PSSM's with pseudocounts

W	**E**	**I**	**R**	**D**	
W	**E**	**I**	**R**	**D**	
W	**E**	**I**	**R**	**E**	
W	**E**	**I**	**Q**	**H**	

D	**1**	**1**	**1**	**1**	**3**	**5**
E	**1**	**5**	**1**	**1**	**2**	**5**
H	**1**	**1**	**1**	**1**	**2**	**5**
I	**1**	**1**	**5**	**1**	**1**	**5**
Q	**1**	**1**	**1**	**2**	**1**	**5**
R	**1**	**1**	**1**	**4**	**1**	**5**
W	**5**	**1**	**1**	**1**	**1**	**5**
SUM	**11**	**11**	**11**	**11**	**11**	

F[i,j]

D	**0.091**	**0.091**	**0.091**	**0.091**	**0.27**
E	**0.091**	**0.455**	**0.091**	**0.091**	**0.18**
H	**0.091**	**0.091**	**0.091**	**0.091**	**0.18**
I	**0.091**	**0.091**	**0.455**	**0.091**	**0.09**
Q	**0.091**	**0.091**	**0.091**	**0.182**	**0.09**
R	**0.091**	**0.091**	**0.091**	**0.364**	**0.09**
W	**0.455**	**0.091**	**0.091**	**0.091**	**0.09**
SUM	1.0	1.0	1.0	1.0	1.0

$$F[i, j] = \frac{n_i + b}{k + |\Sigma| b}$$

Typically, *b=1*

Pretend |Σ| = 7 for simplicity of exposition. DON'T do that at home. Really |Σ| =20 for proteins, 4 for DNA

P[i,j]

D	**1.7**	**1.7**	**1.7**	**1.7**	**5.2**	**D**	**0.052**
E	**1.5**	**7.3**	**1.5**	**1.5**	**2.9**	**E**	**0.062**
H	**4.0**	**4.0**	**4.0**	**4.0**	**7.9**	**H**	**0.023**
I	**1.7**	**1.7**	**8.6**	**1.7**	**1.7**	**I**	**0.053**
Q	**2.2**	**2.2**	**2.2**	**4.4**	**2.2**	**Q**	**0.041**
R	**1.8**	**1.8**	**1.8**	**7.1**	**1.8**	**R**	**0.051**
W	**32.5**	**6.5**	**6.5**	**6.5**	**6.5**	**W**	**0.014**

####

$$P[i, j] = \frac{F[i, j]}{b_i}$$

S[i,j]

D	**0.8**	**0.8**	**0.8**	**0.8**	**2.4**	**D**	**0.052**
E	**0.6**	**2.9**	**0.6**	**0.6**	**1.6**	**E**	**0.062**
H	**2.0**	**2.0**	**2.0**	**2.0**	**3.0**	**H**	**0.023**
I	**0.8**	**0.8**	**3.1**	**0.8**	**0.8**	**I**	**0.053**
Q	**1.1**	**1.1**	**1.1**	**2.1**	**1.1**	**Q**	**0.041**
R	**0.8**	**0.8**	**0.8**	**2.8**	**0.8**	**R**	**0.051**
W	**5.0**	**2.7**	**2.7**	**2.7**	**1.8**	**W**	**0.014**

$$S[i, j] = \log_2 P[i, j]$$

Scoring:

		W	**I**	**W**	**E**	**I**	**R**	**H**
SUM	**9.8**	5.0	0.8	2.7	0.6	0.8		
	5.6		0.8	2.7	0.6	0.8	0.8	
	16.8			5.0	2.9	3.1	2.8	3.0