

(A, G, P, S, T)

Smaller aa some with polar side chain
S and T frequently replace one another

(D, E, N, Q) DE: negative charge NQ: amid of the above.

Large polar - that they end up in one group may be due to early protein sequencing not being able to distinguish between D and N, and E and Q - Else N and Q have high odds to turn into HKR

(H, K, R) positively charged side chain

(F, Y, W) aromatic with extended pi electron system

F and Y change frequently, W is special, but if it changes at all, it is to F and Y

(I, L, M, V) Hydrophobic side chains

(C) Does cystine bridges and coordinates metals

See highlighting in log Odds matrixes. Except for HN-QKR and F-IP the positive log odds are within the groups. (In the Dayhoff 4 recoding Groups 4 and 5 are joined, and C turns in ?)

	G	A	V	L	I	P	S	T	D	E	N	Q	K	R	H	F	Y	W	M	C	B	Z	X	*		
G	5																								G	
A	1	2																								A
V	-1	0	4																							V
L	-4	-2	2	6																						L
I	-3	-1	4	2	5																					I
P	0	1	-1	-3	-2	6																				P
S	1	1	-1	-3	-1	1	2																			S
T	0	1	0	-2	0	0	1	3																		T
D	1	0	-2	-4	-2	-1	0	0	4																	D
E	0	0	-2	-3	-2	-1	0	0	3	4																E
N	0	0	-2	-3	-2	0	1	0	2	1	2															N
Q	-1	0	-2	-2	-2	0	-1	-1	2	2	1	4														Q
K	-2	-1	-2	-3	-2	-1	0	0	0	0	1	1	5													K
R	-3	-2	-2	-3	-2	0	0	-1	-1	-1	0	1	3	6												R
H	-2	-1	-2	-2	-2	0	-1	-1	1	1	2	3	0	2	6											H
F	-5	-3	-1	2	1	-5	-3	-3	-6	-5	-3	-5	-5	-4	-2	9										F
Y	-5	-3	-2	-1	-1	-5	-3	-3	-4	-4	-2	-4	-4	-4	0	7	10									Y
W	-7	-6	-6	-2	-5	-6	-2	-5	-7	-7	-4	-5	-3	-2	-3	0	0	17								W
M	-3	-1	2	4	2	-2	-2	-1	-3	-2	-2	-1	0	0	-2	0	-2	-4	6							M
C	-3	-2	-2	-6	-2	-3	0	-2	-5	-5	-4	-5	-5	-4	-3	-4	0	-8	-5	12						C

PAM 250

This should not take care of all bias, it would be surprising, if halophiles weren't biased under this matrix.