknown	1070	21.8415	98.62	18.9533	98.16	106.3746	99.28
Extracellular	140	19.4465	87.81	17.5376	90.83	95.9992	89.60
Cell Wall	8	20.9739	94.70	17.8469	92.43	103.1890	96.31
Extracellular &Cell Wall	148	19.52908	88.18	17.55428	90.91	96.38786	89.96

	Mean cost	**Spearman**				**Pearson**		
vs Akashi & Gojobori	22.1456	**R**	**t-test**	**p value**	**Gamma**	**R**	**t-test**	**p value**
Craig & Weber	19.3088	0.8744062	112.9059985	→0	0.697734461	0.8728684	112.0696964	→0
Molecular Mass	107.1443	0.7032694	61.97531318	→0	0.532250842	0.7803322	78.17459575	→0

Mann Whitney U Test	**Akashi & Gojobori**		**Craig & Weber**		**Molecular Mass**	
	Z	**p value**	**Z**	**p value**	**Z**	**p value**
vs Cytoplasm						
Cytoplasmic Membrane	10.7167	8.495E-27	19.0031	1.61E-80	14.4926	1.35E-47
Unknown	8.9557	3.375E-19	5.1720	2.316E-07	15.0537	3.27E-51
Extracellular	12.3710	3.752E-35	10.6110	2.65E-26	15.3676	2.70E-53
Cell Wall	2.4495	0.0143057	2.0390	0.0414474	3.4030	0.0006665
Extracellular and Cell Wall	12.5622	3.405E-36	10.7611	5.26E-27	15.6863	1.88E-55
vs All						
Cytoplasm	2.0838	0.0371784	1.7299	0.0836427	11.3553	6.98E-30
Cytoplasmic Membrane	11.7610	6.203E-32	17.8934	1.33E-71	7.8344	4.66E-15
Unknown	7.5395	4.718E-14	6.4176	1.385E-10	7.3012	2.85E-13
Extracellular	11.8675	1.747E-32	10.6696	1.41E-26	13.6224	2.95E-42
Cell Wall	2.2392	0.025142	2.0735	0.0381271	2.7638	0.005713
Extracellular and Cell Wall	12.0425	2.123E-33	10.8408	2.21E-27	13.8675	9.96E-44

Table S19: Comparison of the ASC of Type III secretion effectors with cytoplasmic proteins.

			E. coli **K-12***		***E. coli*** **K-12***			***P. syringae pv. tomato*****		***P. syringae pv. tomato*****		
			Cytoplasm		**Overall**			**Cytoplasm**		**Overall**		
	ASC	**Num.**	**Z**	**P value**	**Z**	**P value**	**Savings**	**Z**	**P value**	**Z**	**P value**	**Savings**
Animal Type III Effectors†	22.233	64	5.835	5.39E-09	6.456	1.08E-10	1.32	4.064	4.81E-05	4.225	2.39E-05	0.81
Plant Type III Effectors†	21.740	38	7.181	6.90E-13	7.375	1.64E-13	1.81	5.878	4.16E-09	5.766	8.11E-09	1.31
Animal and Plant Type III Effectors†	22.049	102	8.928	4.35E-19	9.561	1.17E-21	1.51	6.722	1.79E-11	6.835	8.19E-12	1.00
Type III Secretion System, Chaperones†	23.468	65	***0.957***	***0.3385***	***0.163***	***0.8703***	0.09	2.764	5.71E-03	2.257	0.024037	-0.42
E. coli **Type III Effectors†**	21.564	10	3.100	1.94E-03	3.254	1.14E-03	1.99	2.770	5.60E-03	2.743	6.09E-03	1.48
E. coli&Salmonella **Type III Effectors†**	22.162	20	3.171	1.52E-03	3.539	4.02E-04	1.39	2.197	0.028041415	2.318	0.02046	0.88
P. syringae sp. **Type III Effectors†**	21.740	38	7.181	6.90E-13	7.375	1.64E-13	1.81	5.878	4.16E-09	5.766	8.11E-09	1.31
P. syringae pv tomato **HOPs**	21.757	44	7.966	1.67E-15	8.123	4.54E-16	1.80	6.679	2.40E-11	6.489	8.61E-11	1.29
P. syringae pv tomato **HOP Helpers**	19.745	7	4.450	8.59E-06	4.424	9.67E-06	3.81	4.387	1.15E-05	4.303	1.69E-05	3.30
P. syringae pv tomato **HOP w Helpers**	21.480	51	9.027	1.77E-19	9.170	4.73E-20	2.07	7.801	6.22E-15	7.612	2.70E-14	1.57

			E. coli **K-12***		***E. coli*** **K-12***		***P. syringae pv. tomato*****		***P. syringae pv. tomato*****	
			Cytoplasm		**Overall**		**Cytoplasm**		**Overall**	
	Mean Mass	**Num.**	**Z**	**P value**	**Z**	**P value**	**Z**	**P value**	**Z**	**P value**
Animal Type III Effectors†	109.290	64	4.062	4.87E-05	2.817	4.84E-03	2.661	7.80E-03	***0.789***	***0.429864***
Plant Type III Effectors†	108.600	38	5.529	3.21E-08	4.430	9.43E-06	4.509	6.50E-06	2.695	7.03E-03
Animal and Plant Type III Effectors†	109.033	102	6.535	6.36E-11	4.906	9.28E-07	4.798	1.60E-06	2.259	0.0239
Type III Secretion System, Chaperone†	111.579	65	***0.787***	***0.4310***	***0.827***	***0.408218***	***1.088***	***0.2765355***	3.525	4.23E-04
E. coli **Type III Effectors†**	108.191	10	2.366	0.0180	2.000	0.0455	2.041	0.0412	***1.447***	***0.148***
E. coli&Salmonella **Type III Effectors†**	109.513	20	2.339	0.0193	***1.609***	***0.1076***	***1.596***	***0.1106***	***0.434***	***0.664***
Pseudomonas **Type III Effectors†**	108.600	38	5.529	3.21E-08	4.430	9.43E-06	4.509	6.50E-06	2.695	7.03E-03
P. syringae pv tomato **HOPs**	108.457	44	6.281	3.36E-10	5.165	2.40E-07	5.261	1.43E-07	3.423	6.20E-04
***P. syringae pv tomato*HOP Helpers**	102.782	7	4.192	2.77E-05	4.043	5.28E-05	4.108	3.98E-05	3.863	1.12E-04
P. syringae pv tomato **HOP w Helpers**	107.678	51	7.368	1.73E-13	6.284	3.29E-10	6.384	1.73E-10	4.603	4.17E-06

		Num.	**ASC**	**Mean Mass**
E. coli **K-12***	**Cytoplasm**	2859	23.317	111.609
E. coli **K-12***	**Overall**	4333	23.555	111.005
P. syringae pv. tomato**	**Cytoplasm**	1973	22.961	110.982
P. syringae pv. tomato**	**Overall**	5608	23.047	110.075

BOLD ITALIC **=** Not significant, P > 0.05

* Sequences and locations: Echobase(39, 63)

** Sequences and locations: PSORTDB (81)

† From Arnold et al PLOS Pathology 5 (2009)

Table S20: Amino acid composition and cost savings in different compartments of *P. syringae*.

	Mean ATPs per aa	%A	%R	%N	%D	%C	%E	%Q	%G	%H	%I	%L	%K	%M	%F	%P	%S	%T	%W	%Y	%V	Total Amino Acids	Number
Overall Mean	23.05	10.55	6.45	3.09	5.36	1.13	5.66	4.48	7.47	2.33	5.03	11.20	4.00	2.58	3.67	4.75	6.14	5.07	1.41	2.49	7.13	329.67	5607
Overall Median	22.94	10.37	6.36	2.96	5.44	0.91	5.67	4.36	7.39	2.23	4.92	10.97	3.54	2.37	3.51	4.63	5.93	4.95	1.23	2.41	7.06	282.00	
Overall STDEVP	1.37	3.01	2.48	1.48	1.98	1.09	2.30	1.99	2.43	1.38	1.98	3.28	2.40	1.25	1.74	1.96	2.09	1.89	1.11	1.33	2.19	276.48	
Extracellular all	21.58	10.90	4.93	**4.89**	6.06	**0.60**	4.52	4.86	8.26	2.40	4.13	8.59	4.76	2.22	3.04	4.95	**9.12**	**6.61**	**0.85**	2.30	6.00	433.73	84
Extracellular - HOP	21.74	11.08	**4.00**	**4.85**	**6.76**	**0.69**	**4.03**	4.56	8.82	**1.71**	4.52	8.44	4.16	**1.88**	3.28	4.44	**8.26**	**7.75**	**1.22**	2.77	6.80	428.18	33
HOPs w Helpers	21.48	10.79	5.54	**4.91**	5.60	**0.55**	4.84	5.06	7.90	2.85	3.87	8.69	**5.16**	2.44	2.89	5.29	**9.67**	5.87	**0.61**	2.00	5.48	437.31	51
Outer Membrane	22.25	10.09	5.86	**4.96**	6.43	**0.34**	4.69	5.04	8.99	**1.35**	**3.72**	9.56	3.67	**1.55**	3.47	4.22	**7.84**	**6.45**	1.42	**3.84**	6.52	649.23	115
Periplasm	22.65	11.41	**4.18**	3.85	6.27	**0.63**	5.08	4.55	7.74	**1.50**	4.56	9.37	**6.54**	2.40	3.71	5.19	6.17	5.43	1.38	2.77	7.26	414.95	110
Inner Membrane	23.83	11.36	5.38	2.66	**3.63**	**0.81**	**3.90**	3.72	8.18	1.76	6.18	13.93	**2.80**	3.05	**4.59**	4.43	6.17	5.09	**1.90**	2.41	8.05	412.61	961
Unknown	22.91	10.73	6.46	3.21	5.42	1.20	5.43	4.85	7.28	2.32	4.55	10.56	4.28	2.59	3.50	5.03	6.51	5.31	1.46	2.56	6.74	282.73	2364
Cytoplasm	22.96	9.91	7.19	2.92	5.99	1.29	6.93	4.36	7.20	2.72	5.19	10.96	4.10	2.43	3.48	4.56	5.44	4.63	1.13	2.35	7.21	317.70	1973
Cost Savings (ATP)		15.66	0.06	12.66	14.66	2.66	12.06	11.06	15.66	-10.9	-4.94	0.06	-2.94	-6.94	-24.6	7.06	15.66	8.66	-46.9	-22.6	4.06		

Notes: **Bold** indicates >25% change in amino acid composition from the overall mean.

Underlined indicates >50% change in amino acid composition from the overall mean.

Amino acid percents are on a per protein basis. The percent amino acid composition of each protein in that location were averaged; they are unweighted for chain length. See Table S22B for weighted percents (% of all amino acids) of proteins in nonmembrane locations of *P. syringae*.

	Savings	A	R	N	D	C	E	Q	G	H	I	L	K	M	F	P	S	T	W	Y	V	Count
Extracellular-HOPs	**>1.5 ATPs**	**13.34**	-0.08	**13.86**	8.64	-0.95	-9.75	0.79	8.47	6.45	1.01	-0.08	-0.49	2.28	5.39	-5.29	**25.09**	**12.51**	**16.96**	2.07	-0.23	14
Extracellular-HOPs	**>0 ATPs**	**17.00**	-0.10	**13.25**	**12.36**	-0.48	-11.11	-3.07	**18.47**	2.37	1.13	-0.06	-0.35	3.52	7.54	-4.09	**21.40**	9.43	9.25	5.36	-1.82	23
HOPS and Helpers	**>1.5 ATPs**	**11.61**	-0.03	**10.16**	0.00	-0.93	-6.04	3.22	7.12	1.22	3.18	-0.06	-1.13	1.17	**11.20**	2.36	**25.68**	4.94	**18.70**	**10.16**	-2.53	24
HOPS and Helpers	**>0 ATPs**	-3.64	-0.05	**17.69**	1.88	-1.29	-5.84	6.24	0.45	-12.90	5.33	-0.12	-4.00	1.75	4.86	2.20	**53.36**	1.16	**38.54**	2.77	-8.39	47
Extraceullar All	**>1.5 ATPs**	**12.25**	-0.05	**11.53**	3.19	-0.94	-7.41	2.33	7.62	3.15	2.38	-0.07	-0.90	1.58	9.06	-0.46	**25.46**	7.73	**18.06**	7.18	-1.68	38
Extraceullar All	**>0 ATPs**	3.14	-0.06	**16.23**	5.32	-1.02	-7.57	3.18	6.37	-7.88	3.95	-0.10	-2.80	2.33	5.74	0.13	**42.86**	3.88	**28.92**	3.62	-6.23	70

Table S21: Comparison of the ASC of cellular and extracellular serine proteases.

					U-Test		Student *t* test			Savings (ATP)	
	ASC	Mass	Number	Length	Z	P value	T	Deg. Free	P value	per aa	per protein
vs Cellular Serine proteases	22.547	109.500	5592	494.01							
All Extracelluar Serine proteases	22.157	107.408	191	737.93	4.700	2.60E-06	5.872	206	1.71E-08	0.390	287.97
Extracelluar Serine proteases†	21.827	106.596	82	897.34	6.474	9.52E-11	6.657	84	2.70E-09	0.720	646.09
vs *Bacillus sp.* Cellular Serine proteases	22.454	108.502	447	502.75							
***Bacillus sp.* Extracellular Serine proteases**	21.210	103.899	23	455.61	5.185	2.16E-07	6.558	25	7.19E-07	1.245	567.08
***Bacillus sp.* Extracellular Serine proteases†**	21.273	104.409	18	485.87	4.305	1.67E-05	5.244	19	4.62E-05	1.181	573.85
vs *E. coli* Cellular Serine proteases	23.126	110.089	192	401.68							
***E. coli* Extracellular Serine proteases**	22.381	107.912	10	1335.56	3.856	1.15E-04	6.581	64	1.01E-08	0.745	995.41
***E. coli* Extracellular Serine proteases †**	22.319	107.914	8	1335.60	3.822	1.33E-04	6.729	49	1.74E-08	0.807	1077.88
E. coli Extracellular Serine proteases ‡	21.994	107.306	10	1005.30	4.639	3.51E-06	8.443	30	2.02E-09	1.133	1138.73
E. coli Extracellular Serine proteases †,‡	21.973	107.316	8	1005.69	4.302	1.69E-05	8.461	23	1.62E-08	1.154	1160.14
vs *E. coli* K-12 Cytoplasm	23.555	111.005	4333	312.46							
***E. coli* Cellular Serine proteases**	23.126	110.089	192	401.68	6.214	5.16E-10	4.546	216	9.09E-06	0.429	172.14
***E. coli* Cellular Serine proteases †**	23.142	110.393	26	387.65	2.151	3.15E-02	1.478	25	0.152	0.412	159.88
***E. coli* Extracellular Serine proteases**	22.381	107.912	10	1335.56	3.299	9.70E-04	16.626	11	3.84E-09	1.174	1567.75
***E. coli* Extracellular Serine proteases†**	22.319	107.914	8	1335.60	3.005	2.66E-03	16.631	9	4.59E-08	1.236	1650.24
E. coli Extracellular Serine proteases ‡	21.994	107.306	10	1005.30	4.108	3.99E-05	15.490	10	2.57E-08	1.561	1569.54
E. coli Extracellular Serine proteases †,‡	21.973	107.316	8	1005.69	3.720	1.99E-04	15.259	8	3.37E-07	1.582	1591.13

Note: Cellular serine proteases are serine proteases from Uniprot that are not secreted

† Homologs (Genes with same name and similar ASC, lengths, sequences) were averaged

‡ All the *E. coli* secreted serine proteases were autotransporters. Just the secreted domain was analyzed here.
Autotransporters are bacterial proteins that contain a β-domain which integrates into the outer membrane and assists the translocation of an α-domain to the cell surface where it is often cleaved autolytically or by other proteases (38). While β-domains have high similarity, α-domains are quite diverse and function as enzymes (protease, esterase, lipase) or assist cell aggregation, actin polymerization, or other virulent activities (38). While most β-domains had average ASCs, all the α-domains except IgA1 protease had lower ASCs. On average, α-domains had an 8.7% lower ASC than their respective β-domain (See Table S18C – Tab F in Data Set S1).

Table S22A: Effects of protein structure on ASC in *E. coli*.

Mean values based on per protein amounts (unweighted for length)

		Mean			Filtered Database			Neural Networks only			Unfiltered Database†		
	Count	**Mass/aa**	**ATPs/aa**	**Length**	**%H**	**%E**	**%C**	**%H**	**%E**	**%C**	**%H**	**%E**	**%C**
Overall	3212	111.38	23.26	297.75	39.82	17.54	42.64	41.41	17.80	40.79	33.91	21.14	44.95
Cytoplasmic	2859	111.61	23.32	296.47	40.58	17.08	42.33	42.40	17.27	40.33	ND	ND	ND
Periplasmic	337	109.72	22.85	309.42	33.85	20.96	45.18	33.47	22.00	44.53	34.13	20.88	45.00
Extracellular	16	104.90	20.65	280.44	29.29	26.64	44.07	31.14	24.82	44.04	29.38	26.72	43.90

Linear Correlation	Filtered Database						Neural Netwoks only						Unfiltered Database†					
vs ASC (ATPs/aa)	**%H**		**%E**		**%C**		**%H**		**%E**		**%C**		**%H**		**%E**		**%C**	
	R	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**
Overall	0.0408	0.0208	-0.0291	0.0996	-0.0360	0.0412	-0.0010	0.9527	0.0164	0.3514	-0.0174	0.3249	0.0167	0.3440	-0.0172	0.3290	-0.0072	0.6815
Cytoplasmic	0.0219	0.2408	0.0011	0.9552	-0.0379	0.0427	-0.0311	0.0964	0.0501	**0.0073**	-0.0023	0.9003	ND	ND	ND	ND	ND	ND
Periplasmic	-0.0102	0.8518	-0.0622	0.2547	0.0939	0.0852	-0.0300	0.5826	-0.0184	0.7359	0.0798	0.1439	-0.0121	0.8255	-0.0619	0.2572	0.0963	0.0776
Extracellular	0.2177	0.4179	0.0324	0.9053	-0.4340	0.0931	0.3288	0.2138	-0.0063	0.9816	-0.5458	0.0287	0.2022	0.4528	0.0599	0.8255	-0.4339	0.0931

Notes: ND - Not Done **BOLD:** P < 0.01 **† Overall contains only data from Periplasmic and Extracellular proteins**

Comparing Protein Secondary Structure between Locations

	Filtered Database				Neural Networks Only			
	Student *t**		**Mann Whitney**		**Student *t****		**Mann Whitney**	
vs Cytoplasmic	**T value**	**P value**	**Z**	**P value**	**T value**	**P value**	**Z**	**P value**
%H Periplasmic	6.622	1.16E-10	7.199	6.05E-13	9.639	1.08E-19	10.347	4.32E-25
%E Periplasmic	5.317	1.79E-07	4.551	5.33E-06	7.341	1.21E-12	7.734	1.04E-14
%C Periplasmic	4.768	2.62E-06	5.507	3.65E-08	8.383	8.15E-16	9.792	1.22E-22
%H Extracellular	2.023	**0.061**	2.487	0.013	2.695	0.017	3.058	0.002
%E Extracellular	2.212	0.043	2.108	0.035	2.305	0.036	2.628	0.009
%C Extracellular	0.557	**0.586**	0.656	**0.512**	1.498	**0.155**	1.348	**0.178**

Notes: Comparisons are versus same % secondary structure in cytoplasmic proteins

* Assumes unequal variance, two tailed. **BOLD –** P > 0.01. H-Helix, E- Strand, C-Coil.

Table S22A: Effects of protein structure on ASC in *E. coli* (continued).

Filtered Database (Percent of all amino acids in the Secondary Structure of each location)

		%A	%C	%D	%E	%F	%G	%H	%I	%K	%L	%M	%N	%P	%Q	%R	%S	%T	%V	%W	%Y
Overall	**Helix**	12.77	1.06	4.77	8.46	3.47	3.06	2.23	5.71	13.09	5.21	2.92	3.02	2.24	6.03	7.11	4.63	4.14	5.79	1.51	2.76
	Coil	7.37	1.32	7.75	5.68	2.74	12.23	2.91	3.24	6.23	4.86	2.40	5.73	7.80	4.07	5.20	6.97	5.91	4.29	1.01	2.29
	Strand	6.72	1.70	2.57	4.31	5.42	3.58	2.32	11.40	12.37	3.31	2.60	2.01	1.38	2.96	5.06	4.26	6.51	15.38	1.82	4.31
Cytoplasm	**Helix**	12.63	1.09	4.76	8.74	3.45	3.02	2.33	5.78	13.13	5.09	2.92	2.95	2.23	5.97	7.33	4.52	4.06	5.79	1.50	2.74
	Coil	7.25	1.38	7.69	5.86	2.79	12.33	3.04	3.29	6.38	4.68	2.43	5.51	7.79	3.99	5.40	6.86	5.79	4.30	0.98	2.26
	Strand	6.68	1.86	2.53	4.46	5.43	3.53	2.43	11.71	12.58	3.20	2.63	1.84	1.37	2.81	5.18	4.07	6.27	15.54	1.76	4.13
Extracellular	**Helix**	12.79	0.33	5.53	3.73	3.00	3.26	0.40	5.93	13.32	5.53	1.73	6.53	0.60	9.13	2.93	9.46	7.79	4.86	0.67	2.47
	Coil	12.01	0.46	7.90	2.98	1.74	13.44	0.46	2.46	4.41	4.05	1.64	9.65	4.82	3.85	1.64	10.83	11.29	4.82	0.31	1.23
	Strand	9.74	0.29	3.47	2.22	3.66	4.05	1.06	8.20	9.74	3.09	1.25	3.66	1.54	7.62	2.60	6.17	12.54	13.60	0.77	4.73
Periplasm	**Helix**	14.22	0.86	4.75	5.94	3.74	3.41	1.24	5.04	12.76	6.48	3.01	3.49	2.45	6.52	5.18	5.53	4.83	5.82	1.72	3.01
	Coil	8.08	0.90	8.20	4.50	2.37	11.47	2.05	2.85	5.17	6.28	2.21	7.17	8.03	4.64	3.82	7.67	6.61	4.16	1.26	2.56
	Strand	6.83	0.70	2.79	3.44	5.41	3.85	1.65	9.48	11.13	4.07	2.48	3.09	1.49	3.75	4.38	5.44	7.83	14.42	2.25	5.51

Average Cost

		Filtered Database		**Neural Networks**		**Unfiltered Database†**	
		ATPs/aa	**% total**	**ATPs/aa**	**% total**	**ATPs/aa**	**% total**
Overall	**Helix**	23.904	39.37	24.017	41.21	23.608	32.90
	Coil	21.420	43.41	20.871	41.63	20.902	46.59
	Strand	26.792	17.23	27.607	17.16	26.463	20.50
Cytoplasm	**Helix**	23.936	40.22	24.055	42.30	ND	ND
	Coil	21.490	42.98	20.949	41.08	ND	ND
	Strand	26.844	16.80	27.627	16.63	ND	ND
Extracellular	**Helix**	21.797	33.45	21.746	32.81	21.831	33.41
	Coil	18.255	43.44	17.875	43.37	18.188	43.10
	Strand	23.623	23.11	24.320	23.82	23.616	23.49
Periplasm	**Helix**	23.679	32.67	23.708	32.68	23.685	32.68
	Coil	21.019	46.88	20.426	46.09	21.009	46.09
	Strand	26.599	20.45	27.638	21.22	26.604	21.22

Percent Cost of

		vs Cytoplasmic		**vs Overall**	
		Filtered	**NN**	**Filtered**	**NN**
Extracellular	**Helix**	91.06	90.40	91.19	90.55
	Coil	84.94	85.33	85.22	85.64
	Strand	88.00	88.03	88.17	88.09
Periplasm	**Helix**	98.93	98.56	99.06	98.71
	Coil	97.81	97.50	98.13	97.87
	Strand	99.09	100.04	99.28	100.11

Notes: Values on this page are based on the total amino acids in all of that secondary structure of every protein in that location. Values in the table on the page above are from averaging % secondary structure in each protein. ND – not done. NN – Neural Networks.

† Overall contains only data from Periplasmic and Extracellular proteins

Table S22B: Effects of protein structure on ASC in *P. syringae*.

Mean values based on per protein amounts (unweighted for length)

		Mean			Filtered Database			Neural Networks only			Unfiltered Database†		
	Count	**Mass/aa**	**ATPs/aa**	**Length**	**%H**	**%E**	**%C**	**%H**	**%E**	**%C**	**%H**	**%E**	**%C**
Overall	2167	110.72	22.89	327.13	40.80	16.50	42.70	43.41	15.98	40.61	36.02	15.88	48.10
Cytoplasmic	1973	110.98	22.96	317.70	41.28	16.56	42.16	44.13	15.90	39.97	ND	ND	ND
Periplasmic	110	108.72	22.65	414.95	36.86	18.37	44.78	37.22	17.86	44.92	36.88	18.36	44.75
Extracellular	84	107.30	21.58	433.73	34.68	12.64	52.67	34.60	15.37	50.04	34.88	12.62	52.49
HOPS&Helpers	51	107.68	21.48	437.31	41.31	8.03	50.65	38.23	10.89	50.88	41.57	8.03	50.40
Extracelluar-HOPS	33	106.70	21.74	428.18	24.43	19.77	55.80	28.98	22.28	48.74	24.54	19.73	55.73

Linear Correlation

vs. ASC (ATPs/aa)	Filtered Database %H		%E		%C		Neural Netwoks only %H		%E		%C		Unfiltered Database† %H		%E		%C	
	R	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**	**R**	**P value**
Overall	-0.1329	**5.21E-10**	0.1340	**3.84E-10**	0.0825	**0.0001**	-0.1055	**8.57E-07**	0.1363	**1.90E-10**	0.0376	**7.98E-02**	0.0196	0.3616	0.0071	0.7421	-0.0331	0.1239
Cytoplasmic	-0.1717	**1.60E-14**	0.1259	**2.04E-08**	0.1605	**7.33E-13**	-0.1484	**3.45E-11**	0.1432	**1.64E-10**	0.1069	**1.96E-06**	ND	ND	ND	ND	ND	ND
Periplasmic	0.0346	0.7196	-0.1271	0.1858	0.0619	0.5205	-0.0466	0.6290	-0.0256	0.7909	0.1126	0.2416	0.0275	0.7758	-0.1280	0.1828	0.0707	0.4628
Extracellular	-0.0908	0.4116	0.1980	0.0710	-0.0434	0.6949	-0.1792	0.1029	0.1871	0.0884	0.0705	0.5238	-0.1008	0.3615	0.2057	0.0605	-0.0362	0.7435
HOPS&Helpers	0.1766	0.2152	0.2297	0.1049	-0.3477	0.0124	0.0621	0.6650	0.1996	0.1603	-0.2331	0.0998	0.1668	0.2420	0.2245	0.1132	-0.3377	0.0154
Extracelluar-HOPS	-0.2626	0.1398	0.1538	0.3929	0.2212	0.2161	-0.3455	0.0489	0.1440	0.4241	0.4587	**0.0073**	-0.2704	0.1280	0.1713	0.3405	0.2178	0.2235

Notes: ND - Not Done **BOLD:** $P < 0.01$ **† Overall contains only data from Periplasmic and Extracellular proteins**

Comparing Protein Secondary Structure between Locations

	Filtered Database				Neural Networks Only			
	Student *t**		Mann Whitney		Student *t**		Mann Whitney	
vs Cytoplasmic	**T value**	**P value**	**Z**	**P value**	**T value**	**P value**	**Z**	**P value**
H Periplasmic	3.510	6.21E-04	2.538	1.12E-02	6.302	4.22E-09	5.249	1.53E-07
E Periplasmic	2.356	0.020	2.126	0.034	2.682	8.29E-03	2.705	0.007
C Periplasmic	3.046	0.003	2.243	0.025	7.983	6.56E-13	6.670	2.55E-11
H Extracellular	3.410	0.001	2.721	0.007	6.154	2.00E-08	5.509	3.60E-08
E Extracellular	3.328	0.001	4.377	1.20E-05	0.478	**0.634**	1.436	**0.151**
C Extracellular	8.029	4.32E-12	8.232	2.22E-16	10.689	1.23E-17	9.724	2.37E-22
H HOPs	0.017	**0.987**	0.547	**0.584**	3.629	6.31E-04	3.075	0.002
E HOPs	6.334	2.93E-10	6.580	4.70E-11	5.447	1.23E-06	4.499	6.83E-06
C HOPs	5.469	1.31E-06	5.495	3.92E-08	9.032	3.10E-12	8.055	7.93E-16
H Extracellular - HOPs	5.417	5.38E-06	5.079	3.80E-07	5.528	3.88E-06	5.061	4.18E-07
E Extracellular - HOPs	1.508	**0.141**	1.146	**0.252**	3.337	0.002	5.652	1.59E-08
C Extracellular - HOPs	8.648	1.05E-17	6.438	1.21E-10	5.988	1.00E-06	3.297	0.001

Notes: Comparisons are versus same % secondary structure in cytoplasmic proteins

* Assumes unequal variance, two tailed. **BOLD –** $P > 0.01$. H-Helix, E- Strand, C-Coil.

Table S22B: Effects of protein structure on ASC in *P. syringae* (continued).

Filtered Database (Percent of all amino acids in the Secondary Structure of each location)

		%A	%C	%D	%E	%F	%G	%H	%I	%K	%L	%M	%N	%P	%Q	%R	%S	%T	%V	%W	%Y
Overall	**Helix**	13.61	0.95	4.98	8.44	3.35	3.57	2.34	5.04	13.96	4.51	2.45	2.31	2.33	5.65	8.14	4.68	3.83	6.17	1.26	2.45
	Coil	8.36	1.13	8.12	5.61	2.72	12.52	2.87	2.96	6.79	4.11	2.22	4.40	8.32	3.92	5.91	7.25	5.46	4.57	0.84	1.95
	Strand	6.76	1.67	2.75	4.47	5.59	3.61	2.54	10.66	13.58	3.02	2.25	1.64	1.47	3.09	5.83	4.17	5.41	16.19	1.56	3.75
Cytoplasm	**Helix**	13.46	1.00	4.96	8.67	3.35	3.54	2.41	5.10	14.11	4.29	2.43	2.23	2.31	5.60	8.42	4.51	3.73	6.19	1.25	2.45
	Coil	8.11	1.23	8.10	5.82	2.76	12.57	3.01	3.02	6.96	3.95	2.28	4.17	8.30	3.85	6.19	7.02	5.33	4.57	0.83	1.94
	Strand	6.70	1.78	2.76	4.58	5.46	3.57	2.61	10.82	13.78	2.90	2.21	1.54	1.46	3.04	6.00	4.07	5.17	16.36	1.52	3.67
Periplasm	**Helix**	15.45	0.50	5.28	6.46	3.52	3.89	1.33	4.58	12.51	7.11	2.53	2.78	2.85	5.73	5.35	5.41	4.38	5.97	1.66	2.72
	Coil	9.84	0.51	8.53	4.59	2.79	11.77	1.80	2.57	5.69	6.08	2.10	5.25	8.97	4.45	3.90	7.46	5.70	4.62	1.03	2.36
	Strand	7.52	0.83	2.39	3.93	6.49	4.25	1.57	9.12	11.48	4.38	2.69	2.28	1.86	3.37	4.40	4.79	6.78	15.60	2.15	4.12
Extracellular	**Helix**	14.37	0.51	4.89	6.24	3.03	3.83	2.30	4.38	12.64	5.71	2.76	3.30	1.96	6.49	6.03	7.31	5.20	5.99	0.94	2.09
	Coil	10.06	0.42	7.90	3.87	2.10	12.70	2.18	2.54	5.69	4.17	1.57	6.47	7.86	4.16	4.31	10.18	6.95	4.45	0.70	1.72
	Strand	6.75	0.78	3.09	3.05	6.84	3.38	2.51	9.74	12.70	3.40	2.40	2.73	0.93	3.59	4.59	5.37	8.31	13.46	1.43	4.96
HOPS	**Helix**	13.97	0.49	4.97	6.52	2.98	3.65	2.62	4.07	12.63	6.12	3.02	2.98	2.13	6.50	6.49	7.27	4.77	6.00	0.80	2.04
	Coil	9.60	0.35	7.01	4.08	1.98	12.14	2.63	2.72	5.70	4.48	1.66	6.44	8.84	4.36	4.78	10.92	6.31	4.14	0.53	1.33
	Strand	7.20	0.93	2.24	3.27	7.14	3.22	3.65	9.05	13.30	3.65	2.34	2.45	0.93	3.71	6.27	4.85	6.54	14.50	0.98	3.76
Extracellular - HOPs	**Helix**	15.60	0.56	4.66	5.42	3.17	4.40	1.35	5.32	12.69	4.49	1.98	4.23	1.45	6.48	4.66	7.44	6.48	5.98	1.39	2.25
	Coil	10.69	0.52	9.13	3.58	2.27	13.48	1.56	2.28	5.66	3.75	1.44	6.52	6.52	3.88	3.66	9.17	7.82	4.87	0.94	2.26
	Strand	6.46	0.68	3.66	2.91	6.64	3.48	1.76	10.19	12.31	3.23	2.44	2.91	0.93	3.52	3.48	5.71	9.47	12.77	1.72	5.74

Average Cost

		Filtered Database		Neural Networks		Unfiltered Database†	
		ATPs/aa	% total	ATPs/aa	% total	ATPs/aa	% total
Overall	**Helix**	23.549	40.58	23.691	43.08	22.993	34.90
	Coil	21.138	43.49	20.640	41.58	20.359	50.03
	Strand	26.629	15.93	27.397	15.34	26.658	15.07
Cytoplasm	**Helix**	23.610	41.35	23.752	44.02	ND	ND
	Coil	21.258	42.62	20.751	40.72	ND	ND
	Strand	26.625	16.03	27.429	15.26	ND	ND
Periplasm	**Helix**	23.282	35.95	23.500	36.97	23.287	35.98
	Coil	20.852	46.93	20.373	46.46	20.859	46.96
	Strand	26.590	17.12	27.487	16.57	26.577	17.06
Extracellular	**Helix**	22.601	33.15	22.626	34.584	22.597	33.55
	Coil	19.821	54.17	19.409	50.295	19.813	53.87
	Strand	26.785	12.68	26.713	15.121	26.796	12.58
HOPS	**Helix**	22.664	40.58	22.583	39.21	22.662	41.20
	Coil	19.819	51.19	19.657	50.25	19.809	50.68
	Strand	26.748	8.22	26.667	10.54	26.691	8.12
Extracellular - HOPs	**Helix**	22.413	21.42	22.724	27.28	22.402	21.48
	Coil	19.824	58.86	19.018	50.37	19.819	58.90
	Strand	26.810	19.72	26.746	22.36	26.864	19.62

Percent Cost of

		vs Cytoplasmic		vs Overall	
		Filtered	NN	Filtered	NN
Periplasm	**Helix**	98.61	98.94	98.87	99.19
	Coil	98.09	98.18	98.65	98.71
	Strand	99.87	100.21	99.85	100.33
Extracellular	**Helix**	95.73	95.26	95.97	95.50
	Coil	93.24	93.53	93.77	94.03
	Strand	100.60	97.39	100.59	97.50
HOPS	**Helix**	95.99	95.08	96.24	95.32
	Coil	93.23	94.73	93.76	95.24
	Strand	100.46	97.22	100.45	97.34
Extracellular - HOPs	**Helix**	94.93	95.67	95.18	95.92
	Coil	93.25	91.65	93.78	92.14
	Strand	100.69	97.51	100.68	97.62

Notes: Values on this page are based on the total amino acids in all of that secondary structure of every protein in that location. Values in the table on the page above are from averaging % secondary structure in each protein. ND – not done. NN – Neural Networks.

† Overall contains only data from Periplasmic and Extracellular proteins

Table S23: Disorder prediction and ASC in nonmembrane proteins of *E. coli* and *P. syringae*.

E. coli Mean values vased on per protein amounts (unweighted for length)

	Count	Mass/aa	ATPs/aa	Length	% Ordered	% Disordered
Overall	3212	111.38	23.26	297.75	94.25	5.75
Cytoplasmic	2859	111.61	23.32	296.47	94.50	5.50
Periplasmic	337	109.72	22.85	309.42	92.39	7.61
Extracellular	16	104.90	20.65	280.44	88.81	11.19

E. coli Linear Correlation of ASC (ATPs/aa) and…

	% Ordered		% Disordered	
	R	P value	R	P value
Overall	0.1122	1.81E-10	-0.1122	1.81E-10
Cytoplasmic	0.0719	1.20E-04	-0.0719	1.20E-04
Periplasmic	0.2597	1.35E-06	-0.2597	1.35E-06
Extracellular	0.3342	**0.2058**	-0.3342	**0.2058**

P. syringae Mean values vased on per protein amounts (unweighted for length)

	Count	Mass/aa	ATPs/aa	Length	% Ordered	% Disordered
Overall	2167	110.72	22.89	327.13	93.70	6.30
Cytoplasmic	1973	110.98	22.96	317.70	94.07	5.93
Periplasmic	110	108.72	22.65	414.95	95.62	4.38
Extracellular	84	107.30	21.58	433.73	82.63	17.37
HOPS&Helpers†	51	107.68	21.48	437.31	77.13	22.87
Extracellular-HOPS	33	106.70	21.74	428.18	91.14	8.86

P. syringae Linear Correlation of ASC (ATPs/aa) and…

	% Ordered		% Disordered	
	R	P value	R	P value
Overall	0.1762	1.43E-16	-0.1762	1.43E-16
Cytoplasmic	0.0648	0.0040	-0.0648	0.0040
Periplasmic	0.4273	3.23E-06	-0.4273	3.23E-06
Extracellular	0.4987	1.38E-06	-0.4987	1.38E-06
HOPS&Helpers†	0.5896	5.29E-06	-0.5896	5.29E-06
Extracellular-HOPS	0.4271	**0.0132**	-0.4271	**0.0132**

E. coli Comparing composition (Filtered Database)

	Student *t**		Mann Whitney	
vs Cytoplasmic	T value	P value	Z	P value
Periplasmic				
% Ordered	3.518	4.87E-04	4.337	1.44E-05
% Disordered	3.518	4.87E-04	4.337	1.45E-05
Extracellular				
% Ordered	1.020	**0.324**	0.099	**0.921**
% Disordered	1.020	**0.324**	0.099	**0.921**

P. syringae Comparing composition (Filtered database)

	Student *t**		Mann Whitney	
vs Cytoplasmic	T value	P value	Z	P value
Periplasmic				
% Ordered	1.865	**0.0647**	3.989	6.62E-05
% Disordered	1.865	**0.0647**	3.990	6.62E-05
Extracellular				
% Ordered	5.814	1.07E-07	6.836	8.15E-12
% Disordered	5.814	1.07E-07	6.836	8.16E-12
HOPS†				
% Ordered	8.029	1.31E-10	9.615	6.93E-22
% Disordered	8.029	1.31E-10	9.615	6.93E-22
Extracellular - HOPS				
% Ordered	0.884	**0.3832**	0.962	**0.3360**
% Disordered	0.884	**0.3832**	0.962	**0.3358**

Notes: **BOLD –** P > 0.01 * Assumes unequal variance, two tailed.

† The N terminus (non cleaved signal sequence) of Type III effectors is typically disordered (29); HOPS have higher % disorder in the rest of their proteins. Removing the first 30 aa from the HOP proteins changed the % of ordered and disordered regions to 81.36% and 18.64% respectively.

Table S23: Disorder prediction and ASC in nonmembrane proteins of *E. coli* and *P. syringae* (continued).

Filtered Database (Percent of all amino acids in the (dis)ordered region of each location)

E. coli

		%A	%C	%D	%E	%F	%G	%H	%I	%K	%L	%M	%N	%P	%Q	%R	%S	%T	%V	%W	%Y	ASC	% total
Overall	**Disordered**	9.82	0.44	4.96	7.93	1.42	5.54	2.43	2.71	5.29	8.16	6.98	4.77	5.78	6.87	7.14	8.17	5.95	4.17	0.35	1.12	21.326	4.29
	Ordered	9.36	1.32	5.71	6.48	3.58	7.20	2.55	5.75	10.20	4.58	2.45	3.99	4.45	4.55	5.87	5.47	5.29	6.91	1.39	2.90	23.413	95.71
Cytoplasm	**Disordered**	9.72	0.47	4.97	8.35	1.37	5.46	2.49	2.75	5.23	7.93	6.95	4.65	5.77	6.77	7.53	8.08	5.81	4.20	0.35	1.14	21.344	4.07
	Ordered	9.30	1.38	5.67	6.71	3.59	7.18	2.66	5.83	10.34	4.45	2.48	3.83	4.42	4.50	6.08	5.34	5.15	6.90	1.36	2.83	23.459	95.93
Periplasm	**Disordered**	10.23	0.32	4.82	5.75	1.74	6.11	2.21	2.44	5.41	9.71	7.39	5.48	5.70	7.34	5.30	8.42	6.23	3.95	0.35	1.11	21.338	5.78
	Ordered	9.81	0.88	6.03	4.69	3.54	7.35	1.68	5.07	9.08	5.66	2.22	5.11	4.82	4.93	4.32	6.40	6.28	6.97	1.69	3.45	23.133	94.22
Extracellular	**Disordered**	11.31	0.23	6.11	4.52	1.36	3.62	0.45	2.49	8.14	5.20	4.07	4.30	7.92	7.47	2.26	11.76	12.90	5.20	0.45	0.23	19.763	9.85
	Ordered	11.79	0.40	6.08	2.89	2.74	8.33	0.59	5.22	8.68	4.23	1.31	7.54	2.08	6.38	2.30	9.02	10.14	7.05	0.54	2.69	20.780	90.15

P. syringae

		%A	%C	%D	%E	%F	%G	%H	%I	%K	%L	%M	%N	%P	%Q	%R	%S	%T	%V	%W	%Y	ASC	% total
Overall	**Disordered**	11.95	0.40	5.91	7.22	1.39	6.17	2.21	2.42	5.93	5.39	5.23	3.58	8.27	5.83	7.63	9.08	5.47	4.84	0.26	0.82	20.540	5.73
	Ordered	10.13	1.18	5.99	6.54	3.56	7.55	2.63	5.19	11.07	4.02	2.14	3.08	4.58	4.41	6.75	5.51	4.75	7.20	1.18	2.54	23.140	94.27
Cytoplasm	**Disordered**	11.47	0.48	6.00	7.78	1.40	5.96	2.30	2.39	6.11	5.31	5.92	3.39	7.91	5.51	8.25	8.49	5.37	4.92	0.27	0.78	20.811	5.00
	Ordered	10.02	1.26	5.94	6.75	3.55	7.47	2.72	5.27	11.27	3.85	2.14	2.93	4.56	4.39	7.02	5.35	4.60	7.25	1.16	2.52	23.211	95.00
Periplasm	**Disordered**	15.14	0.12	6.11	6.35	1.72	6.17	1.17	2.56	5.39	6.98	3.94	2.83	10.90	7.34	4.85	7.28	4.73	5.18	0.21	1.02	20.090	7.28
	Ordered	11.17	0.60	6.33	5.06	3.84	7.77	1.62	4.56	9.43	6.09	2.23	3.93	5.13	4.52	4.48	6.19	5.46	7.12	1.55	2.93	22.914	92.72
Extracellular	**Disordered**	12.71	0.13	5.33	4.78	1.14	7.30	2.30	2.52	5.28	4.91	2.32	5.05	8.69	6.66	5.89	13.17	6.44	4.24	0.24	0.89	19.364	16.32
	Ordered	10.75	0.56	6.48	4.51	3.38	8.83	2.25	4.36	9.58	4.52	2.02	4.92	4.31	4.51	4.72	7.73	6.56	6.47	1.00	2.52	22.067	83.68
HOPS	**Disordered**	12.03	0.14	4.85	4.94	1.12	7.76	2.52	2.62	5.24	4.90	2.03	4.96	9.04	6.82	6.24	13.35	6.40	3.90	0.28	0.85	19.417	22.76
	Ordered	10.92	0.55	6.07	5.02	3.31	8.02	2.76	4.13	10.29	5.13	2.33	4.63	4.41	4.69	5.40	7.64	5.50	6.29	0.80	2.11	22.170	77.24
Extracellular - HOPs	**Disordered**	16.69	0.12	8.17	3.80	1.27	4.60	1.04	1.96	5.52	4.95	4.03	5.52	6.67	5.75	3.80	12.08	6.67	6.21	0.00	1.15	19.057	6.15
	Ordered	10.53	0.59	7.02	3.85	3.46	9.89	1.59	4.66	8.67	3.73	1.61	5.30	4.18	4.28	3.84	7.86	7.93	6.70	1.27	3.06	21.933	93.85

Percent Cost of		***E. coli***	
		vs Cytoplasmic	**vs Overall**
Periplasm	**Disordered**	99.98	100.06
	Ordered	98.61	98.81
Extracellular	**Disordered**	92.59	92.67
	Ordered	88.58	88.76

Percent Cost of		***P. syringae***	
		vs Cytoplasmic	**vs Overall**
Periplasm	**Disordered**	96.54	97.81
	Ordered	98.72	99.02
Extracellular	**Disordered**	93.05	94.27
	Ordered	95.07	95.36
HOPS	**Disordered**	93.30	94.53
	Ordered	95.52	95.81
Extracellular - HOPs	**Disordered**	91.57	92.78
	Ordered	94.49	94.78

Notes: Values on this page are based on the total amino acids in all the (dis)ordered region of every protein in that location. Values in the table on the page above are from averaging the % (dis)order in each protein.

Table S24: Amino acid composition of extracellular flagellar subunits in endosymbionts and other *Gammaproteobacteria*.

	Protein	Insect Host		Presence of Amino Acid Pathway – Source (100, 111)																			
Symbiont	**Sequence**	**Name**	**Nutrition**	**A**	**R**	**N**	**D**	**C**	**E**	**Q**	**G**	**H**	**I**	**L**	**K**	**M**	**F**	**P**	**S**	**T**	**W**	**Y**	**V**
Wigglesworthia glossinidia (4)	PSORTDB	*Glossina brevipalpis*	blood				P		Y	Y	Y				P					P			
Sodalis glossinidius str. 'morsitans' (100)	NCBI	*Glossina morsitans morsitans*	blood		Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Buchnera aphidicola str. APS (89)	PSORTDB	*Acyrthosiphon pisum*	phloem		Y			Y			Y	Y	Y	Y	Y		Y			Y	Y		Y
Buchnera aphidicola str. Sg (98)	NCBI	*Schizaphis graminum*	phloem		Y			Y			Y	Y	Y	Y	Y		Y			Y	Y		Y

P=Some genes in pathway are present

			Mean Protein Values																							
	Proteins	**%GC**	**Length**	**Mass/aa**	**ATPs/aa**	**%A**	**%R**	**%N**	**%D**	**%C**	**%E**	**%Q**	**%G**	**%H**	**%I**	**%L**	**%K**	**%M**	**%F**	**%P**	**%S**	**%T**	**%W**	**%Y**	**%V**
W. glossinidia	611	22.50	330.1	114.979	25.585	3.6	3.3	8.4	3.8	1.3	5.0	2.3	5.0	1.6	13.3	9.5	12.1	2.1	5.5	2.7	7.6	3.7	0.8	3.9	4.2
S. glossinidius	2515	54.51	287.1	110.743	23.261	10.1	6.6	3.4	5.3	1.3	5.4	4.7	7.2	2.4	5.6	10.9	4.2	2.8	3.6	4.5	5.8	5.3	1.3	2.8	6.9
B. aphidicola APS	574	26.35	328.0	114.303	25.027	4.5	4.0	7.1	4.1	1.2	5.4	3.2	5.4	2.1	11.7	9.8	10.3	2.3	5.0	2.9	7.2	4.6	0.9	3.5	4.9
B. aphidicola SG	545	25.00	326.2	114.531	25.192	4.4	3.9	7.2	4.0	1.2	5.4	3.0	5.3	2.0	11.8	9.8	11.0	2.2	5.3	2.9	7.0	4.4	0.8	3.5	4.8
E. coli K-12	4333	50.80	312.5	111.005	23.555	9.4	5.6	3.9	5.0	1.3	5.8	4.4	7.1	2.3	6.1	10.6	4.7	3.0	3.9	4.3	5.8	5.4	1.5	2.8	7.1

Amino Acids	**A**	**R**	**N**	**D**	**C**	**E**	**Q**	**G**	**H**	**I**	**L**	**K**	**M**	**F**	**P**	**S**	**T**	**W**	**Y**	**V**	
%aa in highly expressed genes (85)	↑	↑	↓		↓	↑		↑		↓	↓	↑		↓		↓†	↑†	↓	↓	↑	
Source of insect host nutrition																					**BOLD**
Vicia faba phloem (molar %)(109)	**3**	0.6	**68.7**	**3.4**		**3.7**	**6.7**	0.8	**3.3**	0.5	0.6	1	0.2	0.4		**4.4**	1	0.2	0.5	1	>3%
Triticum aestivum (>5mM)(84)	Y			Y		Y	Y		Y*	Y	Y	Y		Y		Y	Y			Y	
Triticum aestivum (>10mM)(84)				Y		Y	Y			Y	Y			Y		Y				Y	
Human whole blood (μM)(79)	**325**	nd	48	129	nd	**208**	**528**	**234**	58	51	114	79	19	47	184	118	124	nd	63	**213**	>200μM
Human plasma (μM)(79)	**208**	nd	15	7	nd	35	**344**	**105**	40	37	80	69	15	34	**130**	64	82	nd	41	**146**	>100μM
Human serum (μM)(80)	**351**	91	59	11	10	99	**748**	**325**	88	89	**135**	**187**	8	63	**254**	**155**	**186**	73	70	**274**	>100μM
Mature human serum albumin (%aa)	**10.6**	4.1	2.9	6.2	6.0	**10.6**	3.4	2.1	2.7	1.4	**10.4**	**10.1**	1.0	5.3	4.1	4.1	4.8	0.2	3.1	7.0	>10%
%aa in ext. flagellar prot. **																					
Buchnera aphidicola str. APS	↓		↑		↑			↓	↑	↑		↑					↓			↓	
Buchnera aphidicola str. Sg	↓		↑		↑			↓	↑	↑		↑			↓		↓			↓	
Wigglesworthia glossinidia	↓		↑	↓		↑	↓	↓	↑	↑		↑	↓	↑	↓		↓		↑	↓	

Notes: Amino acid percents are on a per protein basis. The percent amino acid composition of each protein were averaged; they are unweighted for length.
Selected papers related to amino acid composition in phloem sap – (3, 24, 83, 84, 109), in human blood – (68, 69, 79, 80)

Legend: nd – not done; † only in *W. glossinidia*; * Histidine in stylus from *B. aphidicola* str. Sg feeding on *T. aestivium*is is ~4.9mM (Adobe Photoshop used to measure graph in (84))
** Trend in amino acid composition in at least 3 of 4 (*Buchnera)* or 4 of 7 (*Wigglesworthia*) in flagellar proteins compared to other *Gammaproteobacteria* (see continued table below)

Table S24: Amino acid composition of extracellular flagellar subunits in endosymbionts and other *Gammaproteobacteria* (continued).

		Homologs		Mean Values		Economic Rank **																					
		Total	Size*	Mass/aa	ATPs/aa	Organism	Homologs	%A	%R	%N	%D	%C	%E	%Q	%G	%H	%I	%L	%K	%M	%F	%P	%S	%T	%W	%Y	%V
FliC	γ Proteobacteria	544	516	104.682	19.662			13.1	3.8	8.2	6.5	0.0	4.5	6.5	8.0	0.4	5.8	8.5	4.1	2.5	2.0	0.7	10.0	8.7	0.1	1.1	5.5
	W. glossinidia			109.134	21.100	4/611	514/516	8.8	4.0	8.8	6.7	0.0	6.2	5.5	5.2	0.5	9.5	8.3	7.1	1.4	2.1	1.0	11.9	4.8	0.0	**2.6**	5.7
	S. glossinidius			106.962	20.633	48/2515	489/516	10.7	3.3	6.9	6.7	0.0	6.9	5.7	8.1	**0.0**	9.1	7.2	7.4	**1.0**	1.9	1.4	7.9	8.1	0.0	2.6	5.0
FliD	γ Proteobacteria	284	277	104.682	19.662			10.2	2.9	5.8	7.6	0.0	4.9	5.3	7.5	0.4	5.4	8.9	6.4	2.1	2.5	1.8	10.3	10.2	0.2	1.9	5.8
	W. glossinidia			116.578	24.725	185/611	277/277	**1.6**	2.6	**13.5**	4.2	**0.4**	7.9	**2.0**	4.0	**1.0**	**12.1**	9.7	**15.8**	1.6	**5.3**	1.6	8.7	**2.2**	**0.4**	3.2	**2.4**
	S. glossinidius			108.020	20.970	92/2515	224/277	9.1	3.1	5.6	6.0	0.0	7.6	5.1	7.2	**1.0**	6.2	9.9	8.6	2.5	1.4	2.3	9.7	8.6	**0.4**	1.9	3.9
FliK	γ Proteobacteria	208	160	104.442	19.734			14.1	2.7	3.8	6.2	0.0	5.3	8.2	6.4	2.0	3.2	10.6	4.1	1.9	1.8	6.4	9.6	7.7	0.4	0.4	4.9
	W. glossinidia			120.408	25.627	328/611	155/160	**0.5**	**0.2**	**17.0**	9.0	**0.5**	7.8	**1.9**	**1.0**	**0.5**	**11.2**	8.8	**16.5**	**0.5**	**8.0**	**0.2**	6.6	**1.7**	**0.0**	**5.8**	**2.2**
	S. glossinidius			105.888	20.726	57/2515	151/160	14.2	5.4	2.4	5.6	**0.5**	3.8	8.6	7.5	**4.3**	3.5	11.8	**1.9**	1.1	1.3	7.3	5.9	7.8	0.5	**1.3**	5.1
	B. aphidicola APS			118.654	26.538	479/574	160/160	**0.9**	**0.4**	**11.9**	4.3	**0.4**	4.7	**1.3**	**0.9**	3.0	**14.0**	7.7	**20.4**	2.1	**4.7**	**2.1**	9.8	**3.4**	**0.9**	**4.7**	2.6
	B. aphidicola SG			119.464	26.081	408/545	158/160	**1.3**	2.4	**15.5**	4.2	**1.0**	4.7	**3.1**	**1.3**	2.4	**10.8**	7.1	**17.6**	1.8	**8.4**	**1.3**	7.3	3.9	0.3	**3.1**	**2.4**
FlgD	γ Proteobacteria	232	232	104.719	20.362			8.2	1.7	6.6	5.7	0.0	4.6	7.0	9.6	0.6	4.7	9.3	4.3	2.8	2.5	3.5	10.1	8.9	0.4	1.4	8.0
	W. glossinidia			115.399	23.758	81/611	232/232	**1.4**	1.8	**16.1**	4.7	0.0	7.2	4.3	**3.9**	**1.4**	**11.8**	7.5	**11.5**	2.2	**5.7**	2.5	6.8	6.5	0.4	1.8	**2.5**
	S. glossinidius			107.147	21.125	15/2515	213/232	11.5	2.6	7.8	6.3	**0.4**	4.8	5.6	7.1	0.7	3.0	9.3	6.3	4.1	2.6	4.1	7.8	5.6	0.4	2.6	7.4
	B. aphidicola APS			110.889	22.048	8/574	230/232	**3.4**	1.3	**14.0**	5.9	0.0	3.4	7.2	**4.2**	**1.3**	**12.3**	8.5	8.5	2.1	2.5	4.7	8.9	5.9	0.4	0.8	4.7
	B. aphidicola SG			109.811	22.002	8/545	228/232	**3.5**	1.3	12.4	4.4	0.0	4.0	6.2	4.9	**2.2**	**14.2**	8.0	8.4	2.2	**0.9**	2.2	9.7	8.0	0.4	0.9	6.2
FlgE	γ Proteobacteria	260	253	105.108	20.884			9.5	2.2	9.1	5.9	0.0	3.1	5.8	10.1	0.6	4.3	7.0	3.1	1.9	4.3	3.5	9.5	10.0	0.6	2.8	6.6
	W. glossinidia			108.765	22.121	13/611	250/253	**3.9**	1.2	11.2	4.9	**0.2**	4.4	6.1	9.0	**1.2**	**9.5**	6.3	**6.3**	2.7	4.4	2.7	12.9	**4.6**	0.5	2.9	4.9
	S. glossinidius			107.360	22.062	446/2515	249/253	7.9	1.5	7.7	5.4	**0.5**	4.5	6.2	11.1	**0.2**	5.2	6.7	**6.4**	3.2	4.5	4.2	7.7	7.4	0.7	4.0	5.0
	B. aphidicola APS			112.277	22.797	24/574	252/253	**2.9**	1.7	12.5	6.4	**0.5**	6.1	4.4	6.1	0.7	**9.1**	7.6	**8.1**	2.5	4.4	**1.7**	10.5	7.1	0.7	3.4	3.4
	B. aphidicola SG			112.295	23.113	43/545	253/253	**2.9**	1.5	12.2	6.6	**0.2**	5.6	5.1	6.1	**1.7**	**12.0**	9.3	**7.6**	1.2	5.1	2.0	10.3	**4.4**	0.5	2.7	**2.9**
FlgK	γ Proteobacteria	292	275	106.450	20.722			10.3	3.4	8.0	6.5	0.0	4.2	6.5	8.3	0.7	4.9	9.1	3.8	2.3	3.1	2.5	9.3	7.6	0.2	2.5	6.6
	W. glossinidia			115.250	24.132	118/611	275/275	**3.9**	1.8	12.8	5.3	0.0	**9.0**	**3.1**	4.4	**1.7**	**11.9**	8.3	**10.3**	**0.6**	**6.6**	**0.9**	9.7	**2.6**	0.2	4.4	**2.6**
	S. glossinidius			108.524	20.891	76/2515	202/275	8.5	5.0	9.5	6.8	0.0	5.6	6.1	9.0	1.1	4.7	10.4	4.1	**0.9**	2.5	2.0	7.9	6.7	0.2	3.4	5.6
	B. aphidicola APS			115.122	23.590	88/574	272/275	**4.2**	2.9	11.4	7.2	**0.9**	7.9	3.5	**2.8**	**1.7**	**11.0**	8.8	**9.8**	1.8	3.7	2.0	6.6	4.2	0.2	3.7	5.5
	B. aphidicola SG			116.360	24.132	140/545	274/275	**2.8**	2.6	11.0	8.3	**0.9**	7.7	3.5	**2.8**	**1.7**	**11.6**	8.6	**10.1**	1.7	4.2	1.5	6.8	5.0	0.4	4.4	4.6
FlgL	γ Proteobacteria	264	238	108.528	21.057			9.6	3.7	6.9	7.3	0.0	5.2	7.4	6.6	0.8	4.9	8.9	4.3	2.9	3.3	2.3	10.0	7.0	0.3	2.6	5.9
	W. glossinidia			115.438	23.440	58/611	238/238	**1.0**	3.2	12.8	**1.9**	**1.0**	9.3	4.8	**2.9**	**1.6**	**10.2**	12.1	**8.9**	2.2	3.5	1.3	13.1	**3.5**	0.3	3.2	3.2
	S. glossinidius			107.083	21.309	146/2515	176/238	11.2	3.8	4.8	8.0	0.0	3.5	7.3	9.6	1.3	5.1	9.3	4.8	4.2	2.6	1.9	7.3	5.4	0.3	2.9	6.7

BOLD = Amino acid % is 2X or 1/2 that of same protein average in γ Proteobacteria

Notes: * Homologs whose lengths were 50% (+/-) their respective *E. coli* counterpart were removed. Comparisons were done on the remaining homologs.

** Economic Rank in ATPs/aa (least expensive/total) amongst homologs and within each organism.

Amino acid percents in γproteobacteria are on a per protein basis. The amino acid compositions of each protein are averaged and unweighted for length.

Buchnera aphidicola strains lack FliA, FliC, FliD, FliL, FliO, FliS, FliT, FlgM and FlgL (56).

Table S26: Comparison of the ASC of extracellular protein with total proteins in 26 Gram-negative bacteria.*

				Total Proteins		Extracellular Proteins							U test (ASC)		% Savings	
Organism	**Group**	**% GC**	**Number Proteins**	**ASC per aa**	**Mass per aa**	**#**	**# ASC < Org. ASC**	**% ASC < Org. ASC**	**ASC per aa**	**Mass per aa**	**Avgerage Percentile**	**ATP Savings**	**vs Cyto. Proteins**	**vs All Proteins**	**ASC per aa**	**Mass per aa**
Arcobacter butzleri RM4018	Epsilonproteobacteria	27	2259	24.608	113.851	13	12	92.31	21.736	108.451	90	2.87	1.75E-07	1.12E-07	11.67	4.74
Flavobacterium johnsoniae UW101	Bacteroidetes/Chlorobi	34.11	5017	24.443	113.337	35	29	82.86	23.245	109.862	74.5	1.2	3.27E-05	5.72E-07	4.9	3.07
Trichodesmium erythraeum IMS101	Cyanobacteria	34.14	4451	24.086	112.993	46	46	100	21.195	106.929	93	2.89	1.75E-22	9.50E-24	12	5.37
Colwellia psychrerythraea 34H	Gammaproteobacteria	38	4910	23.864	111.778	37	31	83.78	21.894	107.82	80.3	1.97	1.14E-07	2.17E-10	8.26	3.54
Cytophaga hutchinsonii ATCC 33406	Bacteroidetes/Chlorobi	38.8	3785	24.53	112.817	40	39	97.50	22.602	106.967	85.2	1.93	2.28E-12	1.67E-14	7.86	5.19
Acinetobacter sp. ADP1	Gammaproteobacteria	40.4	3325	23.947	112.141	11	10	90.91	22.095	106.614	81	1.85	1.35E-03	3.80E-04	7.73	4.93
Photorhabdus luminescens laumondii TTO1	Gammaproteobacteria	42.8	4683	23.863	112.161	33	31	93.94	22.02	108.192	83.5	1.84	2.23E-10	3.11E-11	7.72	3.54
Aquifex aeolicus VF5[1]	Aquificae	43.3	1560	24.774	114.13	11	10	90.91	23.214	111.694	81.8	1.56	7.80E-04	2.78E-04	6.3	2.13
Shewanella oneidensis MR-1	Gammaproteobacteria	45.9	4471	23.54	111.158	36	34	94.44	21.357	106.088	89.1	2.18	4.85E-16	6.09E-16	9.27	4.56
Synechocystis sp. PCC 6803	Cyanobacteria	47.3	3564	23.583	111.507	12	10	83.33	22.151	107.61	77.7	1.43	2.50E-03	9.17E-04	6.07	3.49
Vibrio cholerae O1 biovar eltor N16961	Gammaproteobacteria	47.5	3835	23.759	111.782	50	45	90.00	22.278	108.608	77.3	1.48	8.97E-09	2.96E-11	6.23	2.84
Wolinella succinogenes DSM 1740	Epsilonproteobacteria	48.5	2044	24.042	112.206	13	13	100	21.439	107.44	93.7	2.6	3.40E-08	5.48E-08	10.83	4.25
Escherichia coli K-12 MG1655	Gammaproteobacteria	50.8	4333	23.555	111.005	16	16	100	20.652	104.9	93.4	2.9	2.19E-09	1.96E-09	12.32	5.5
Salmonella typhimurium LT2	Gammaproteobacteria	52.2	4527	23.532	110.908	58	51	87.93	21.676	107.196	81.4	1.86	1.43E-14	2.22E-16	7.89	3.35
Magnetococcus MC-1	Other	54.2	3716	23.277	110.848	37	37	100	20.261	105.888	94.2	3.02	1.08E-20	2.00E-20	12.96	4.47
Rhodopirellula baltica SH 1	Plantomycetes	55.4	7325	23.011	110.75	24	21	87.50	20.872	106.693	85.9	2.14	1.54E-09	1.15E-09	9.3	3.66
Marinobacter aquaeolei VT8	Gammaproteobacteria	57	4272	23.077	110.776	19	19	100	20.671	106.006	93.2	2.41	1.95E-11	7.50E-11	10.43	4.31
Pseudomonas syringae pv. tomato DC3000	Gammaproteobacteria	58.3	5608	23.047	110.075	84	70	83.33	21.581	107.296	77.1	1.47	9.94E-18	1.25E-17	6.36	2.52
Silicibacter sp. TM1040	Alphaproteobacteria	60.1	3864	22.754	109.053	29	29	100	20.397	104.447	92.7	2.36	4.34E-16	2.22E-15	10.36	4.22
Roseiflexus castenholzii DSM 13941	Chloroflexi	60.7	4330	23.072	109.936	22	22	100	21.316	105.67	89.3	1.76	1.13E-10	2.02E-10	7.61	3.88
Burkholderia xenovorans LB400	Betaproteobacteria	62.6	8702	22.762	108.813	45	39	86.67	21.063	103.631	79.1	1.7	5.80E-12	1.45E-11	7.46	4.76
Ralstonia eutropha JMP134	Betaproteobacteria	64.5	6446	22.648	108.437	30	27	90.00	20.62	104.01	86.3	2.03	2.17E-12	6.63E-12	8.95	4.08
Pseudomonas aeruginosa PAO1	Gammaproteobacteria	66.6	5667	22.881	109.787	48	41	85.42	21.392	105.903	79.2	1.49	2.89E-12	3.56E-12	6.51	3.54
Caulobacter cresentus CB15	Alphaproteobacteria	67.2	3737	22.423	107.884	27	22	81.48	20.955	103.973	79	1.47	2.95E-07	1.95E-07	6.55	3.63
Rhodobacter sphaeroides 2.4.1	Alphaproteobacteria	68.8	4242	22.481	108.15	29	29	100	20.251	102.962	90.1	2.23	7.22E-15	9.25E-14	9.92	4.8
Anaeromyxobacter dehalogenans 2CP-C	Deltaproteobacteria	74.9	4346	21.913	106.92	28	27	96.43	20.249	102.949	84.4	1.66	2.54E-12	2.00E-10	7.59	3.71

* Each is capable of aerobic growth and has the capacity to synthesize all 20 amino acids (See notes in Table S25 – Tab H in Data Set S1).

[1] *Aquifex aeolicus* VF5 was not graphed in Fig. 5B and was not used in later analysis. Many hyperthermophiles (>96°C) do not follow the %GC cost trends of other organisms (See Fig. 5A) possibly due to different amino acid biases needed for protein folding at higher temperatures. Many thermophiles display weak cost selection relative to mesophiles (97), perhaps because the Gibbs free energy of amino acid formation is radically different in hydrothermal vents (Table S1) (6).

References

1. http://redpoll.pharmacy.ualberta.ca/CCDB/cgi-bin/STAT_NEW.cgi.
2. **Akashi, H., and T. Gojobori.** 2002. Metabolic efficiency and amino acid composition in the proteomes of Escherichia coli and Bacillus subtilis. Proc Natl Acad Sci U S A **99:**3695-700.
3. **Akman Gunduz, E., and A. E. Douglas.** 2009. Symbiotic bacteria enable insect to use a nutritionally inadequate diet. Proc Biol Sci **276:**987-91.
4. **Akman, L., A. Yamashita, H. Watanabe, K. Oshima, T. Shiba, M. Hattori, and S. Aksoy.** 2002. Genome sequence of the endocellular obligate symbiont of tsetse flies, Wigglesworthia glossinidia. Nat Genet **32:**402-7.
5. **Alm, R. A., U. H. Stroeher, and P. A. Manning.** 1988. Extracellular proteins of Vibrio cholerae: nucleotide sequence of the structural gene (hlyA) for the haemolysin of the haemolytic El Tor strain 017 and characterization of the hlyA mutation in the non-haemolytic classical strain 569B. Mol Microbiol **2:**481-8.
6. **Amend, J. P., and E. L. Shock.** 1998. Energetics of amino acid synthesis in hydrothermal ecosystems. Science **281:**1659-62.
7. **Arnold, R., S. Brandmaier, F. Kleine, P. Tischler, E. Heinz, S. Behrens, A. Niinikoski, H. W. Mewes, M. Horn, and T. Rattei.** 2009. Sequence-based prediction of type III secreted proteins. PLoS Pathog **5:**e1000376.
8. **Atkinson, D. E.** 1977. Cellular Energy Metabolism and Its Regulation, vol. Academic Press, New York.
9. **Baar, C., M. Eppinger, G. Raddatz, J. Simon, C. Lanz, O. Klimmek, R. Nandakumar, R. Gross, A. Rosinus, H. Keller, P. Jagtap, B. Linke, F. Meyer, H. Lederer, and S. C. Schuster.** 2003. Complete genome sequence and analysis of Wolinella succinogenes. Proc Natl Acad Sci U S A **100:**11690-5.
10. **Bauchop, T., and S. R. Elsden.** 1960. The growth of micro-organisms in relation to their energy supply. J Gen Microbiol **23:**457-69.
11. **Berg, H. C.** 2000. Constraints on models for the flagellar rotary motor. Philos Trans R Soc Lond B Biol Sci **355:**491-501.
12. **Berg, H. C.** 1974. Dynamic properties of bacterial flagellar motors. Nature **249:**77-9.
13. **Berg, H. C.** 2003. The rotary motor of bacterial flagella. Annu Rev Biochem **72:**19-54.
14. **Berg, H. C., and L. Turner.** 1993. Torque generated by the flagellar motor of Escherichia coli. Biophys J **65:**2201-16.
15. **Beskhlebnaya, V. A., E. V. Trinchina, P. Aprikyan, V. Chesnokova, and E. V. Sokurenko.** 2006. Molecular genetic analysis of Escherichia coli type I adhesins. Bull Exp Biol Med **141:**339-42.
16. **Bingle, L. E., C. M. Bailey, and M. J. Pallen.** 2008. Type VI secretion: a beginner's guide. Curr Opin Microbiol **11:**3-8.
17. **Bragg, J. G., and A. Wagner.** 2007. Protein carbon content evolves in response to carbon availability and may influence the fate of duplicated genes. Proc Biol Sci **274:**1063-70.
18. **Capitani, G., O. Eidam, R. Glockshuber, and M. G. Grutter.** 2006. Structural and functional insights into the assembly of type 1 pili from Escherichia coli. Microbes Infect **8:**2284-90.

19. **Chen, X., and H. C. Berg.** 2000. Torque-speed relationship of the flagellar rotary motor of Escherichia coli. Biophys J **78:**1036-41.
20. **Chevance, F. F., and K. T. Hughes.** 2008. Coordinating assembly of a bacterial macromolecular machine. Nat Rev Microbiol **6:**455-65.
21. **Craig, C. L., and R. S. Weber.** 1998. Selection costs of amino acid substitutions in ColE1 and ColIa gene clusters harbored by Escherichia coli. Mol Biol Evol **15:**774-6.
22. **DeRosier, D. J.** 1995. Spinning tails. Curr Opin Struct Biol **5:**187-93.
23. **DeRosier, D. J.** 1998. The turn of the screw: the bacterial flagellar motor. Cell **93:**17-20.
24. **Douglas, A. E.** 2006. Phloem-sap feeding by animals: problems and solutions. J Exp Bot **57:**747-54.
25. **Duong, F., E. Bonnet, V. Geli, A. Lazdunski, M. Murgier, and A. Filloux.** 2001. The AprX protein of Pseudomonas aeruginosa: a new substrate for the Apr type I secretion system. Gene **262:**147-53.
26. **Farmer, I. S., and C. W. Jones.** 1976. The energetics of Escherichia coli during aerobic growth in continuous culture. Eur J Biochem **67:**115-22.
27. **Feist, A. M., C. S. Henry, J. L. Reed, M. Krummenacker, A. R. Joyce, P. D. Karp, L. J. Broadbelt, V. Hatzimanikatis, and B. O. Palsson.** 2007. A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. Mol Syst Biol **3:**121.
28. **Gabel, C. V., and H. C. Berg.** 2003. The speed of the flagellar rotary motor of Escherichia coli varies linearly with protonmotive force. Proc Natl Acad Sci U S A **100:**8748-51.
29. **Galan, J. E., and H. Wolf-Watz.** 2006. Protein delivery into eukaryotic cells by type III secretion machines. Nature **444:**567-73.
30. **Gohl, O., A. Friedrich, M. Hoppert, and B. Averhoff.** 2006. The thin pili of Acinetobacter sp. strain BD413 mediate adhesion to biotic and abiotic surfaces. Appl Environ Microbiol **72:**1394-401.
31. **Guo, M., F. Tian, Y. Wamboldt, and J. R. Alfano.** 2009. The majority of the type III effector inventory of Pseudomonas syringae pv. tomato DC3000 can suppress plant immunity. Mol Plant Microbe Interact **22:**1069-80.
32. **Haddix, P. L., S. Jones, P. Patel, S. Burnham, K. Knights, J. N. Powell, and A. LaForm.** 2008. Kinetic analysis of growth rate, ATP, and pigmentation suggests an energy-spilling function for the pigment prodigiosin of Serratia marcescens. J Bacteriol **190:**7453-63.
33. **Hahn, E., P. Wild, U. Hermanns, P. Sebbel, R. Glockshuber, M. Haner, N. Taschner, P. Burkhard, U. Aebi, and S. A. Muller.** 2002. Exploring the 3D molecular architecture of Escherichia coli type 1 pili. J Mol Biol **323:**845-57.
34. **Hasegawa, K., I. Yamashita, and K. Namba.** 1998. Quasi- and nonequivalence in the structure of bacterial flagellar filament. Biophys J **74:**569-75.
35. **Hauser, A. R.** 2009. The type III secretion system of Pseudomonas aeruginosa: infection by injection. Nat Rev Microbiol **7:**654-65.
36. **Heizer, E. M., Jr., D. W. Raiford, M. L. Raymer, T. E. Doom, R. V. Miller, and D. E. Krane.** 2006. Amino acid cost and codon-usage biases in 6 prokaryotic genomes: a whole-genome analysis. Mol Biol Evol **23:**1670-80.

37. **Hempfling, W. P., and S. E. Mainzer.** 1975. Effects of varying the carbon source limiting growth on yield and maintenance characteristics of Escherichia coli in continuous culture. J Bacteriol **123:**1076-87.
38. **Henderson, I. R., F. Navarro-Garcia, M. Desvaux, R. C. Fernandez, and D. Ala'Aldeen.** 2004. Type V protein secretion pathway: the autotransporter story. Microbiol Mol Biol Rev **68:**692-744.
39. **Horler, R. S., A. Butcher, N. Papangelopoulos, P. D. Ashton, and G. H. Thomas.** 2009. EchoLOCATION: an in silico analysis of the subcellular locations of Escherichia coli proteins and comparison with experimentally derived locations. Bioinformatics **25:**163-6.
40. **Hughes, K. T., K. L. Gillen, M. J. Semon, and J. E. Karlinsey.** 1993. Sensing structural intermediates in bacterial flagellar assembly by export of a negative regulator. Science **262:**1277-80.
41. **Iino, T.** 1969. Polarity of flagellar growth in salmonella. J Gen Microbiol **56:**227-39.
42. **Ishihama, Y., T. Schmidt, J. Rappsilber, M. Mann, F. U. Hartl, M. J. Kerner, and D. Frishman.** 2008. Protein abundance profiling of the Escherichia coli cytosol. BMC Genomics **9:**102.
43. **Jamir, Y., M. Guo, H. S. Oh, T. Petnicki-Ocwieja, S. Chen, X. Tang, M. B. Dickman, A. Collmer, and J. R. Alfano.** 2004. Identification of Pseudomonas syringae type III effectors that can suppress programmed cell death in plants and yeast. Plant J **37:**554-65.
44. **Jarrell, K. F., and M. J. McBride.** 2008. The surprisingly diverse ways that prokaryotes move. Nat Rev Microbiol **6:**466-76.
45. **Karp, P. D., I. M. Keseler, A. Shearer, M. Latendresse, M. Krummenacker, S. M. Paley, I. Paulsen, J. Collado-Vides, S. Gama-Castro, M. Peralta-Gil, A. Santos-Zavaleta, M. I. Penaloza-Spinola, C. Bonavides-Martinez, and J. Ingraham.** 2007. Multidimensional annotation of the Escherichia coli K-12 genome. Nucleic Acids Res **35:**7577-90.
46. **Khan, S., D. Pierce, and R. D. Vale.** 2000. Interactions of the chemotaxis signal protein CheY with bacterial flagellar motors visualized by evanescent wave microscopy. Curr Biol **10:**927-30.
47. **Krogfelt, K. A., and P. Klemm.** 1988. Investigation of minor components of Escherichia coli type 1 fimbriae: protein chemical and immunological aspects. Microb Pathog **4:**231-8.
48. **Kutsukake, K.** 1994. Excretion of the anti-sigma factor through a flagellar substructure couples flagellar gene expression with flagellar assembly in Salmonella typhimurium. Mol Gen Genet **243:**605-12.
49. **Kutsukake, K., and T. Iino.** 1994. Role of the FliA-FlgM regulatory system on the transcriptional control of the flagellar regulon and flagellar formation in Salmonella typhimurium. J Bacteriol **176:**3598-605.
50. **Li, N., J. Lv, and D. K. Niu.** 2009. Low contents of carbon and nitrogen in highly abundant proteins: evidence of selection for the economy of atomic composition. J Mol Evol **68:**248-55.
51. **Link, C., S. Eickernjager, D. Porstendorfer, and B. Averhoff.** 1998. Identification and characterization of a novel competence gene, comC, required for DNA binding and uptake in Acinetobacter sp. strain BD413. J Bacteriol **180:**1592-5.

52. **Liu, D., and P. R. Reeves.** 1994. Escherichia coli K12 regains its O antigen. Microbiology **140 (Pt 1):**49-57.
53. **Lowe, G., Meister, M., & Berg, H.C.** 1987. Rapid rotation of flagellar bundles in swimming bacteria. Nature **325:**637 - 640.
54. **Lowe, M. A., S. C. Holt, and B. I. Eisenstein.** 1987. Immunoelectron microscopic analysis of elongation of type 1 fimbriae in Escherichia coli. J Bacteriol **169:**157-63.
55. **Macnab, R. M.** 2003. How bacteria assemble flagella. Annu Rev Microbiol **57:**77-100.
56. **Maezawa, K., S. Shigenobu, H. Taniguchi, T. Kubo, S. Aizawa, and M. Morioka.** 2006. Hundreds of flagellar basal bodies cover the cell surface of the endosymbiotic bacterium Buchnera aphidicola sp. strain APS. J Bacteriol **188:**6539-43.
57. **Masuda, T., N. Saito, M. Tomita, and Y. Ishihama.** 2009. Unbiased quantitation of Escherichia coli membrane proteome using phase-transfer surfactants. Mol Cell Proteomics.
58. **McGuffin, L. J., K. Bryson, and D. T. Jones.** 2000. The PSIPRED protein structure prediction server. Bioinformatics **16:**404-5.
59. **Meister, M., S. R. Caplan, and H. C. Berg.** 1989. Dynamics of a tightly coupled mechanism for flagellar rotation. Bacterial motility, chemiosmotic coupling, protonmotive force. Biophys J **55:**905-14.
60. **Meister, M., G. Lowe, and H. C. Berg.** 1987. The proton flux through the bacterial flagellar motor. Cell **49:**643-50.
61. **Minamino, T., B. Gonzalez-Pedrajo, K. Yamaguchi, S. I. Aizawa, and R. M. Macnab.** 1999. FliK, the protein responsible for flagellar hook length control in Salmonella, is exported during hook assembly. Mol Microbiol **34:**295-304.
62. **Minamino, T., S. Yamaguchi, and R. M. Macnab.** 2000. Interaction between FliE and FlgB, a proximal rod component of the flagellar basal body of Salmonella. J Bacteriol **182:**3029-36.
63. **Misra, R. V., R. S. Horler, W. Reindl, Goryanin, II, and G. H. Thomas.** 2005. EchoBASE: an integrated post-genomic database for Escherichia coli. Nucleic Acids Res **33:**D329-33.
64. **Mongiovi, P., and R. Milo.** http://openwetware.org/wiki/BioNumber_Of_The_Month.
65. **Nakamura, S., N. Kami-ike, J. P. Yokota, S. Kudo, T. Minamino, and K. Namba.** 2009. Effect of intracellular pH on the torque-speed relationship of bacterial proton-driven flagellar motor. J Mol Biol **386:**332-8.
66. **Namba, K., and F. Vonderviszt.** 1997. Molecular architecture of bacterial flagellum. Q Rev Biophys **30:**1-65.
67. **Nambu, T., T. Minamino, R. M. Macnab, and K. Kutsukake.** 1999. Peptidoglycan-hydrolyzing activity of the FlgJ protein, essential for flagellar rod formation in Salmonella typhimurium. J Bacteriol **181:**1555-61.
68. **Nasset, E. S., F. P. Heald, D. H. Calloway, S. Margen, and P. Schneeman.** 1979. Amino acids in human blood plasma after single meals of meat, oil, sucrose and whiskey. J Nutr **109:**621-30.
69. **Nasset, E. S., and J. S. Ju.** 1969. Amino acids and glucose in human blood plasma after beef and nonprotein meals. Proc Soc Exp Biol Med **132:**1077-80.
70. **Neidhardt, F. C.** 1987. Escherichia coli and Salmonella typhimurium : cellular and molecular biology, vol. American Society for Microbiology, Washington, D.C.

71. **Neidhardt, F. C., R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low Jr., B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger.** 1996. Escherichia coli and Salmonella: Cellular and Molecular Biology, 2nd ed, vol. ASM Press, Washington, D.C.
72. **Neidhardt, F. C., J. L. Ingraham, and M. Schaechter.** 1990. Physiology of the bacterial cell : a molecular approach, vol. Sinauer Associates, Sunderland, Mass.
73. **Nenninger, A. A., L. S. Robinson, and S. J. Hultgren.** 2009. Localized and efficient curli nucleation requires the chaperone-like amyloid assembly protein CsgF. Proc Natl Acad Sci U S A **106:**900-5.
74. **Nishiyama, M., M. Vetsch, C. Puorger, I. Jelesarov, and R. Glockshuber.** 2003. Identification and characterization of the chaperone-subunit complex-binding domain from the type 1 pilus assembly platform FimD. J Mol Biol **330:**513-25.
75. **Oh, H. S., B. H. Kvitko, J. E. Morello, and A. Collmer.** 2007. Pseudomonas syringae lytic transglycosylases coregulated with the type III secretion system contribute to the translocation of effector proteins into plant cells. J Bacteriol **189:**8277-89.
76. **Ohnishi, K., Y. Ohto, S. Aizawa, R. M. Macnab, and T. Iino.** 1994. FlgD is a scaffolding protein needed for flagellar hook assembly in Salmonella typhimurium. J Bacteriol **176:**2272-81.
77. **Pirt, S. J.** 1965. The maintenance energy of bacteria in growing cultures. Proc R Soc Lond B Biol Sci **163:**224-31.
78. **Porstendorfer, D., O. Gohl, F. Mayer, and B. Averhoff.** 2000. ComP, a pilin-like protein essential for natural competence in Acinetobacter sp. Strain BD413: regulation, modification, and cellular localization. J Bacteriol **182:**3673-80.
79. **Proenza, A. M., C. Crespi, P. Roca, and A. Palou.** 2001. Gender related differences in the effect of aging on blood amino acid compartmentation*. J Nutr Biochem **12:**431-440.
80. **Qu, J., Y. Wang, G. Luo, Z. Wu, and C. Yang.** 2002. Validated quantitation of underivatized amino acids in human blood samples by volatile ion-pair reversed-phase liquid chromatography coupled to isotope dilution tandem mass spectrometry. Anal Chem **74:**2034-40.
81. **Rey, S., M. Acab, J. L. Gardy, M. R. Laird, K. deFays, C. Lambert, and F. S. Brinkman.** 2005. PSORTdb: a protein subcellular localization database for bacteria. Nucleic Acids Res **33:**D164-8.
82. **Rosu, V., and K. T. Hughes.** 2006. sigma28-dependent transcription in Salmonella enterica is independent of flagellar shearing. J Bacteriol **188:**5196-203.
83. **Sandström, J., and N. Moran.** 1999. How nutritionally imbalanced is phloem sap for aphids? Entomologia Experimentalis et Applicata **91:**203-210.
84. **Sandstrom, J., A. Telang, and N. A. Moran.** 2000. Nutritional enhancement of host plants by aphids - a comparison of three aphid species on grasses. J Insect Physiol **46:**33-40.
85. **Schaber, J., C. Rispe, J. Wernegreen, A. Buness, F. Delmotte, F. J. Silva, and A. Moya.** 2005. Gene expression levels influence amino acid usage and evolutionary rates in endosymbiotic bacteria. Gene **352:**109-17.
86. **Schechter, L. M., M. Vencato, K. L. Jordan, S. E. Schneider, D. J. Schneider, and A. Collmer.** 2006. Multiple approaches to a complete inventory of Pseudomonas syringae pv. tomato DC3000 type III secretion system effector proteins. Mol Plant Microbe Interact **19:**1180-92.

87. **Schumann, W.** 2006. Dynamics of the Bacterial Chromosome. Structure and Function., vol. Wiley-VCH, Weinheim, Germany.
88. **Schuster, S. C., and S. Khan.** 1994. The bacterial flagellar motor. Annu Rev Biophys Biomol Struct **23:**509-39.
89. **Shigenobu, S., H. Watanabe, M. Hattori, Y. Sakaki, and H. Ishikawa.** 2000. Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS. Nature **407:**81-6.
90. **Sokurenko, E. V., H. S. Courtney, J. Maslow, A. Siitonen, and D. L. Hasty.** 1995. Quantitative differences in adhesiveness of type 1 fimbriated Escherichia coli due to structural differences in fimH genes. J Bacteriol **177:**3680-6.
91. **Sowa, Y., and R. M. Berry.** 2008. Bacterial flagellar motor. Q Rev Biophys **41:**103-32.
92. **Steigmiller, S., P. Turina, and P. Graber.** 2008. The thermodynamic H+/ATP ratios of the H+-ATPsynthases from chloroplasts and Escherichia coli. Proc Natl Acad Sci U S A **105:**3745-50.
93. **Stouthamer, A. H.** 1979. The search for correlation between theoretical and experimental growth yields, p. 1-47. *In* J. R. Quayle (ed.), Int. Rev. Biochem. Microb. Biochem, vol. 21. University Park Press.
94. **Stouthamer, A. H.** 1973. A theoretical study on the amount of ATP required for synthesis of microbial cell material. Antonie Van Leeuwenhoek **39:**545-65.
95. **Stouthamer, A. H., and C. Bettenhaussen.** 1973. Utilization of energy for growth and maintenance in continuous and batch cultures of microorganisms. A reevaluation of the method for the determination of ATP production by measuring molar growth yields. Biochim Biophys Acta **301:**53-70.
96. **Stouthamer, A. H., and C. W. Bettenhaussen.** 1977. A continuous culture study of an ATPase-negative mutant of Escherichia coli. Arch Microbiol **113:**185-9.
97. **Swire, J.** 2007. Selection on synthesis cost affects interprotein amino acid usage in all three domains of life. J Mol Evol **64:**558-71.
98. **Tamas, I., L. Klasson, B. Canback, A. K. Naslund, A. S. Eriksson, J. J. Wernegreen, J. P. Sandstrom, N. A. Moran, and S. G. Andersson.** 2002. 50 million years of genomic stasis in endosymbiotic bacteria. Science **296:**2376-9.
99. **Toh, E., H. D. Kurtz, Jr., and Y. V. Brun.** 2008. Characterization of the Caulobacter crescentus holdfast polysaccharide biosynthesis pathway reveals significant redundancy in the initiating glycosyltransferase and polymerase steps. J Bacteriol **190:**7219-31.
100. **Toh, H., B. L. Weiss, S. A. Perkin, A. Yamashita, K. Oshima, M. Hattori, and S. Aksoy.** 2006. Massive genome erosion and functional adaptations provide insights into the symbiotic lifestyle of Sodalis glossinidius in the tsetse host. Genome Res **16:**149-56.
101. **Toma, C., H. Kuroki, N. Nakasone, M. Ehara, and M. Iwanaga.** 2002. Minor pilin subunits are conserved in Vibrio cholerae type IV pili. FEMS Immunol Med Microbiol **33:**35-40.
102. **Tran, Q. H., and G. Unden.** 1998. Changes in the proton potential and the cellular energetics of Escherichia coli during growth by aerobic and anaerobic respiration or by fermentation. Eur J Biochem **251:**538-43.
103. **Trinchina, E. V.** 2003. Structural and functional study of the receptor binding site for FimH adhesin in uropathogenic strains of Escherichia coli. Bull Exp Biol Med **136:**380-4.

104. **Turina, P., D. Samoray, and P. Graber.** 2003. H+/ATP ratio of proton transport-coupled ATP synthesis and hydrolysis catalysed by CF0F1-liposomes. Embo J **22:**418-26.
105. **Wagner, A.** 2005. Energy constraints on the evolution of gene expression. Mol Biol Evol **22:**1365-74.
106. **Walz, D., and S. R. Caplan.** 2000. An electrostatic mechanism closely reproducing observed behavior in the bacterial flagellar motor. Biophys J **78:**626-51.
107. **Ward, J. J., J. S. Sodhi, L. J. McGuffin, B. F. Buxton, and D. T. Jones.** 2004. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J Mol Biol **337:**635-45.
108. **Waters, R. C., P. W. O'Toole, and K. A. Ryan.** 2007. The FliK protein and flagellar hook-length control. Protein Sci **16:**769-80.
109. **Wilkinson, T. L., D. Adams, L. B. Minto, and A. E. Douglas.** 2001. The impact of host plant on the abundance and function of symbiotic bacteria in an aphid. J Exp Biol **204:**3027-38.
110. **Yonekura, K., S. Maki, D. G. Morgan, D. J. DeRosier, F. Vonderviszt, K. Imada, and K. Namba.** 2000. The bacterial flagellar cap as the rotary promoter of flagellin self-assembly. Science **290:**2148-52.
111. **Zientz, E., T. Dandekar, and R. Gross.** 2004. Metabolic interdependence of obligate intracellular bacteria and their insect hosts. Microbiol Mol Biol Rev **68:**745-70.