

Table 1: Frequency of mutations in conserved amino acids in HetR

Function ^a	Number of sites ^b	Primary HetR's (74) ^c		Secondary HetR's (9) ^c		Unicellular strains (2) ^c		Closest relative (2) ^c	
		Mutations ^d	Rate ^e	Mutations ^d	Rate ^e	Mutations ^d	Rate ^e	Mutations ^d	Rate ^e
DNA-binding	19 (11%)	3 (3 + 0) [14]**	0.002	13 (8 + 5) [28]*	0.076	1 (0 + 1) [8]*	0.026	0 (0 + 0) [3]	0.000
PatS-binding	7 (4%)	5 (5 + 0) [5]	0.010	4 (4 + 0) [10]	0.063	6 (3 + 3) [3]†	0.429	1 (1 + 0) [1]	0.071
Subunit interaction	14 (8%)	9 (6 + 3) [10]	0.009	29 (11 + 18) [20] †	0.230	10 (5 + 5) [6] †	0.357	3 (2 + 1) [2]	0.107
No information	132 (77%)	99 (55 + 44)	0.010	192 (70 + 122)	0.162	52 (25 + 27)	0.197	20 (13 + 7)	0.076
TOTAL	172 (100%)	116 (69 + 47)	0.009	239 (93 + 146)	0.154	69 (33 + 36)	0.201	24 (16 + 8)	0.070

^a Amino acid residues associated with different functions. See Supporting Information Fig. 3 for evidence

^b Conserved amino acid residues in each functional class. See Fig. 4 for definition of “conserved”.

^c Primary HetR's are those HetR that are in filamentous cyanobacteria and are phylogenetically most related, as described in the text. Secondary HetR's are very similar proteins in organisms with primary HetR's. There are two unicellular strains (*Synechococcus* PCC 7002 and *Synechococcus* PCC 7335) considered here bearing HetR-like proteins and two strains most closely related to them (*Leptolyngbya* PCC 7376 and *Leptolyngbya* PCC, respectively) .

^d Number of mutational events in conserved sites of each functional class. The first number represents conserved amino acid changes and the second non-conserved, as described in Fig. 4. Events were estimated with guidance from the organismal (Fig. 3) and HetR (Suppl. Fig. 2) phylogenetic trees. The second and third numbers (in parentheses) indicate the number of conserved and non-conserved substitutions, respectively. The last number (in brackets) indicates the expected number of substitutions, if the total number of mutations for the class (e.g. primary HetR's) were distributed proportionally to the number of sites in each functional category relative to the number of mutations in the *No information* category. For example, the expected number of mutations in Primary HetR's at amino acid positions related to DNA-binding would be 99 (19/132). Values that are significantly less than expected according to a Chi-squared test are marked with one asterisk (p <0.05) or two (p<0.01). Values that are significantly more than expected are marked with a dagger (p<0.05).

^e The rate is defined as the number of mutational events divided by the total number of possible events (the number of sites in the functional category times the number of proteins in the class).