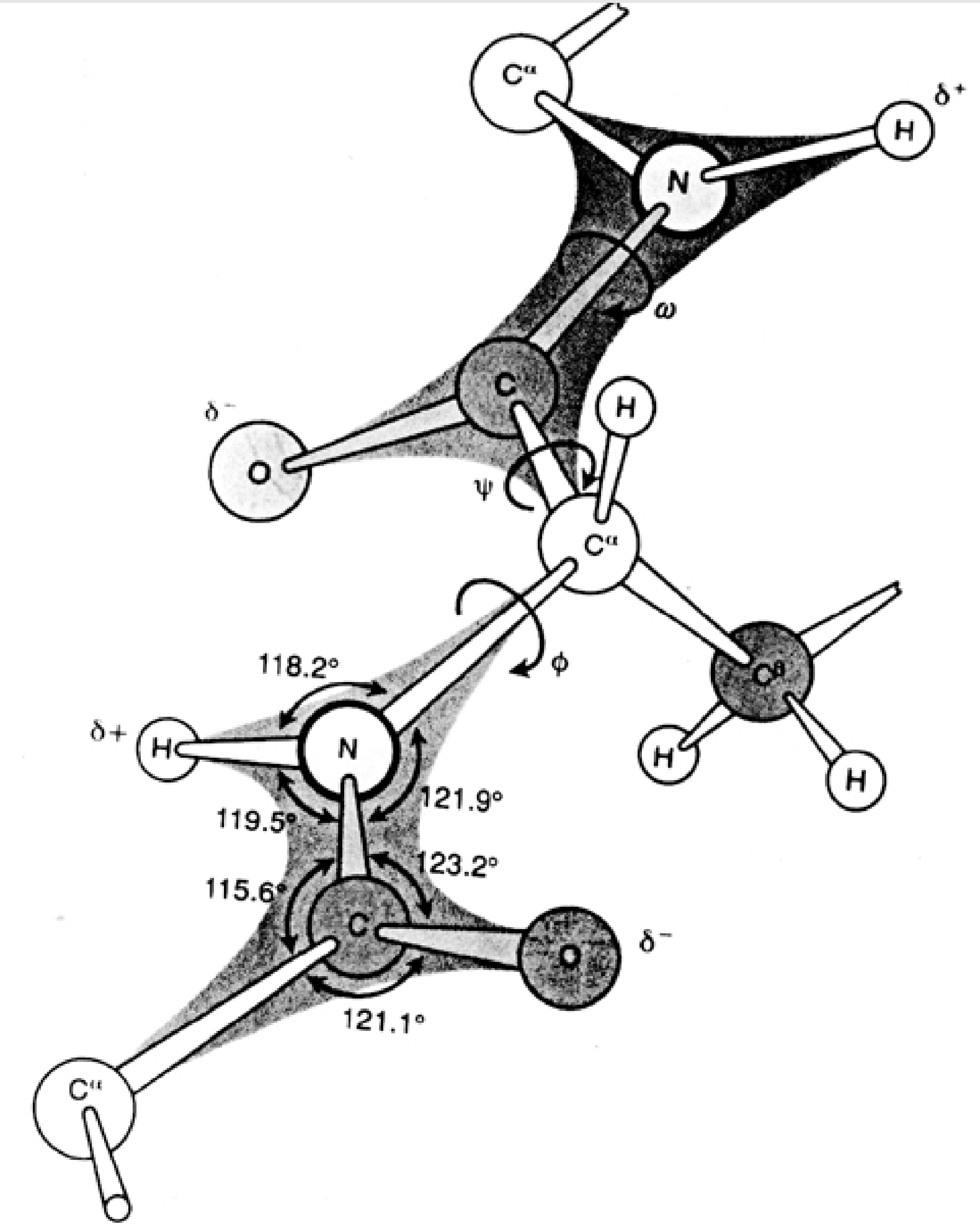
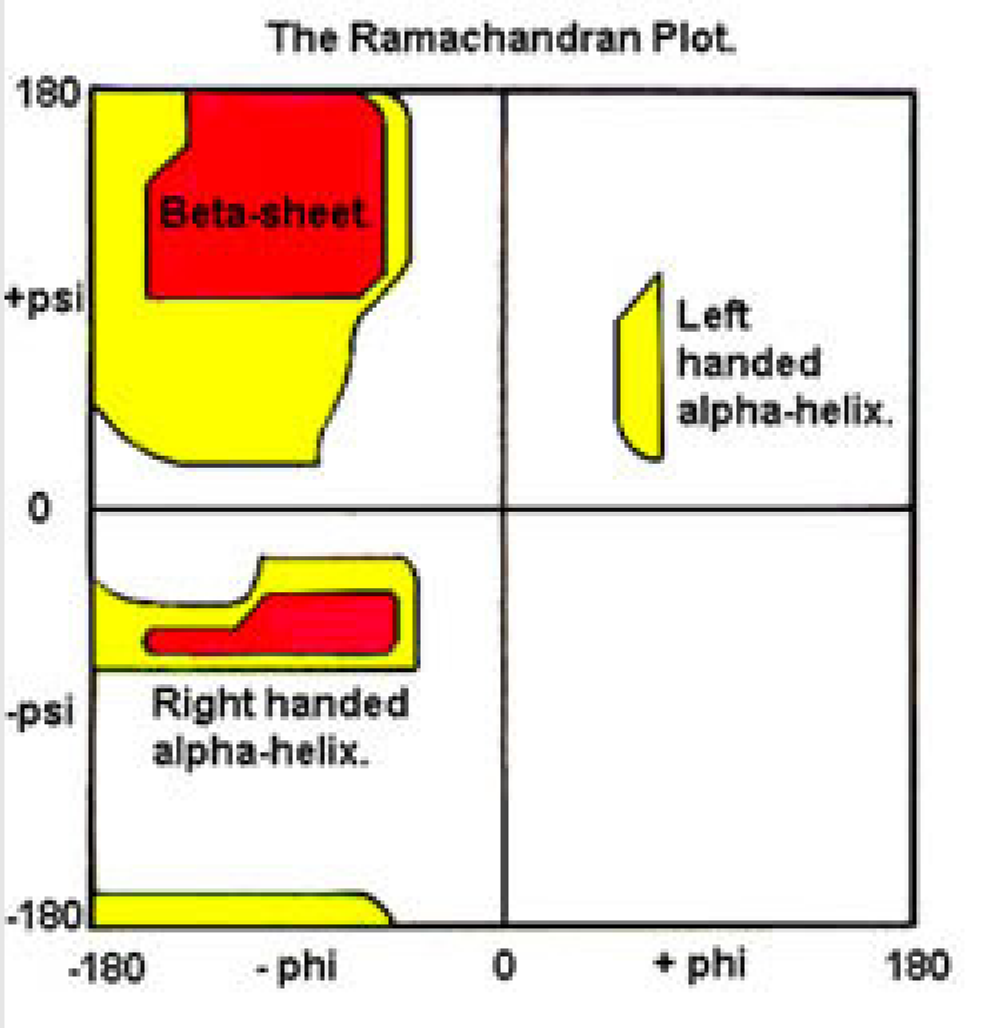
Makromolekulák

Simon István

A peptid kötés

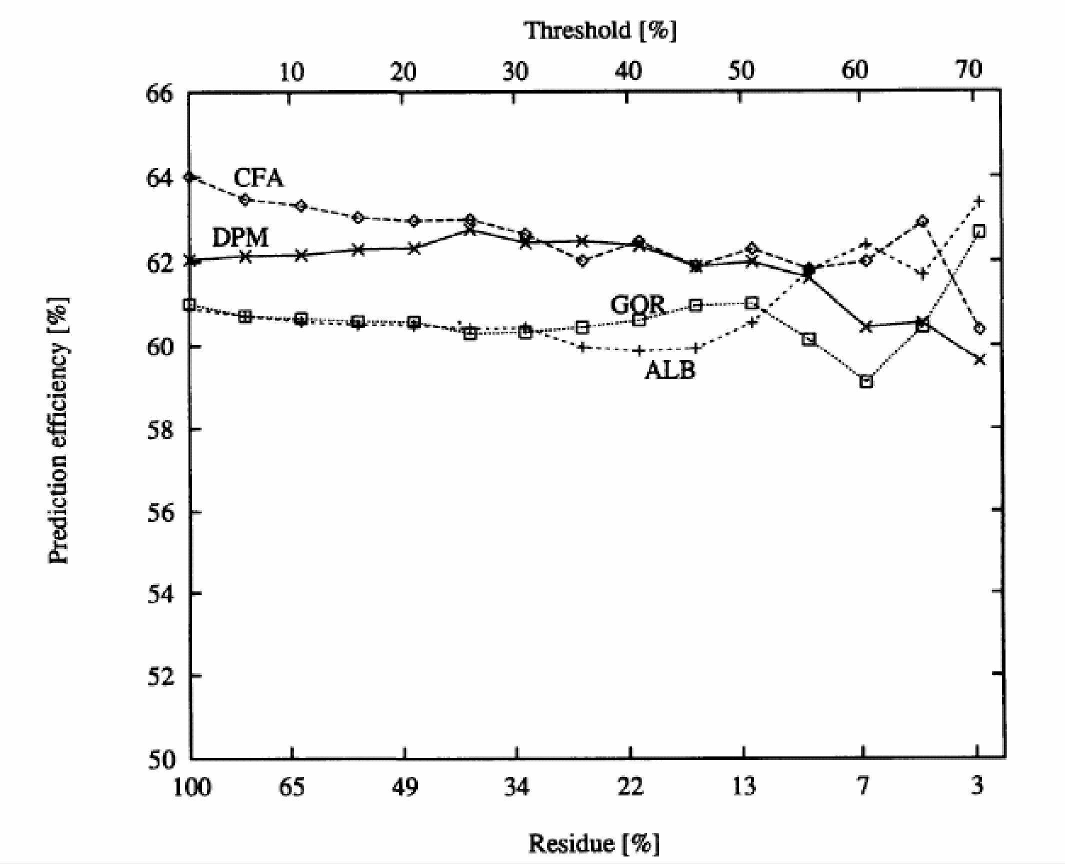



The Ramachandran Plot.
180
+psi
0
-psi
-180
-180
- phi
0
+ phi
180
Beta-sheet.
Left handed alpha-helix.
Right handed alpha-helix.

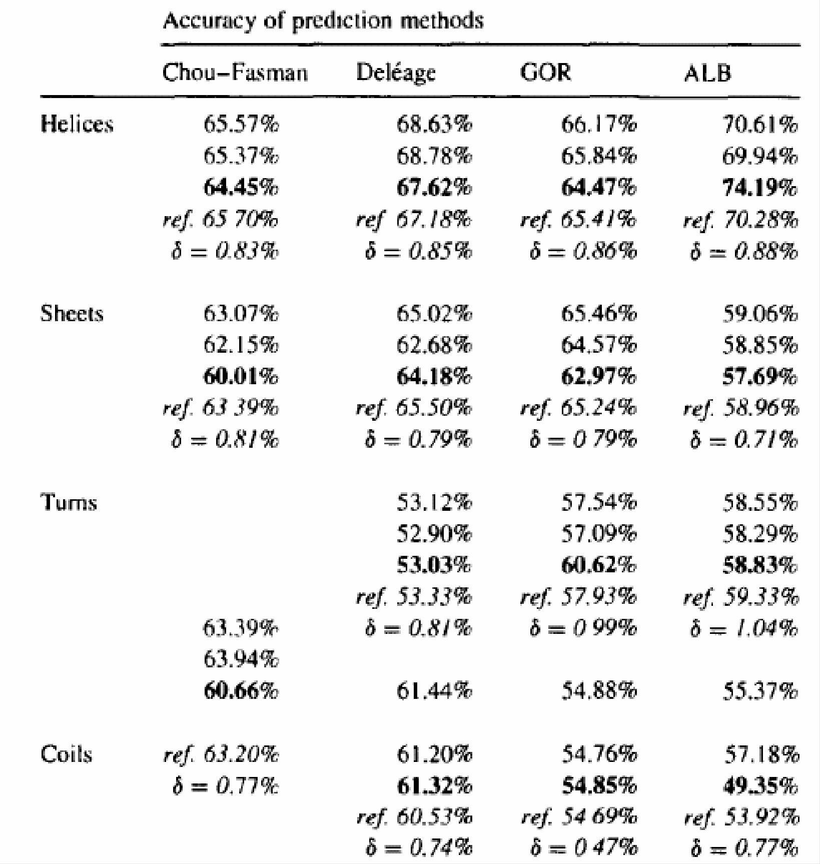
Chou-Fasman Paraméterek

Aminosav	P(a)	P(b)
Alanine	142	83
Arginine	98	93
Aspartic Acid	101	54
Asparagine	67	89
Cysteine	70	119
Glutamic Acid	151	037
Glutamine	111	110
Glycine	57	75
Histidine	100	87
Isoleucine	108	160
Leucine	121	130
Lysine	114	74
Methionine	145	105
Phenylalanine	113	138
Proline	57	55
Serine	77	75
Threonine	83	119
Tryptophan	108	137
Tyrosine	69	147
Valine	106	170

Másodlagos szerkezetbecslő módszerek pontossága



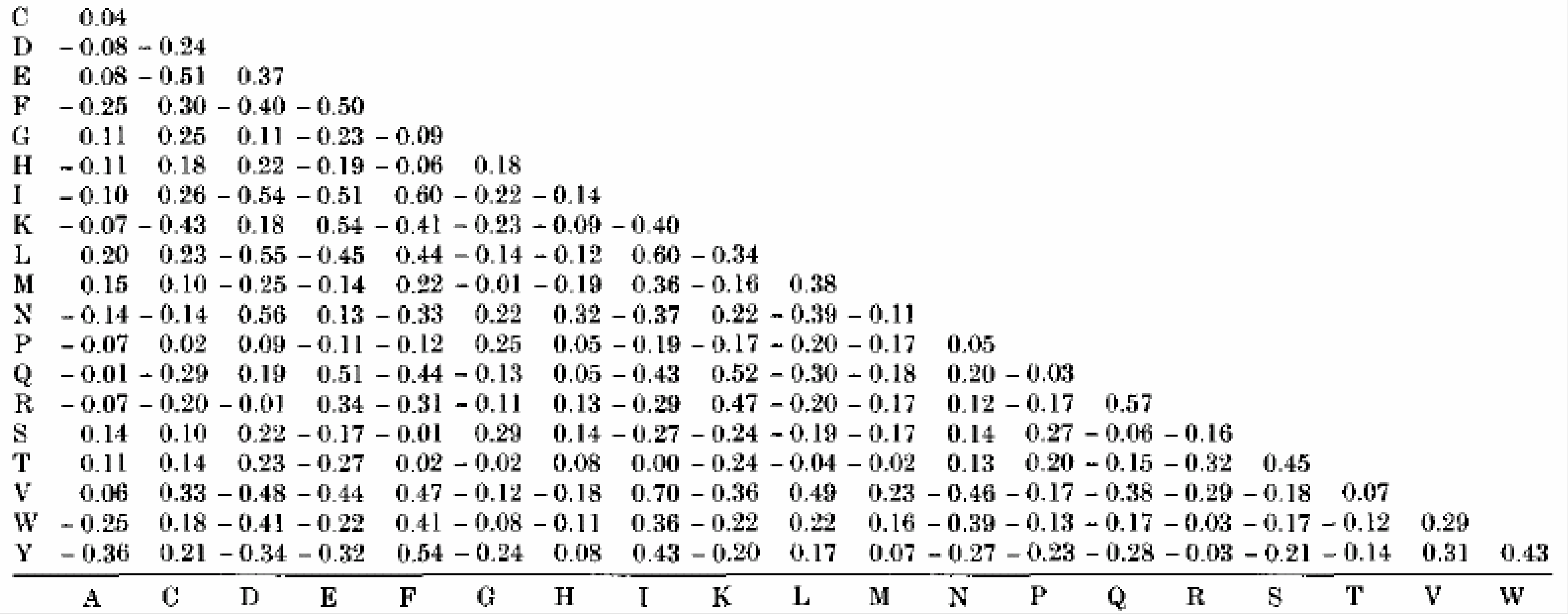
Másodlagos szerkezetbecslő módszerek pontossága



Accuracy of prediction methods

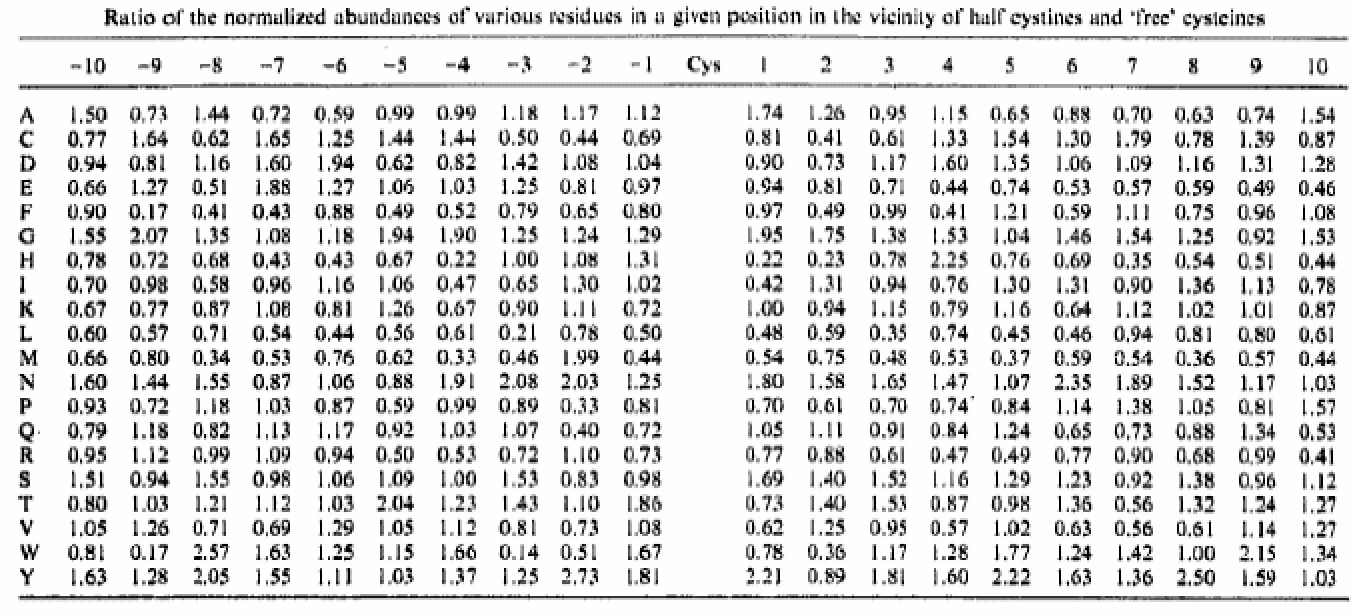
	Chou–Fasman	Deléage	GOR	ALB
Helices	65.57%	68.63%	66.17%	70.61%
	65.37%	68.78%	65.84%	69.94%
	64.45%	**67.62%**	**64.47%**	**74.19%**
	ref. 65 70%	*ref 67.18%*	*ref. 65.41%*	*ref. 70.28%*
	δ = 0.83%	*δ = 0.85%*	*δ = 0.86%*	*δ = 0.88%*
Sheets	63.07%	65.02%	65.46%	59.06%
	62.15%	62.68%	64.57%	58.85%
	60.01%	**64.18%**	**62.97%**	**57.69%**
	ref. 63 39%	*ref. 65.50%*	*ref. 65.24%*	*ref. 58.96%*
	δ = 0.81%	*δ = 0.79%*	*δ = 0 79%*	*δ = 0.71%*
Turns		53.12%	57.54%	58.55%
		52.90%	57.09%	58.29%
		53.03%	**60.62%**	**58.83%**
		ref. 53.33%	*ref. 57.93%*	*ref. 59.33%*
	63.39%	*δ = 0.81%*	*δ = 0 99%*	*δ = 1.04%*
	63.94%			
	60.66%	61.44%	54.88%	55.37%
Coils	*ref. 63.20%*	61.20%	54.76%	57.18%
	δ = 0.77%	**61.32%**	**54.85%**	**49.35%**
		ref. 60.53%	*ref. 54 69%*	*ref. 53.92%*
		δ = 0.74%	*δ = 0 47%*	*δ = 0.77%*

Aminosav helyettesítési mátrix



	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W
C	0.04																		
D	−0.08	−0.24																	
E	0.08	−0.51	0.37																
F	−0.25	0.30	−0.40	−0.50															
G	0.11	0.25	0.11	−0.23	−0.09														
H	−0.11	0.18	0.22	−0.19	−0.06	0.18													
I	−0.10	0.26	−0.54	−0.51	0.60	−0.22	−0.14												
K	−0.07	−0.43	0.18	0.54	−0.41	−0.23	−0.09	−0.40											
L	0.20	0.23	−0.55	−0.45	0.44	−0.14	−0.12	0.60	−0.34										
M	0.15	0.10	−0.25	−0.14	0.22	−0.01	−0.19	0.36	−0.16	0.38									
N	−0.14	−0.14	0.56	0.13	−0.33	0.22	0.32	−0.37	0.22	−0.39	−0.11								
P	−0.07	0.02	0.09	−0.11	−0.12	0.25	0.05	−0.19	−0.17	−0.20	−0.17	0.05							
Q	−0.01	−0.29	0.19	0.51	−0.44	−0.13	0.05	−0.43	0.52	−0.30	−0.18	0.20	−0.03						
R	−0.07	−0.20	−0.01	0.34	−0.31	−0.11	0.13	−0.29	0.47	−0.20	−0.17	0.12	−0.17	0.57					
S	0.14	0.10	0.22	−0.17	−0.01	0.29	0.14	−0.27	−0.24	−0.19	−0.17	0.14	0.27	−0.06	−0.16				
T	0.11	0.14	0.23	−0.27	0.02	−0.02	0.08	0.00	−0.24	−0.04	−0.02	0.13	0.20	−0.15	−0.32	0.45			
V	0.06	0.33	−0.48	−0.44	0.47	−0.12	−0.18	0.70	−0.36	0.49	0.23	−0.46	−0.17	−0.38	−0.29	−0.18	0.07		
W	−0.25	0.18	−0.41	−0.22	0.41	−0.08	−0.11	0.36	−0.22	0.22	0.16	−0.39	−0.13	−0.17	−0.03	−0.17	−0.12	0.29	
Y	−0.36	0.21	−0.34	−0.32	0.54	−0.24	0.08	0.43	−0.20	0.17	0.07	−0.27	−0.23	−0.28	−0.03	−0.21	−0.14	0.31	0.43

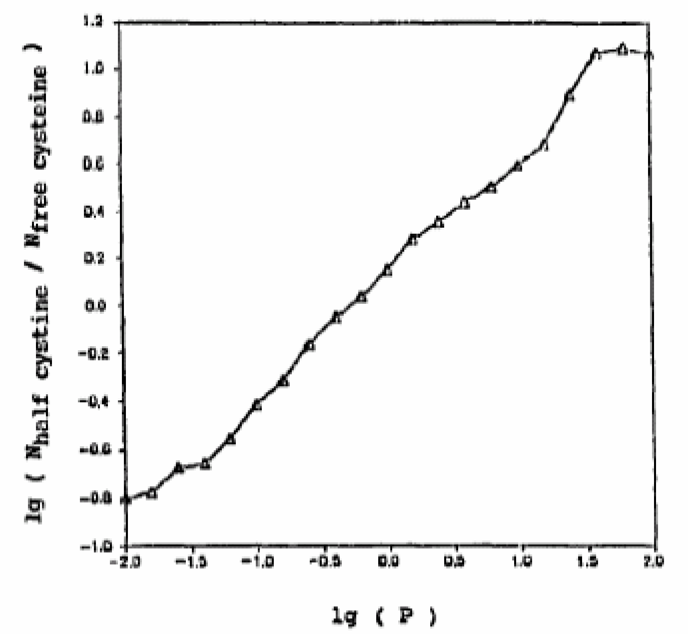
Diszulfid becslő potenciál



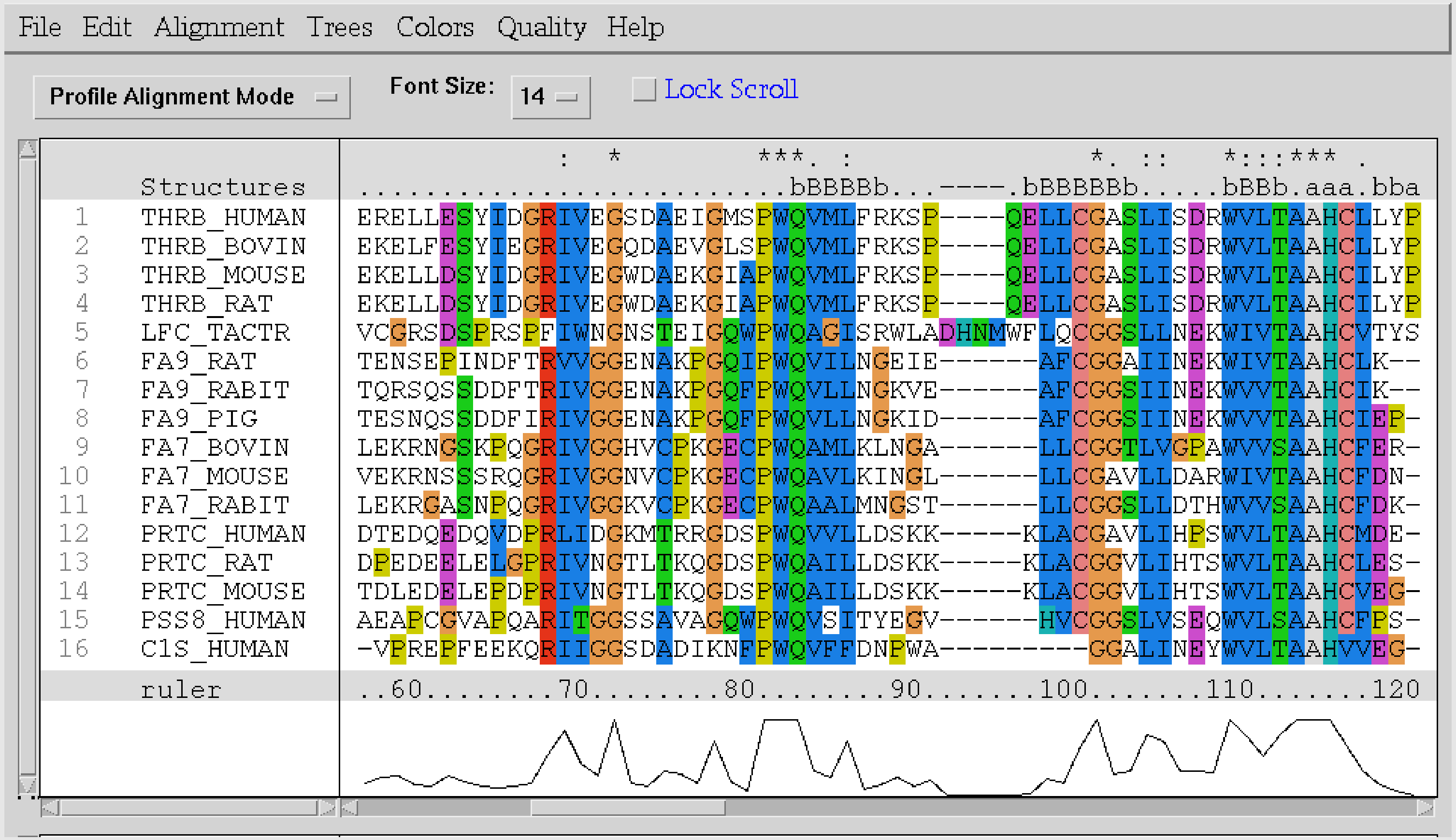
Ratio of the normalized abundances of various residues in a given position in the vicinity of half cystines and 'free' cysteines

	−10	−9	−8	−7	−6	−5	−4	−3	−2	−1	Cys	1	2	3	4	5	6	7	8	9	10
A	1.50	0.73	1.44	0.72	0.59	0.99	0.99	1.18	1.17	1.12		1.74	1.26	0.95	1.15	0.65	0.88	0.70	0.63	0.74	1.54
C	0.77	1.64	0.62	1.65	1.25	1.44	1.44	0.50	0.44	0.69		0.81	0.41	0.61	1.33	1.54	1.30	1.79	0.78	1.39	0.87
D	0.94	0.81	1.16	1.60	1.94	0.62	0.82	1.42	1.08	1.04		0.90	0.73	1.17	1.60	1.35	1.06	1.09	1.16	1.31	1.28
E	0.66	1.27	0.51	1.88	1.27	1.06	1.03	1.25	0.81	0.97		0.94	0.81	0.71	0.44	0.74	0.53	0.57	0.59	0.49	0.46
F	0.90	0.17	0.41	0.43	0.88	0.49	0.52	0.79	0.65	0.80		0.97	0.49	0.99	0.41	1.21	0.59	1.11	0.75	0.96	1.08
G	1.55	2.07	1.35	1.08	1.18	1.94	1.90	1.25	1.24	1.29		1.95	1.75	1.38	1.53	1.04	1.46	1.54	1.25	0.92	1.53
H	0.78	0.72	0.68	0.43	0.43	0.67	0.22	1.00	1.08	1.31		0.22	0.23	0.78	2.25	0.76	0.69	0.35	0.54	0.51	0.44
I	0.70	0.98	0.58	0.96	1.16	1.06	0.47	0.65	1.30	1.02		0.42	1.31	0.94	0.76	1.30	1.31	0.90	1.36	1.13	0.78
K	0.67	0.77	0.87	1.08	0.81	1.26	0.67	0.90	1.11	0.72		1.00	0.94	1.15	0.79	1.16	0.64	1.12	1.02	1.01	0.87
L	0.60	0.57	0.71	0.54	0.44	0.56	0.61	0.21	0.78	0.50		0.48	0.59	0.35	0.74	0.45	0.46	0.94	0.81	0.80	0.61
M	0.66	0.80	0.34	0.53	0.76	0.62	0.33	0.46	1.99	0.44		0.54	0.75	0.48	0.53	0.37	0.59	0.54	0.36	0.57	0.44
N	1.60	1.44	1.55	0.87	1.06	0.88	1.91	2.08	2.03	1.25		1.80	1.58	1.65	1.47	1.07	2.35	1.89	1.52	1.17	1.03
P	0.93	0.72	1.18	1.03	0.87	0.59	0.99	0.89	0.33	0.81		0.70	0.61	0.70	0.74	0.84	1.14	1.38	1.05	0.81	1.57
Q	0.79	1.18	0.82	1.13	1.17	0.92	1.03	1.07	0.40	0.72		1.05	1.11	0.91	0.84	1.24	0.65	0.73	0.88	1.34	0.53
R	0.95	1.12	0.99	1.09	0.94	0.50	0.53	0.72	1.10	0.73		0.77	0.88	0.61	0.47	0.49	0.77	0.90	0.68	0.99	0.41
S	1.51	0.94	1.55	0.98	1.06	1.09	1.00	1.53	0.83	0.98		1.69	1.40	1.52	1.16	1.29	1.23	0.92	1.38	0.96	1.12
T	0.80	1.03	1.21	1.12	1.03	2.04	1.23	1.43	1.10	1.86		0.73	1.40	1.53	0.87	0.98	1.36	0.56	1.32	1.24	1.27
V	1.05	1.26	0.71	0.69	1.29	1.05	1.12	0.81	0.73	1.08		0.62	1.25	0.95	0.57	1.02	0.63	0.56	0.61	1.14	1.27
W	0.81	0.17	2.57	1.63	1.25	1.15	1.66	0.14	0.51	1.67		0.78	0.36	1.17	1.28	1.77	1.24	1.42	1.00	2.15	1.34
Y	1.63	1.28	2.05	1.55	1.11	1.03	1.37	1.25	2.73	1.81		2.21	0.89	1.81	1.60	2.22	1.63	1.36	2.50	1.59	1.03

Diszulfid becslés pontossága



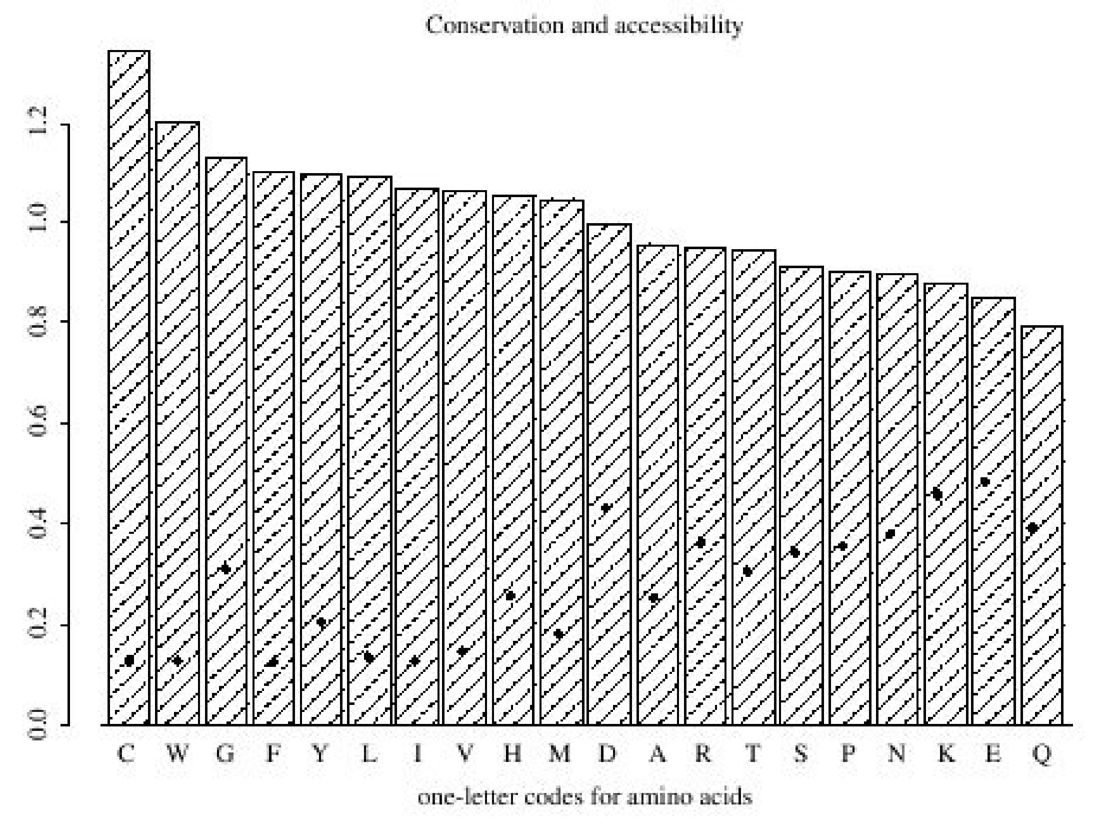
Szerin proteázok szekvenciáinak összehasonlítása



Font Size: 14

```
                             :  *        ***. :              *. ::   *:::*** .
     Structures   ..........................bBBBBb...----.bBBBBBb.....bBBb.aaa.bba
 1   THRB_HUMAN   ERELLESYIDGRIVEGSDAEIGMSPWQVMLFRKSP----QELLCGASLISDRWVLTAAHCLLYP
 2   THRB_BOVIN   EKELFESYIEGRIVEGQDAEVGLSPWQVMLFRKSP----QELLCGASLISDRWVLTAAHCLLYP
 3   THRB_MOUSE   EKELLDSYIDGRIVEGWDAEKGIAPWQVMLFRKSP----QELLCGASLISDRWVLTAAHCILYP
 4   THRB_RAT     EKELLDSYIDGRIVEGWDAEKGIAPWQVMLFRKSP----QELLCGASLISDRWVLTAAHCILYP
 5   LFC_TACTR    VCGRSDSPRSPFIWNGNSTEIGQWPWQAGISRWLADHNMWFLQCGGSLLNEKWIVTAAHCVTYS
 6   FA9_RAT      TENSEPINDFTRVVGGENAKPGQIPWQVILNGEIE------AFCGGAIINEKWIVTAAHCLK--
 7   FA9_RABIT    TQRSQSSDDFTRIVGGENAKPGQFPWQVLLNGKVE------AFCGGSIINEKWVVTAAHCIK--
 8   FA9_PIG      TESNQSSDDFIRIVGGENAKPGQFPWQVLLNGKID------AFCGGSIINEKWVVTAAHCIEP-
 9   FA7_BOVIN    LEKRNGSKPQGRIVGGHVCPKGECPWQAMLKLNGA------LLCGGTLVGPAWVVSAAHCFER-
10   FA7_MOUSE    VEKRNSSSRQGRIVGGNVCPKGECPWQAVLKINGL------LLCGAVLLDARWIVTAAHCFDN-
11   FA7_RABIT    LEKRGASNPQGRIVGGKVCPKGECPWQAALMNGST------LLCGGSLLDTHWVVSAAHCFDK-
12   PRTC_HUMAN   DTEDQEDQVDPRLIDGKMTRRGDSPWQVVLLDSKK-----KLACGAVLIHPSWVLTAAHCMDE-
13   PRTC_RAT     DPEDEELELGPRIVNGTLTKQGDSPWQAILLDSKK-----KLACGGVLIHTSWVLTAAHCLES-
14   PRTC_MOUSE   TDLEDELEPDPRIVNGTLTKQGDSPWQAILLDSKK-----KLACGGVLIHTSWVLTAAHCVEG-
15   PSS8_HUMAN   AEAPCGVAPQARITGGSSAVAGQWPWQVSITYEGV------HVCGGSLVSEQWVLSAAHCFPS-
16   C1S_HUMAN    -VPREPFEEKQRIIGGSDADIKNFPWQVFFDNPWA---------GGALINEYWVLTAAHVVEG-

     ruler        ..60.......70.......80.......90.......100.......110.......120
```

Aminosavak konzervativitása és eltemetettsége



Diszulfid hidak előfordulása másodlagos szerkezet szerint

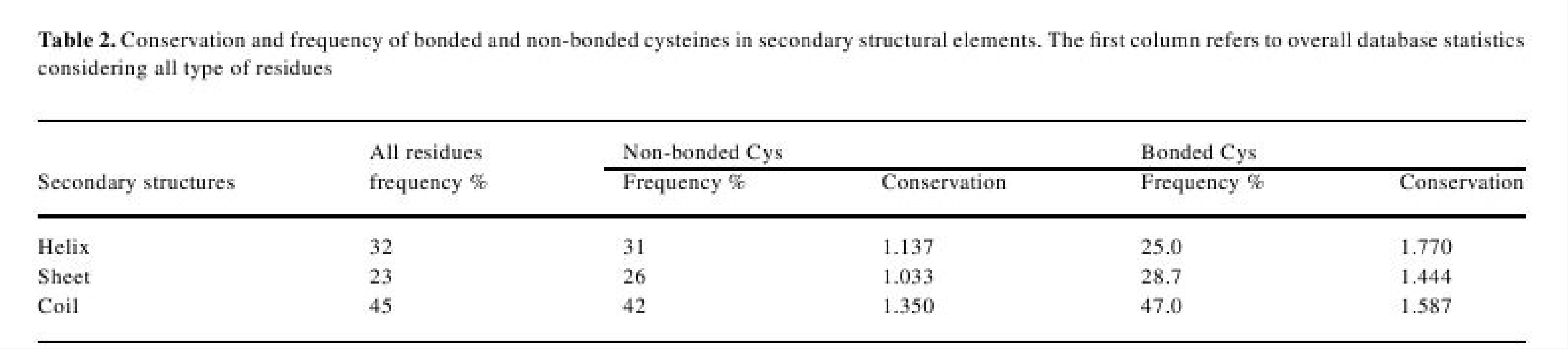
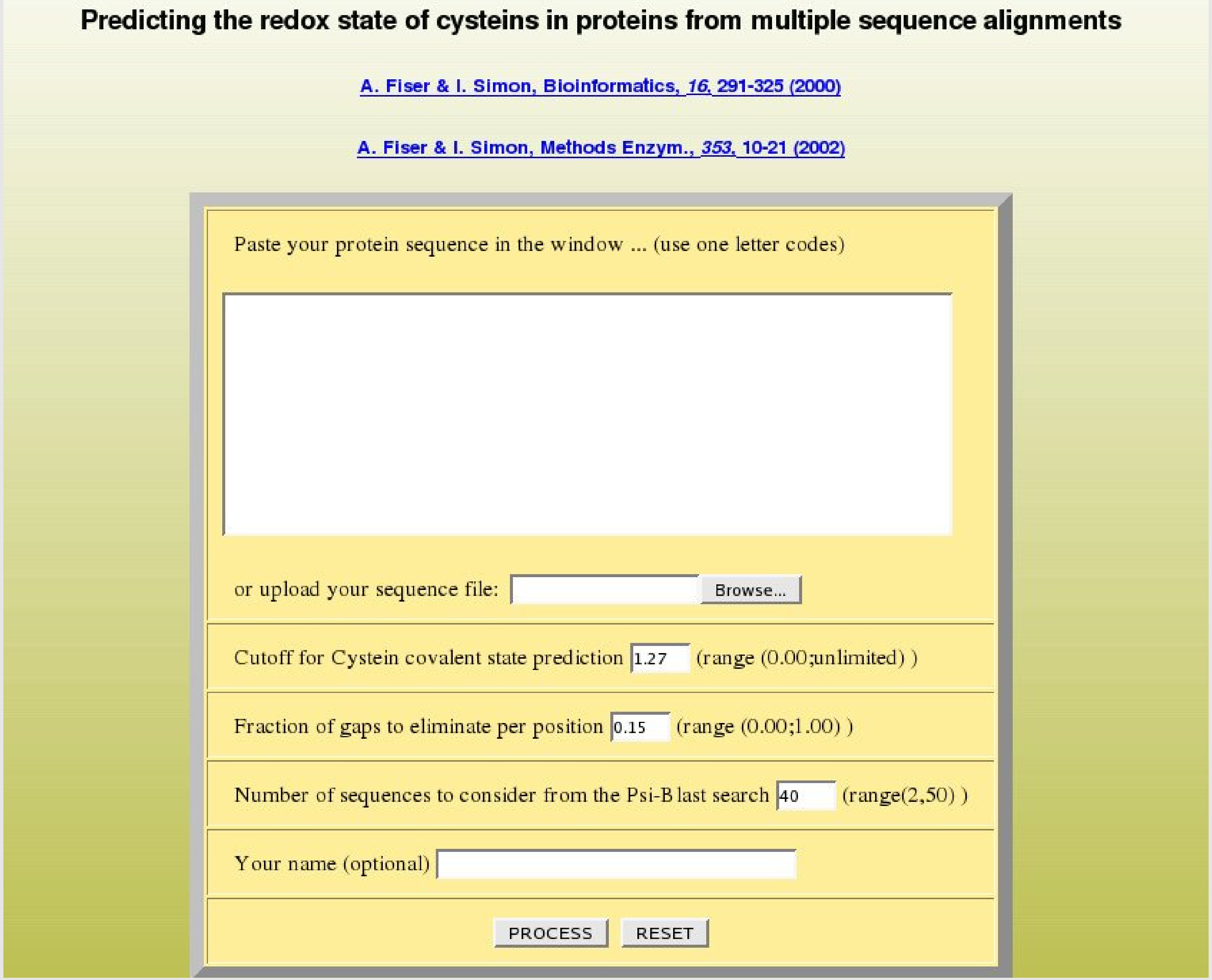


Table 2. Conservation and frequency of bonded and non-bonded cysteines in secondary structural elements. The first column refers to overall database statistics considering all type of residues

	All residues	Non-bonded Cys		Bonded Cys	
Secondary structures	frequency %	Frequency %	Conservation	Frequency %	Conservation
Helix	32	31	1.137	25.0	1.770
Sheet	23	26	1.033	28.7	1.444
Coil	45	42	1.350	47.0	1.587

http://manaslu.aecom.yu.edu/cysredox.html



Predicting the redox state of cysteins in proteins from multiple sequence alignments

A. Fiser & I. Simon, Bioinformatics, *16*, 291-325 (2000)

A. Fiser & I. Simon, Methods Enzym., *353*, 10-21 (2002)

Paste your protein sequence in the window ... (use one letter codes)

or upload your sequence file:

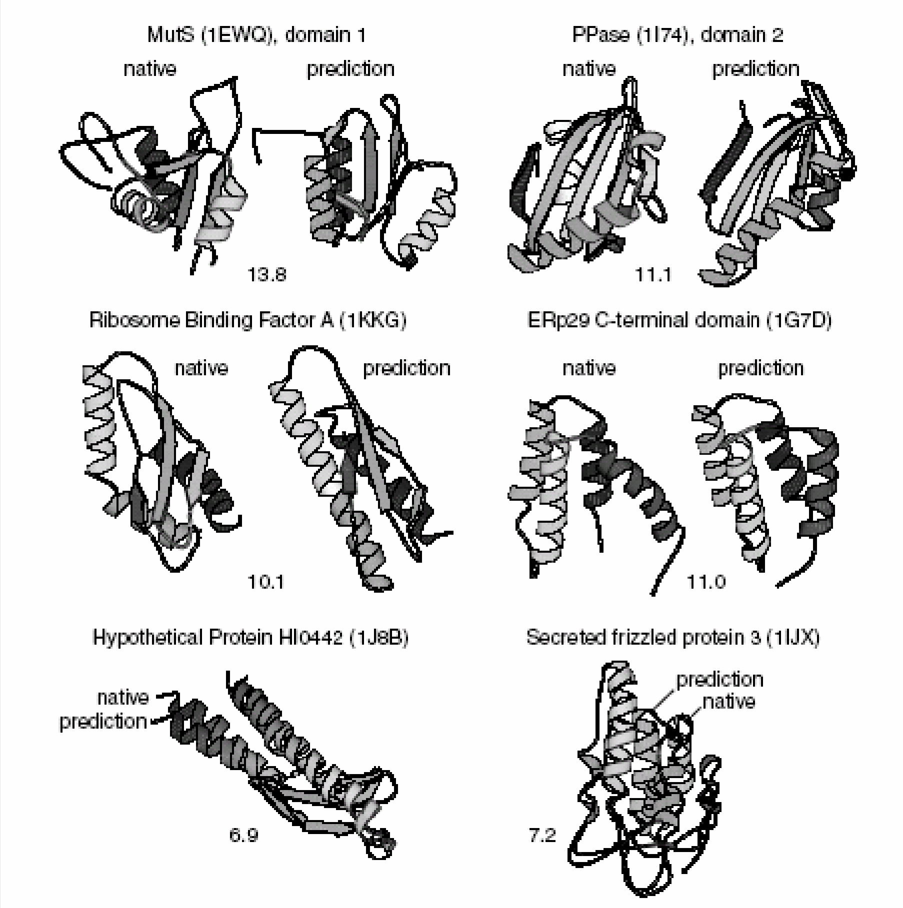
Cutoff for Cystein covalent state prediction 1.27 (range (0.00;unlimited))

Fraction of gaps to eliminate per position 0.15 (range (0.00;1.00))

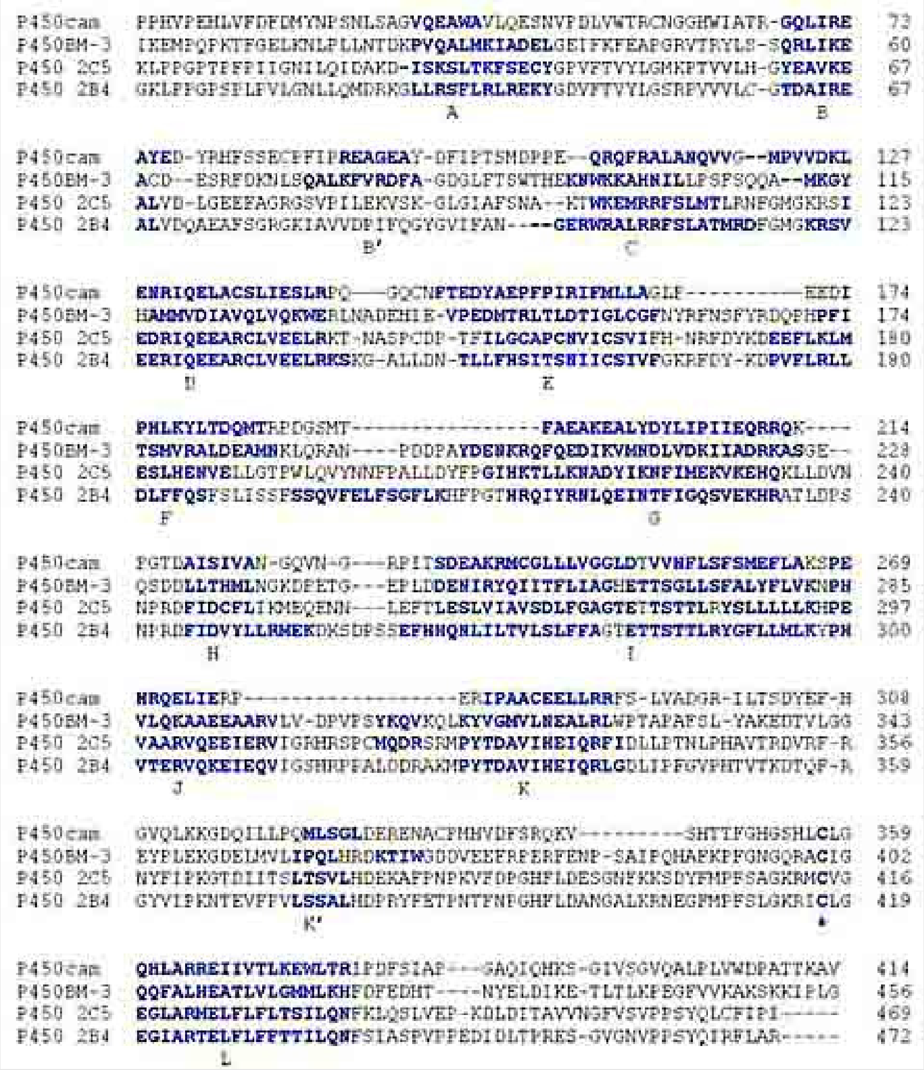
Number of sequences to consider from the Psi-B last search 40 (range(2,50))

Your name (optional)

A rozetta módszer

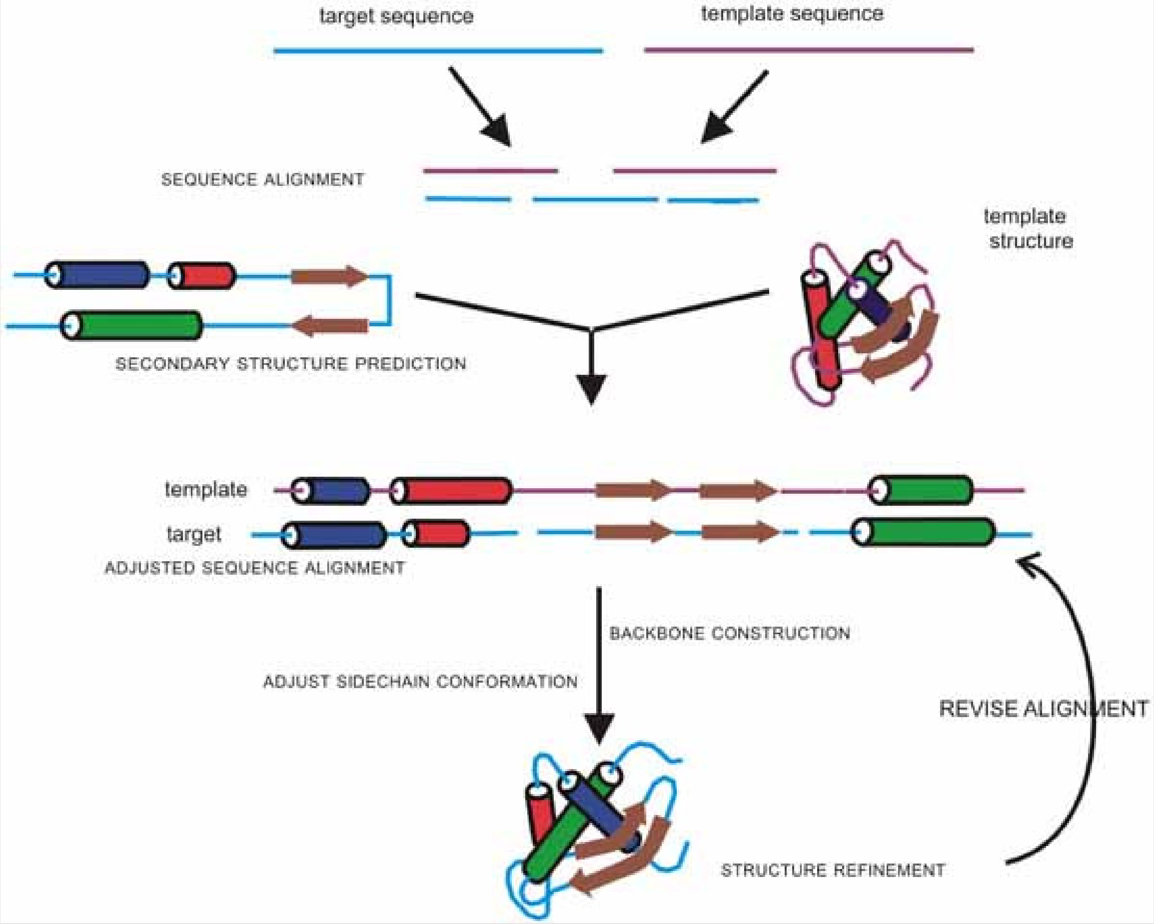


Homológia meghatározása

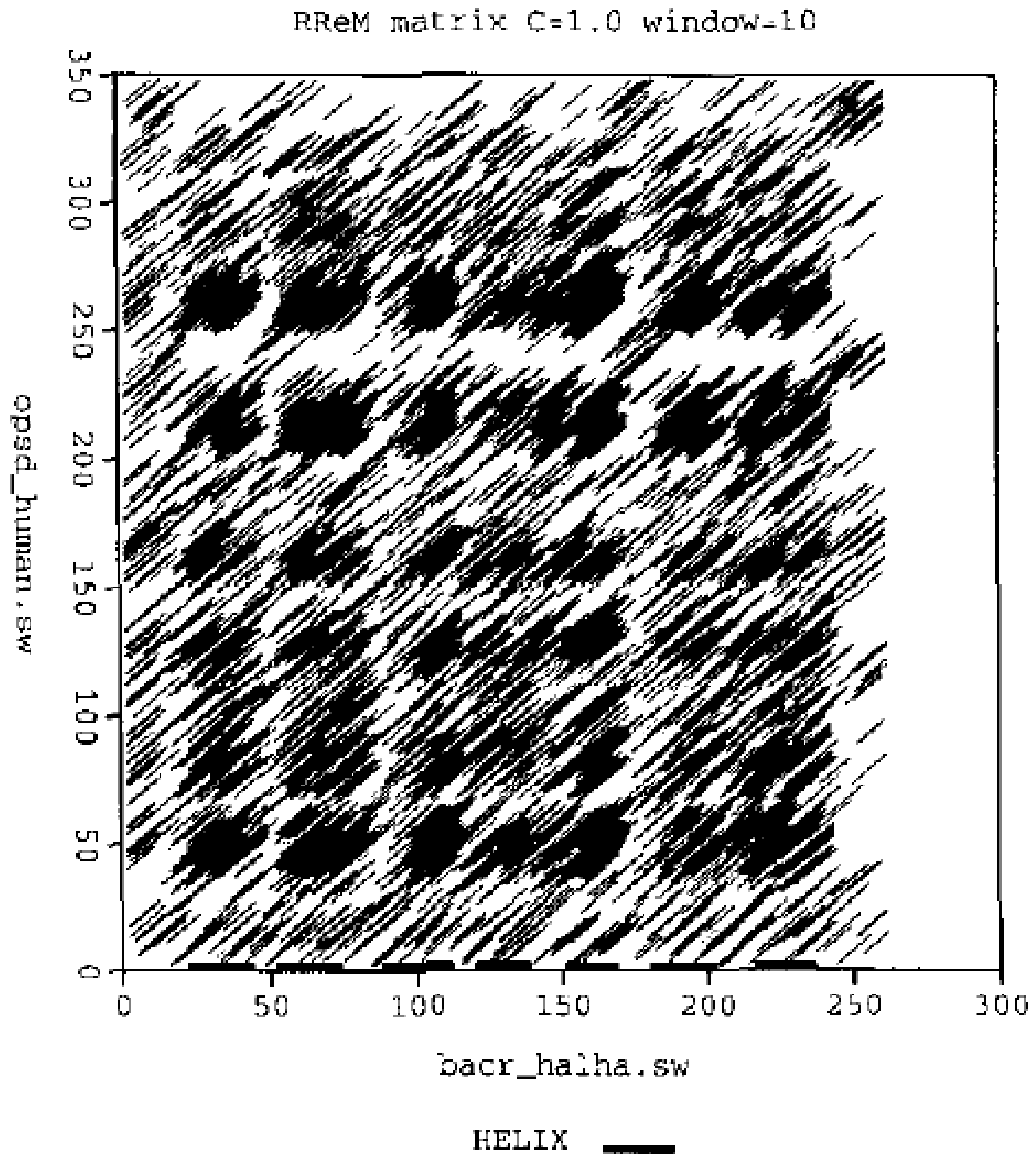


```
P450cam    PPHVPEHLVFDFDMYNPSNLSAGVQEAWAVLQESNVPDLVWTRCNGGHWIATR-GQLIRE   73
P450BM-3   IKEMPQPKTFGELKNLPLLNTDKPVQALMKIADELGEIFKFEAPGRVTRYLS-SQRLIKE   60
P450 2C5   KLPPGPTPFPIIGNILQIDAKD-ISKSLTKFSECYGPVFTVYLGMKPTVVLH-GYEAVKE   67
P450 2B4   GKLPPGPSPLPVLGNLLQMDRKGLLRSFLRLREKYGDVFTVYLGSRPVVVLC-GTDAIRE   67
                                      A                               B

P450cam    AYED-YRHFSSECPFIPREAGEAY-DFIPTSMDPPE--QRQFRALANQVVG--MPVVDKL  127
P450BM-3   ACD--ESRFDKNLSQALKFVRDFA-GDGLFTSWTHEKNWKKAHNILLPSFSQQA--MKGY  115
P450 2C5   ALVD-LGEEFAGRGSVPILEKVSK-GLGIAFSNA--KTWKEMRRFSLMTLRNFGMGKRSI  123
P450 2B4   ALVDQAEAFSGRGKIAVVDPIFQGYGVIFAN----GERWRALRRFSLATMRDFGMGKRSV  123
                               B'                     C

P450cam    ENRIQELACSLIESLRPQ---GQCNFTEDYAEPFPIRIFMLLAGLP----------EEDI  174
P450BM-3   HAMMVDIAVQLVQKWERLNADEHIE-VPEDMTRLTLDTIGLCGFNYRFNSFYRDQPHPFI  174
P450 2C5   EDRIQEEARCLVEELRKT-NASPCDP-TFILGCAPCNVICSVIFH-NRFDYKDEEFLKLM  180
P450 2B4   EERIQEEARCLVEELRKSKG-ALLDN-TLLFHSITSNIICSIVFGKRFDY-KDPVFLRLL  190
                D                              E

P450cam    PHLKYLTDQMTRPDGSMT-----------------FAEAKEALYDYLIPIIEQRRQK----  214
P450BM-3   TSMVRALDEAMNKLQRAN----PDDPAYDENKRQFQEDIKVMNDLVDKIIADRKASGE--  228
P450 2C5   ESLHENVELLGTPWLQVYNNFPALLDYFPGIHKTLLKNADYIKNFIMEKVKEHQKLLDVN  240
P450 2B4   DLFFQSFSLISSFSSQVFELFSGFLKHFPGTHRQIYRNLQEINTFIGQSVEKHRATLDPS  240
             F                                          G

P450cam    PGTDAISIVAN-GQVN-G---RPITSDEAKRMCGLLLVGGLDTVVNFLSFSMEFLAKSPE  269
P450BM-3   QSDDLLTHMLNGKDPETG---EPLDDENIRYQIITFLIAGHETTSGLLSFALYFLVKNPH  285
P450 2C5   NPRDFIDCFLIKMEQENN---LEFTLESLVIAVSDLFGAGTETTSTTLRYSLLLLLKHPE  297
P450 2B4   NPRDFIDVYLLRMEKDKSDPSSEFHHQNLILTVLSLFFAGTETTSTTLRYGFLLMLKYPH  300
                  H                                    I

P450cam    HRQELIERP-----------------ERIPAACEELLRRFS-LVADGR-ILTSDYEF-H  308
P450BM-3   VLQKAAEEAARVLV-DPVPSYKQVKQLKYVGMVLNEALRLWPTAPAFSL-YAKEDTVLGG  343
P450 2C5   VAARVQEEIERVIGRHRSPCMQDRSRMPYTDAVIHEIQRFIDLLPTNLPHAVTRDVRF-R  356
P450 2B4   VTERVQKEIEQVIGSHRPPALDDRAKMPYTDAVIHEIQRLGDLIPFGVPHTVTKDTQF-R  359
               J                             K

P450cam    GVQLKKGDQILLPQMLSGLDERENACPMHVDFSRQKV---------SHTTFGHGSHLCLG  359
P450BM-3   EYPLEKGDELMVLIPQLHRDKTIWGDDVEEFRPERFENP-SAIPQHAFKPFGNGQRACIG  402
P450 2C5   NYFIPKGTDIITSLTSVLHDEKAFPNPKVFDPGHFLDESGNFKKSDYFMPFSAGKRMCVG  416
P450 2B4   GYVIPKNTEVFPVLSSALHDPRYFETPNTFNPGHFLDANGALKRNEGFMPFSLGKRICLG  419
                          K'                                          *

P450cam    QHLARREIIVTLKEWLTRIPDFSIAP---GAQIQHKS-GIVSGVQALPLVWDPATTKAV   414
P450BM-3   QQFALHEATLVLGMMLKHFDFEDHT----NYELDIKE-TLTLKPEGFVVKAKSKKIPLG   456
P450 2C5   EGLARMELFLFLTSILQNFKLQSLVEP-KDLDITAVVNGFVSVPPSYQLCFIPI-----   469
P450 2B4   EGIARTELFLFFTTILQNFSIASPVPPEDIDLTPRES-GVGNVPPSYQIRFLAR-----   472
                   L
```

Szerkezetbecslés homológia alapján



Az emberi rodopszin és a bakteriorodopszin aminosav-sorrendjeinek összehasonlítása



A DAS szerver algoritmusa