

Figure 1. A diagram of the RDHSP homology recognition algorithm. The broken arrows indicate the calculation of environment-specific amino substitution tables. The continuous arrows indicate the generation of the profile library, which represents all known protein structures, from individual structures using environment specific substitution tables and structure-dependent gap penalties. The dash-dot arrows show the flow of the alignment and homology recognition procedure using multiple sequence/structure information.

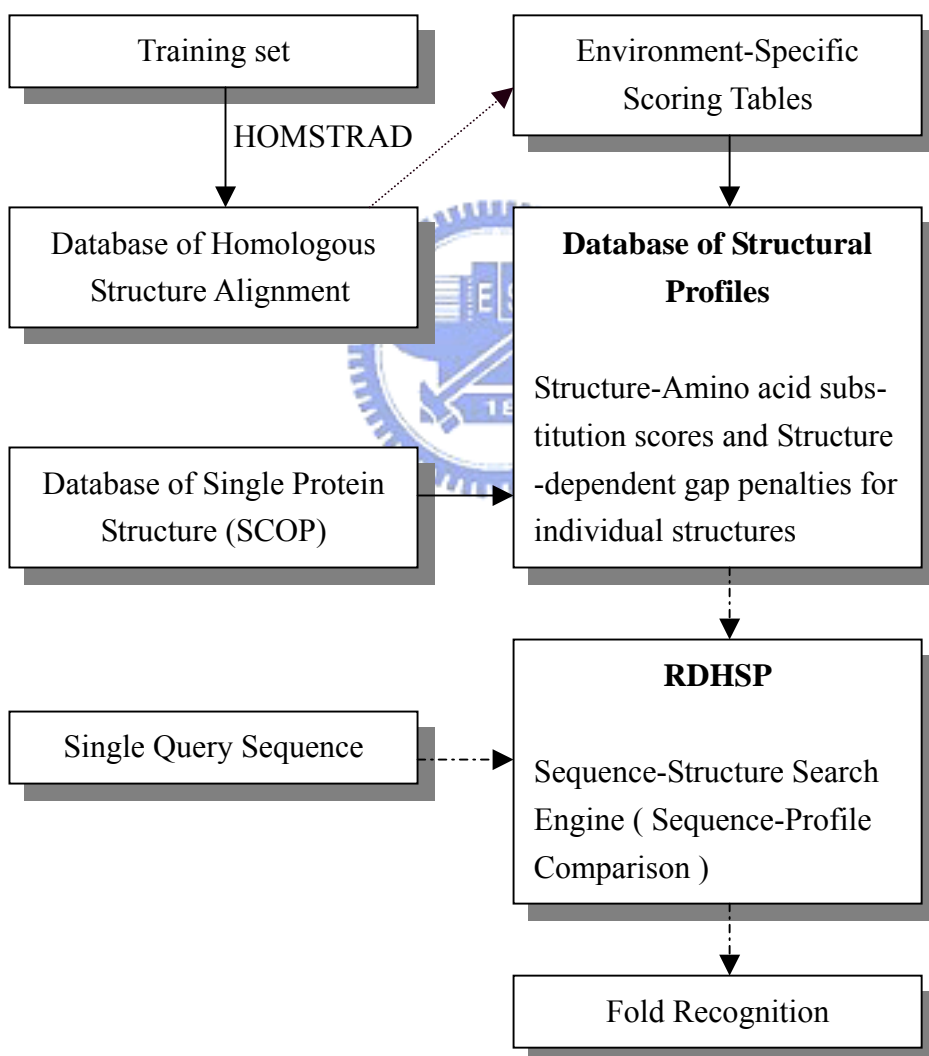


Figure 2. JOY – protein structure and alignment analysis (Overington et al, 1990; Mizuguchi et al., 1998)

solvent inaccessible	UPPER CASE	X
solvent accesible	lower case	x
positive ϕ	<i>italic</i>	x
<i>cis</i> -peptide	breve	˘
hydrogen bond to other sidechain	tilde	˜
hydrogen bond to mainchain amide	bold	x
hydrogen bond to mainchain carbonyl	<u>underline</u>	<u>x</u>
disulphide bond	cedilla	ç
α -helix	red	x
β -strand	blue	x
3_{10} -helix	maroon	x



```

10      20      30      40      50
laqt ( 2 )  styhLdVvSaeqMfsglVekIqVtGsegeIgyppghaplltaIkpgmIr
            bbbbbbb  bbbbbbbbbbbbbbb  bbbb  bbbbb  bbbb

60      70      80      90      100
laqt ( 52 )  IvkqhgheefIYLsggiLeVqpgnVtVlAdtAirGqlddearAmeakrka
            bbb  bbbbbbb  bbbbbbb  bbbbbbb  bbb333  aaaaaaaaaa

110      120      130
laqt ( 102 )  eehisshgdvyaqAsaelakAiaqlrVieltkk
            aaaa  aaaaaaaaaaaaaaaaaaaaaaaaaa

```

Figure 3 Solvent accessibility – buried surface area

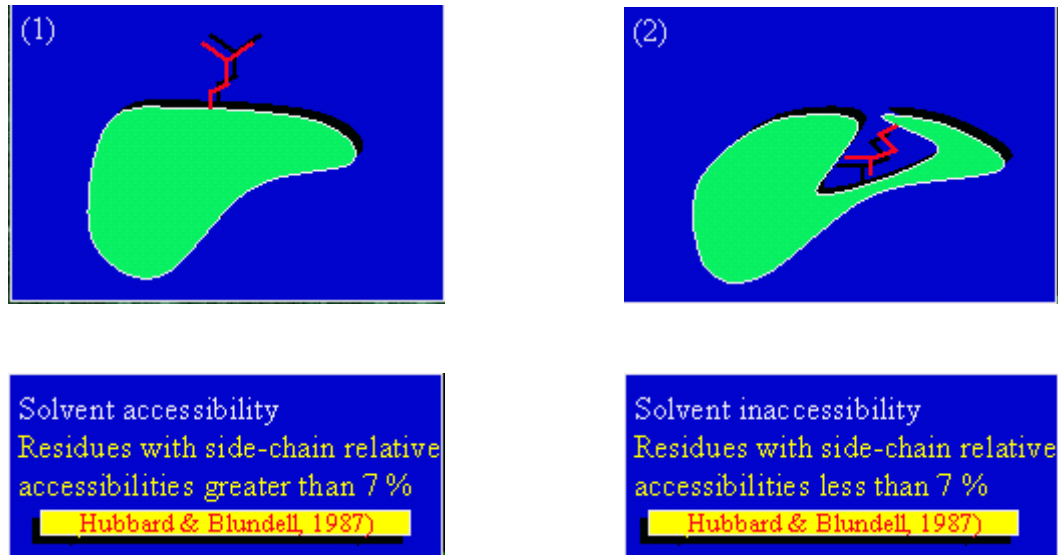


Figure 4 Estimate contact energies – Coarse grained model

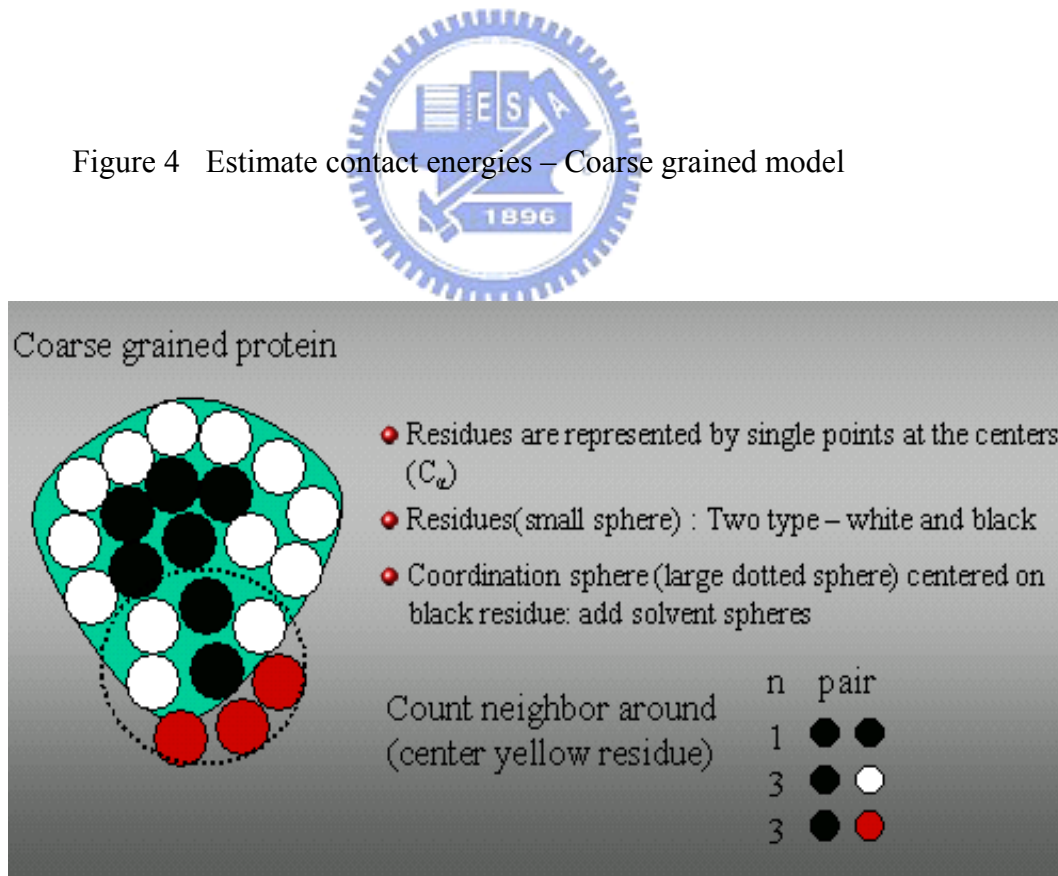
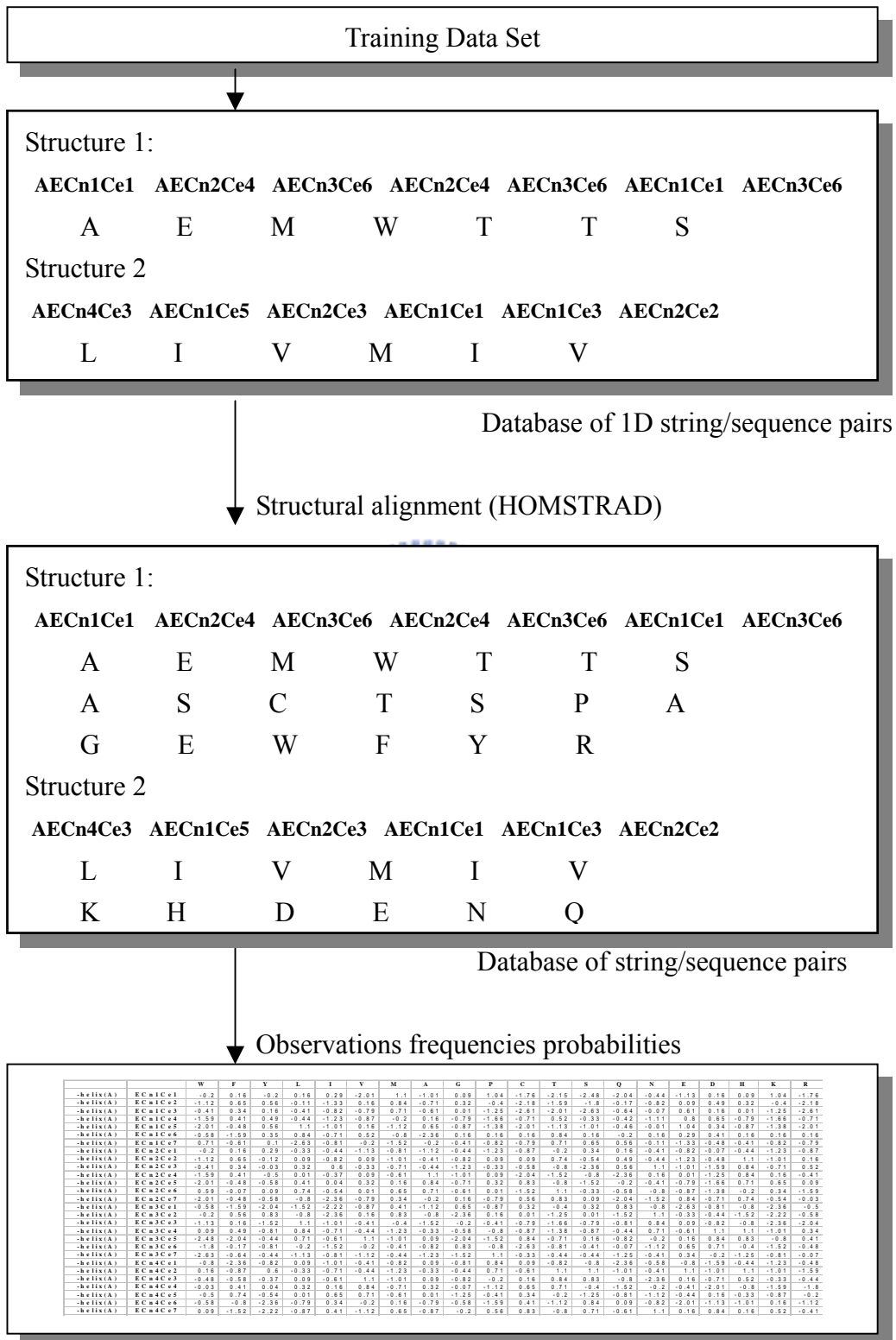


Figure 5 Calculation of environment-specific amino acid Scoring matrix



3D-1D Scoring Table

Figure 6 Example to calculation of environment-specific amino acid Scoring matrix

Structure 1:						
AECn1Ce1	AECn2Ce4	AECn3Ce6	AECn2Ce4	AECn3Ce6	AECn1Ce1	AECn3Ce6
A	E	M	W	T	T	S
A	S	C	T	S	P	A
G	E	W	F	Y	R	
Structure 2						
AECn4Ce3	AECn1Ce5	AECn2Ce3	AECn1Ce1	AECn1Ce3	AECn2Ce2	
L	I	V	M	I	V	
K	H	D	E	N	Q	
Structure 3 :						
AECn2Ce6	AECn3Ce3	AECn1Ce1	AECn2Ce4	AECn3Ce4	AECn1Ce1	AECn1Ce4
Q	N	E	D	H	K	S
L	I	V	F	Y	V	A
M	M	A	S	T	Y	L

Explane:

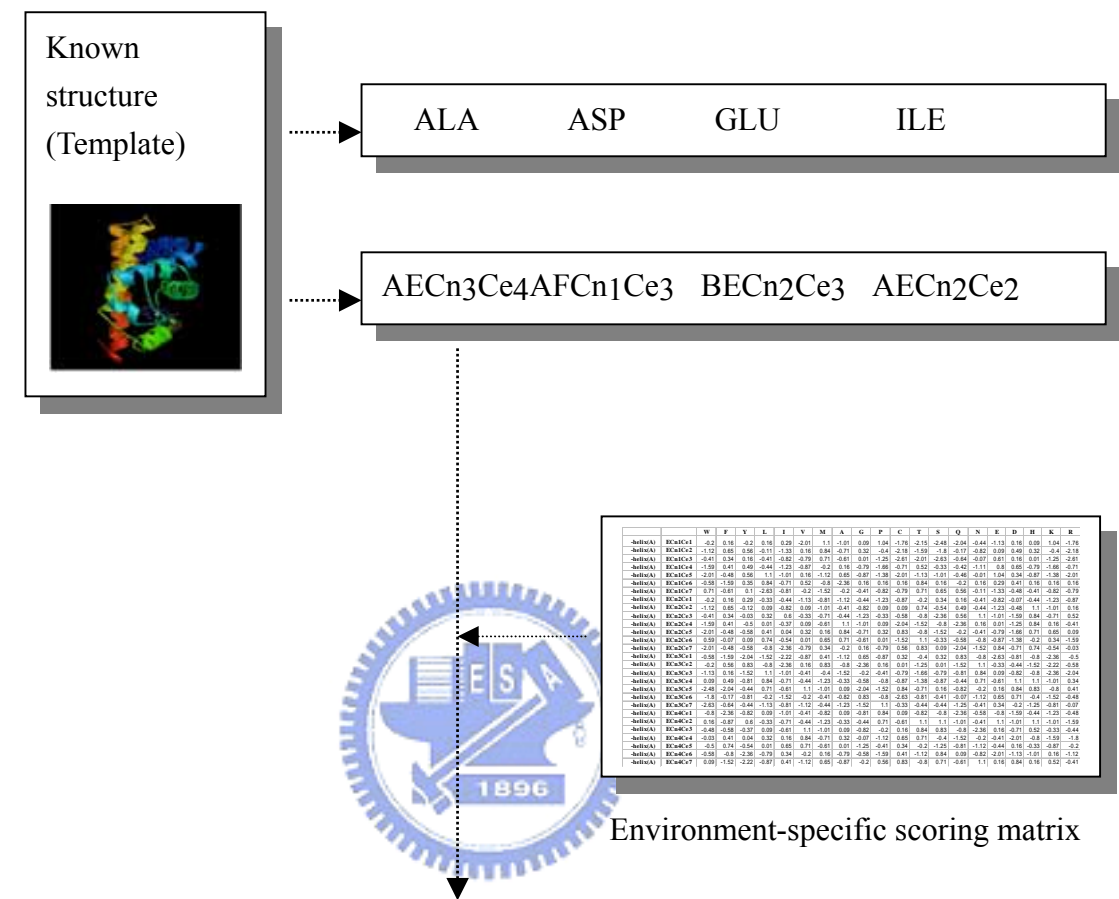
A residue type was counted only once per position.

The probability of finding an amino acid of type “A” in the environment of type “AECn1Ce1”: $2/4=0.5$

The probability of finding an amino acid of type A in any environment: $4/20 =0.2$

$$Z(A, AECn1Ce1)=-\ln[P(i,m)/P(i)] = -\ln(0.5/0.2)= -0.92$$

Figure 7 Generation of a position-dependent comparison matrix known as the 3D profile



Environment Class	(13*Si)				Gap Penalties				
	W	F	Y	H	K	R	ω_0	ω_1
AECn3Ce4	0.117	0.637	-1.053	1.452	0.453	-1.123	1.2	0.7
AFCn1Ce3	1.143	-0.554	2.453	-3.456	1.234	2.234	1.2	0.7
BECn2Ce3	0.315	0.123	3.234	1.543	-3.234	0.543	1.2	0.7
AECn2Ce2	2.345	1.456	-5.126	0.543	3.235	-3.245	1.2	0.7
...								
...								

1D – 3D profile

Figure 8 A sequence profile alignment method using a global dynamic programming algorithm was employed to find the minimum of the total score that aligns the query sequence with a template in the template library.

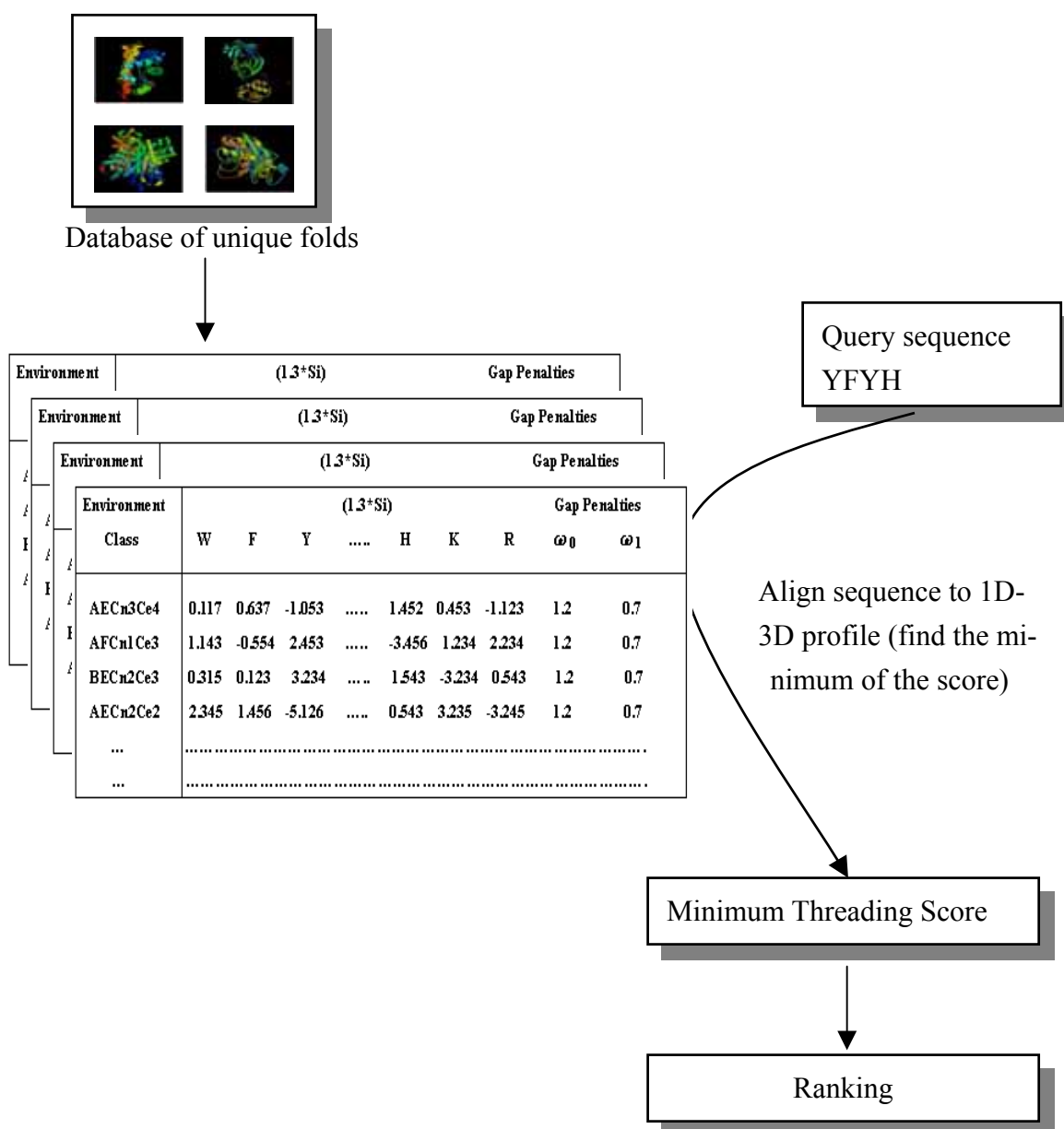
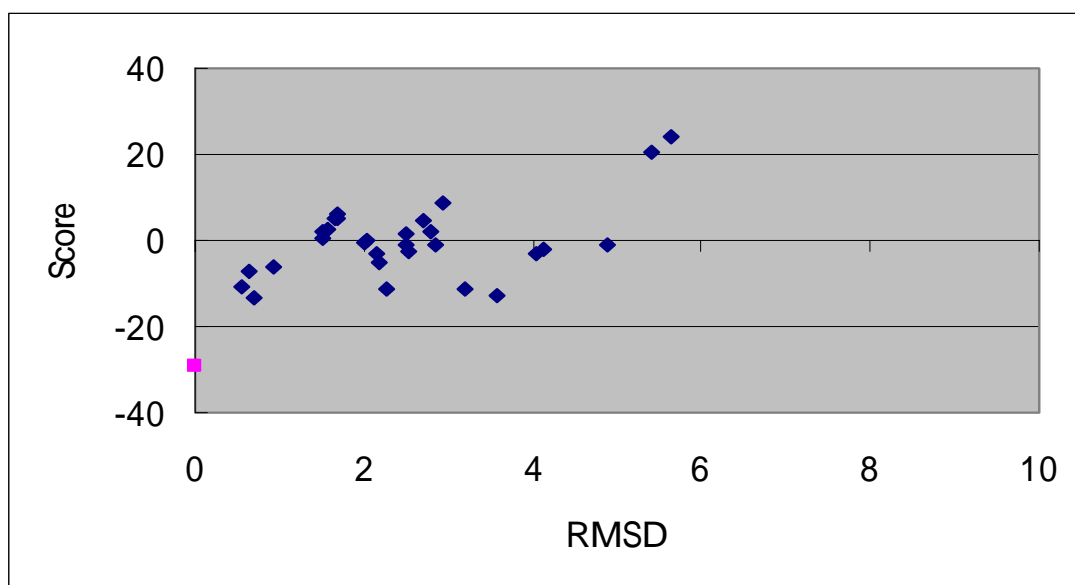


Figure 9 Evaluation of the RDHSP on decoy sets from Decoys'R'Us

1hda-B



RMSD = 0 : Native structure

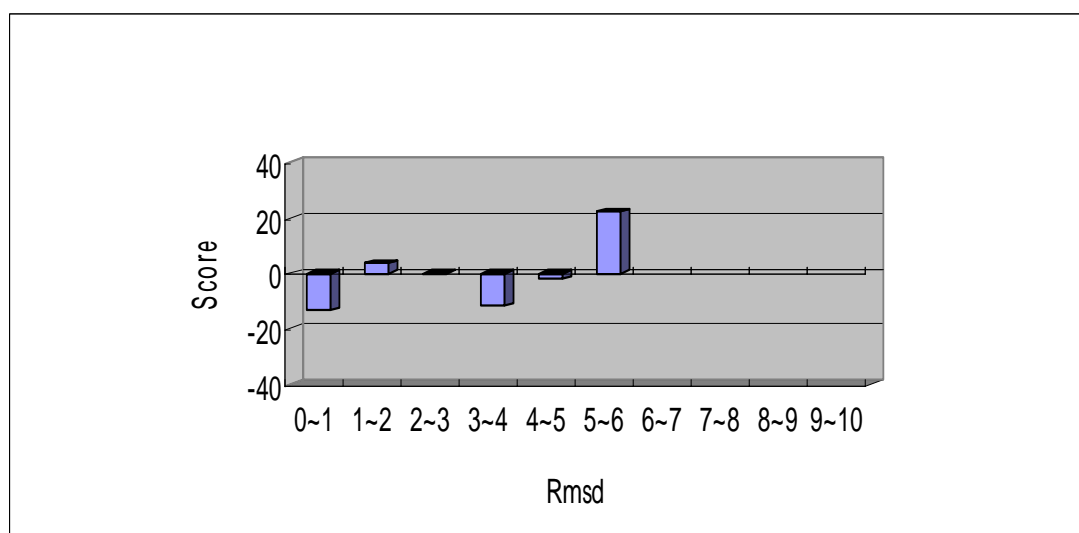


Figure 10 A specificity-sensitivity curve at Family level

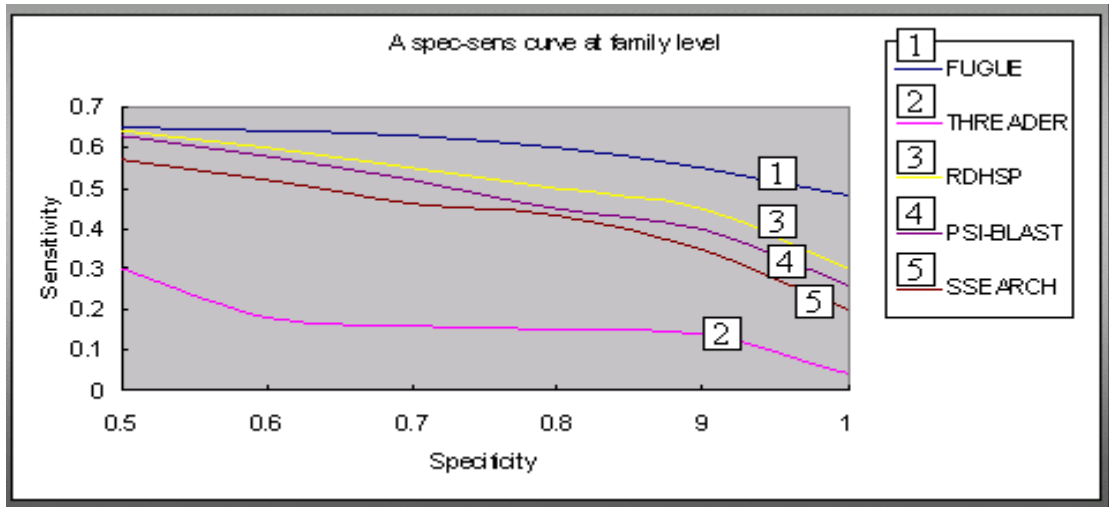


Figure 11 A specificity-sensitivity curve at Superfamily level

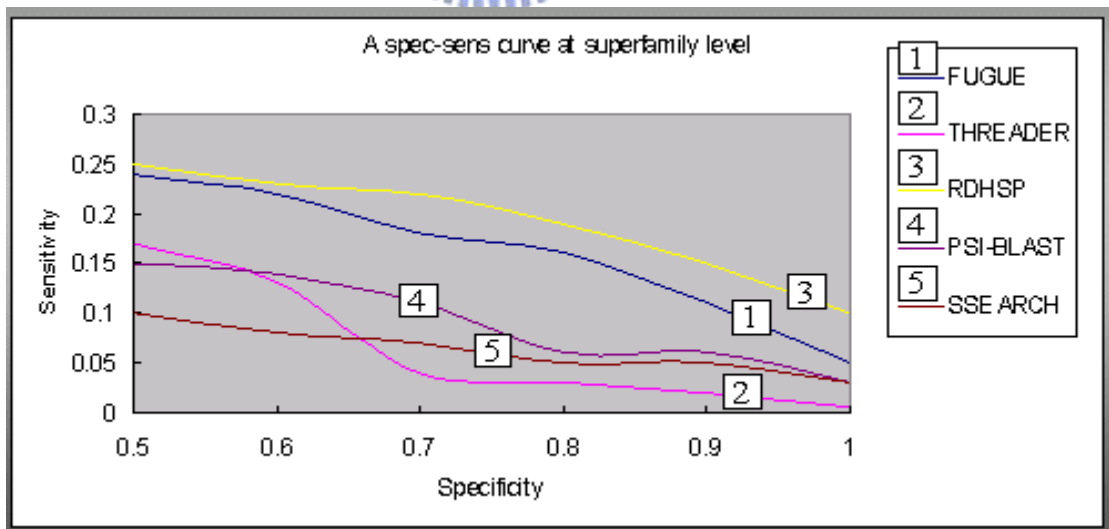


Figure 12 A specificity-sensitivity curve at fold level

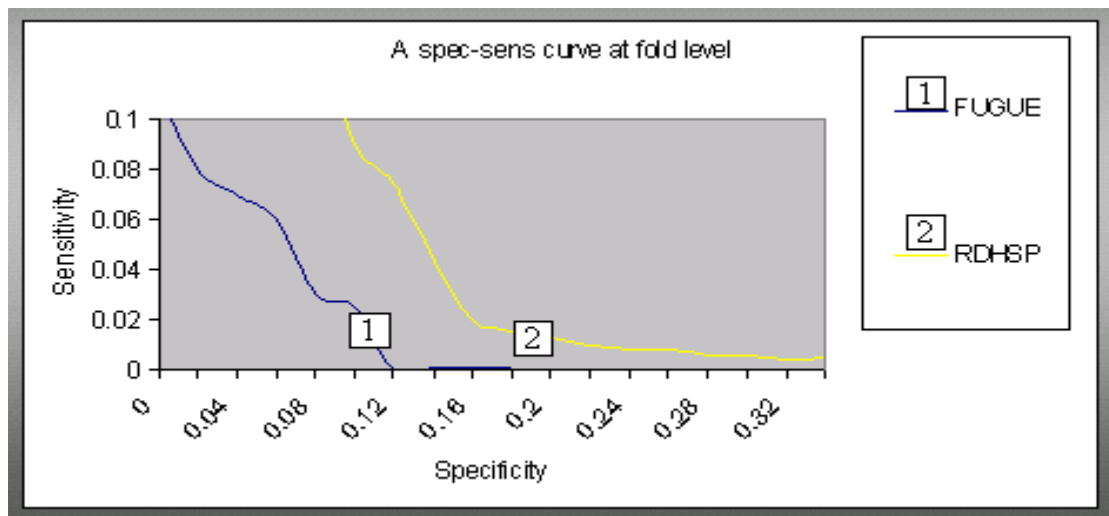


Table 1 The 387 proteins for training set. Those consisting of sequences varying in length from 44 to 1017, with low sequence homology and covering many different three-dimensional-folds according to the Structure Classification of Protein (SCOP) classification [3]. Additional criteria used in selecting the proteins in the training set were follows:

- (1)The protein structure was obtained through x-ray crystallography
- (2)The structures were monomeric.
- (3)The determined structures missed no more than two amino acid

1CE	1KCW	4HB1	1DHX	1BLE	1SQC	1AB4	1PKP	1FY	1RGS	1TDJ	1FSZ	8LDH	6ICD
1AEP	1ANV	5PTD	1AGI	1B6E	1SIG	1EDV	1KXU	1BVB	1BAJ	1A6F	1LRV	2ITG	1C8Y
1LXA	1914	1CC5	1AGP	1OHK	1JON	1PR	1A7J	1AUA	1GLN	1HF	1HLE	1ZAP	2STV
1A17	1RDR	1RLW	1UBY	2SAS	1BCO	1ASY	1AX8	2LIV	1ANS	2OMF	1A41	1C25	1AK5
1AQT	1A36	2FXB	1BOB	1INP	1CYX	1XSM	1BIA	1CPT	1OPR	1PLQ	1AUQ	1KIT	1CTN
1DHR	1OBR	1RCB	1A26	1CTY	1GPC	1PFO	1GRJ	1BY9	1MAZ	1LBA	1KTE	1AM2	1BB9
1HTP	1BIX	1TUL	1DRW	1AQE	2GSQ	1DHS	5EAU	1CFR	1GEN	1BR9	1ACC	1YGS	1B5L
1HAM	1A32	1RMD	1PEA	1SEK	1KLO	1OXA	1CRB	2TCT	1ESC	1TFR	2NG1	1LCI	1PHT
1ALY	1VIN	1A6Q	1A76	1A1X	1CSN	1TIG	1ASH	1BTN	1CDY	1CFB	1MSC	1AMX	1HOE
1UOX	2PGD	1BV1	2FLC	4MT2	1SRA	1DDT	1NSJ	1UOK	1POC	1SUR	1GOX	1GSA	1MJC
2PIA	1LKI	1BY2	1SKF	1BIF	1PBV	1ALO	1RMG	2IB	1DPE	1AJ2	4PAH	1FCE	1PNE
1BF2	1AZ9	1A53	3TDT	7TAA	1OPC	1PTQ	1BEA	1PUC	1FUA	1RSS	1ECL	1SKZ	1NEU
1ALU	1CUK	1CA1	1MAI	1AD2	1OPY	1EDT	1BHE	1DW	1PHM	1DXY	1VOM	1CEO	1A8L
1TMY	1SVB	1AIL	1WHO	1JDC	1SFP	3TSS	1DUN	1IOW	1PBE	1GPR	1A48	4ENL	2PII
3GCB	1BG7	1VLS	1FUD	2ABK	1MDL	1RKD	1EUR	1DMR	1GND	1UCH	1BG2	1AK0	1UXY
2GAR	1LCL	1MML	1POT	1QNF	1NFK	1AYL	1TIF	1BDS	1BDO	1BG6	1C3D	1HYF	2POR
1UAE	1BJ7	1TML	1TYV	1HCL	2SAK	1FNA	1AL3	2TGI	2ACY	1LST	1LBU	1AMP	1NAR
1FAS	2CBP	1FMB	1AXN	1TUD	1PDA	1HAI	1CV8	1CHD	1AMF	1USH	1CPQ	1BMS	1XWL
1BGC	1AJJ	1TPE	1NKR	1IDO	1VJS	1BHP	1WAB	1VIE	1VHH	1GCA	1PDO	1FDR	1PMI
1SBP	1GOF	1AKO	1MOF	2GDM	1FXD	1FNC	1GAI	2HFT	1OSA	1VNS	3CHY	1ERV	1DHN
1AQB	1CNV	119L	1CEM	1CXK	1VCC	1GVP	2DRI	1MBA	1A3C	1EDG	1PHF	16PK	451C
1B6A	1BKF	1RZL	5NUL	1AOP	1ABE	1CVL	1ARV	1NF	3CYR	1MRJ	1ZIN	1LAM	1CSH
1KUH	1PTF	1BFG	1BFD	3PTE	2AYH	2MYR	1NOX	1AKR	2AOB	1ASD	1MOQ	1HFC	1RA9
1TCA	3GR5	2CBA	1KFF	5CB	1AIE	1KOE	1WHI	1RIE	1MLA	1HKA	1OPD	1FLP	2MCM
1CYO	1FOA	1BRT	2HBG	2SNS	1XNB	2RN2	3SEB	1BGF	2END	1YGE	3VUB	2CTC	1HMT
1PFT	1BQK	1UTG	1PLC	1BK0	1DC5	1CS2	7RSA	1OAA	1MSI	1YCC	2PTH	2SN3	1ANM
1BX7	1ATG	2KNT	1MUN	1A75	1CTJ	1BS9	2IGD	1NKD	3SIL	2ERL	1A6M	1CEX	1DGH
1BYI	1AHO	1NLS	2FDN	3LZT	1RB9	3PYF	1CBN	1GCI					

Table 2 The environment-specific substitution tables (α -helix). The environment scores for each amino acid. Large negative values indicate a strong preference for the particular environment whereas large positive values indicate an aversion.

		W	F	Y	L	I	V	M	A	G	P	C	T	S	Q	N	E	D	H	K	R
α -helix(A)	ECn1Ce1	-0.2	0.16	-0.2	0.16	0.29	-2.01	1.1	-1.01	0.09	1.04	-1.76	-2.15	-2.48	-2.04	-0.44	-1.13	0.16	0.09	1.04	-1.76
α -helix(A)	ECn1Ce2	-1.12	0.65	0.56	-0.11	-1.33	0.16	0.84	-0.71	0.32	-0.4	-2.18	-1.59	-1.8	-0.17	-0.82	0.09	0.49	0.32	-0.4	-2.18
α -helix(A)	ECn1Ce3	-0.41	0.34	0.16	-0.41	-0.82	-0.79	0.71	-0.61	0.01	-1.25	-2.61	-2.01	-2.63	-0.64	-0.07	0.61	0.16	0.01	-1.25	-2.61
α -helix(A)	ECn1Ce4	-1.59	0.41	0.49	-0.44	-1.23	-0.87	-0.2	0.16	-0.79	-1.66	-0.71	0.52	-0.33	-0.42	-1.11	0.8	0.65	-0.79	-1.66	-0.71
α -helix(A)	ECn1Ce5	-2.01	-0.48	0.56	1.13	-1.01	0.16	-1.12	0.65	-0.87	-1.38	-2.01	-1.13	-1.01	-0.46	-0.01	1.04	0.34	-0.87	-1.38	-2.01
α -helix(A)	ECn1Ce6	-0.58	-1.59	0.35	0.84	-0.71	0.52	-0.8	-2.36	0.16	0.16	0.16	0.84	0.16	-0.2	0.16	0.29	0.41	0.16	0.16	0.16
α -helix(A)	ECn1Ce7	0.71	-0.61	0.1	-2.63	-0.81	-0.2	-1.52	-0.2	-0.41	-0.82	-0.79	0.71	0.65	0.56	-0.11	-1.33	-0.48	-0.41	-0.82	-0.79
α -helix(A)	ECn2Ce1	-0.2	0.16	0.29	-0.33	-0.44	-1.13	-0.81	-1.12	-0.44	-1.23	-0.87	-0.2	0.34	0.16	-0.41	-0.82	-0.07	-0.44	-1.23	-0.87
α -helix(A)	ECn2Ce2	-1.12	0.65	-0.12	0.09	-0.82	0.09	-1.01	-0.41	-0.82	0.09	0.09	0.74	-0.54	0.49	-0.44	-1.23	-0.48	1.96	-1.01	0.16
α -helix(A)	ECn2Ce3	-0.41	0.34	-0.03	0.32	0.6	-0.33	-0.71	-0.44	-1.23	-0.33	-0.58	-0.8	-2.36	0.56	1.65	-1.01	-1.59	0.84	-0.71	0.52
α -helix(A)	ECn2Ce4	-1.59	0.41	-0.5	0.01	-0.37	0.09	-0.61	1.45	-1.01	0.09	-2.04	-1.52	-0.8	-2.36	0.16	0.01	-1.25	0.84	0.16	-0.41
α -helix(A)	ECn2Ce5	-2.01	-0.48	-0.58	0.41	0.04	0.32	0.16	0.84	-0.71	0.32	0.83	-0.8	-1.52	-0.2	-0.41	-0.79	-1.66	0.71	0.65	0.09
α -helix(A)	ECn2Ce6	0.39	-0.07	0.09	0.74	-0.54	0.01	0.65	0.71	-0.61	0.01	-1.52	1.15	-0.33	-0.58	-0.8	-0.87	-1.38	-0.2	0.34	-1.59
α -helix(A)	ECn2Ce7	-2.01	-0.48	-0.58	-0.8	-2.36	-0.79	0.34	-0.2	0.16	-0.79	0.56	0.83	0.09	-2.04	-1.52	0.84	-0.71	0.74	-0.54	-0.03
α -helix(A)	ECn3Ce1	-0.58	-1.59	-2.04	-1.52	-2.22	-0.87	0.41	-1.12	0.65	-0.87	0.32	-0.4	0.32	0.83	-0.8	-2.63	-0.81	-0.8	-2.36	-0.5
α -helix(A)	ECn3Ce2	-0.2	0.56	0.83	-0.8	-2.36	0.16	0.83	-0.8	-2.36	0.16	0.01	-1.25	0.01	-1.52	1.85	-0.33	-0.44	-1.52	-2.22	-0.58
α -helix(A)	ECn3Ce3	-1.13	0.16	-1.52	1.73	-1.01	-0.41	-0.4	-1.52	-0.2	-0.41	-0.79	-1.66	-0.79	-0.81	0.84	0.09	-0.82	-0.8	-2.36	-2.04
α -helix(A)	ECn3Ce4	0.09	0.49	-0.81	0.84	0.71	-0.44	-1.23	-0.33	-0.58	-0.8	0.87	-1.38	-0.87	-0.44	0.71	0.61	1.84	1.39	-1.01	0.34
α -helix(A)	ECn3Ce5	-2.48	-2.04	-0.44	0.71	-0.61	1.86	-1.01	0.09	-2.04	-1.52	0.84	-0.71	0.16	-0.82	-0.2	0.16	0.84	0.83	-0.8	0.41
α -helix(A)	ECn3Ce6	-1.8	-0.17	-0.81	-0.2	-1.52	-0.2	-0.41	-0.82	0.83	-0.8	-2.63	-0.81	-0.41	-0.07	-1.12	0.65	0.71	-0.4	-1.52	-0.48
α -helix(A)	ECn3Ce7	-2.63	-0.64	-0.44	-1.13	-0.81	-1.12	-0.44	-1.23	-1.52	1.1	-0.33	-0.44	-0.44	-1.25	-0.41	0.34	-0.2	-1.25	-0.81	-0.07
α -helix(A)	ECn4Ce1	-0.8	-2.36	-0.82	0.09	-1.01	-0.41	-0.82	0.09	-0.81	0.84	0.09	-0.82	-0.8	-2.36	-0.58	-0.8	-1.59	-0.44	-1.23	-0.48
α -helix(A)	ECn4Ce2	0.16	-0.87	0.6	-0.33	-0.71	-0.44	-1.23	-0.33	-0.44	0.71	-0.61	1.21	1.64	-1.01	-0.41	1.78	-1.01	1.94	-1.01	-1.59
α -helix(A)	ECn4Ce3	-0.48	-0.58	-0.37	0.09	0.61	1.1	-1.01	0.09	-0.82	-0.2	0.16	0.84	0.83	-0.8	-2.36	0.16	-0.71	0.52	-0.33	-0.44
α -helix(A)	ECn4Ce4	-0.03	0.41	0.04	0.32	0.16	0.84	-0.71	0.52	-0.07	-1.12	0.65	0.71	-0.4	-1.52	-0.2	-0.41	-2.01	-0.8	-1.59	-1.8
α -helix(A)	ECn4Ce5	-0.5	0.74	-0.54	0.01	0.65	0.71	-0.61	0.01	-1.25	-0.41	0.34	-0.2	-1.25	-0.81	-1.12	-0.44	0.16	-0.33	-0.87	-0.2
α -helix(A)	ECn4Ce6	-0.58	-0.8	-2.36	-0.79	0.34	-0.2	0.16	-0.79	-0.58	-1.59	0.41	-1.12	0.84	0.09	-0.82	-2.01	-1.13	-1.01	0.16	-1.12
α -helix(A)	ECn4Ce7	0.09	-1.52	-2.22	-0.87	0.41	-1.12	0.65	-0.87	-0.2	0.56	0.83	-0.8	0.71	-0.61	1.54	0.16	0.84	0.16	0.52	-0.41
α -helix(A)	EG	-0.03	0.32	0.6	-0.33	-0.71	-2.04	-1.52	0.84	-0.71	0.74	-0.54	0.32	0.16	0.84	-0.71	-1.01	-0.41	-0.4	-1.52	-0.2
α -helix(A)	FCn1Ce1	-0.5	0.01	-0.37	0.09	-0.61	0.83	-0.8	-2.63	-0.81	-0.8	-2.36	0.01	0.65	0.71	-0.61	-0.71	-0.44	-1.23	-0.81	-1.12
α -helix(A)	FCn1Ce2	-0.58	0.41	0.04	0.32	0.16	-1.52	1.1	-0.33	-0.44	-1.52	-2.22	-0.79	0.34	-0.2	0.16	-0.61	1.15	-1.01	-1.01	0.41
α -helix(A)	FCn1Ce3	0.09	0.74	-0.54	0.01	0.65	-0.82	-0.79	0.71	0.65	0.56	-0.11	-0.87	0.41	-1.12	0.65	-1.52	-0.2	-0.41	-0.71	-0.44
α -helix(A)	FCn1Ce4	-0.58	-0.8	-2.36	-0.79	0.34	-1.23	-0.87	-0.2	0.34	0.16	-0.41	0.16	0.83	-0.8	-2.36	-0.81	-1.12	-0.44	-0.61	1.19
α -helix(A)	FCn1Ce5	-2.04	-1.52	-2.22	-0.87	0.41	0.09	0.09	0.74	-0.54	0.49	-0.44	-0.41	-0.4	-1.52	-0.2	-1.01	-0.41	-0.82	0.16	0.84
α -helix(A)	FCn1Ce6	0.83	-0.8	-2.36	0.16	0.83	-0.33	-0.58	-0.8	-2.36	0.56	1.1	-0.44	-1.23	-0.33	-0.58	-0.71	-0.44	-1.23	0.65	0.71
α -helix(A)	FCn1Ce7	-1.52	1.94	-1.01	-0.41	-0.4	0.09	-2.04	-1.52	-0.8	-2.36	0.16	2.54	-1.01	0.09	-2.04	-0.61	1.36	-1.01	0.34	-0.2
α -helix(A)	FCn2Ce1	-0.81	0.84	0.71	-0.44	-1.23	0.32	0.83	-0.8	-1.52	-0.2	-0.41	-0.2	-0.41	-0.82	0.83	0.16	0.84	0.71	0.41	-1.12
α -helix(A)	FCn2Ce2	-0.44	0.71	-0.61	1.42	-1.01	0.01	-1.52	1.08	-0.33	-0.58	-0.8	-1.12	-0.44	-1.23	-1.52	0.65	0.71	-0.61	-0.71	-2.04
α -helix(A)	FCn2Ce3	-0.81	-0.2	-1.52	-0.2	-0.41	-0.79	0.56	0.83	0.09	-2.04	-1.52	-0.41	-0.82	0.09	-0.81	0.34	-0.2	0.16	-0.79	-0.33
α -helix(A)	FCn2Ce4	0.74	-0.54	0.49	-0.71	0.16	-0.44	-1.23	-0.4	0.32	0.83	-0.8	-0.61	1.85	-1.01	0.09	0.41	-1.12	0.65	0.87	0.09
α -helix(A)	FCn2Ce5	-0.8	-2.36	0.56	-0.81	-0.41	1.53	-1.01	-1.25	0.01	-1.52	1.66	0.16	0.84	-0.71	0.32	0.83	-0.82	-2.01	-1.13	0.32
α -helix(A)	FCn2Ce6	-1.52	-0.8	-2.36	-0.44	-0.44	0.16	0.01	0.09	-2.04	-1.52	0.16	-0.82	0.71	-0.61	0.01	-1.52	1.88	0.16	0.84	0.01
α -helix(A)	FCn2Ce7	-0.8	-1.52	-0.2	0.32	0.6	0.32	-0.4	0.32	0.83	-0.8	-0.41	0.07	-0.2	0.16	-0.79	0.56	-0.71	-1.01	-0.41	0.79
α -helix(A)	FCn3Ce1	1.17	-0.33	-0.58	0.01	-0.37	0.01	-1.25	0.01	-1.52	1.1	-0.44	-1.25	-1.12	0.65	-0.87	0.32	-0.61	-0.71	-0.44	-0.87
α -helix(A)	FCn3Ce2	0.83	0.09	-2.04	0.41	0.04	-0.79	-1.66	-0.79	-0.11	-1.33	-0.8	-2.36	-0.8	-2.36	0.16	0.01	0.16	-0.61	1.76	0.84
α -helix(A)	FCn3Ce3	-0.4	0.32	0.83	0.74	-0.54	-0.87	-1.38	-0.87	-0.41	-0.82	1.76	-1.01	-1.23	0.49	-0.71	0.16	-0.2	-0.82	0.09	-1.01
α -helix(A)	FCn3Ce4	-1.25	0.01	-1.52	-0.8	-2.36	0.84	-0.71	0.16	-0.44	-1.23	0.83	-0.8	-1.01	0.56	-0.81	-0.41	1.23	0.6	0.33	0.71
α -helix(A)	FCn3Ce5	-1.66	-0.79	-0.81	-1.52	-2.22	-2.63	-0.81	-0.41	1.1	-1.01	-1.59	0.84	-0.71	-2.36	-0.44	-0.44	0.16	-0.37	0.09	-0.61
α -helix(A)	FCn3Ce6	-1.38	-0.87	-0.44	-0.8	-2.36	-0.33	-0.44	-0.44	0.16	0.01	-1.25	0.84	0.16	-0.2	0.32	0.6	0.32	0.04	0.32	0.16
α -helix(A)	FCn3Ce7	-0.71	0.16	-0.82	1.43	-1.01	0.09	-0.82	-0.8	-0.41	-0.79	-1.66	0.71	0.65	-0.58	0.01	-0.37	0.01	-0.54	0.01	0.65
α -helix(A)	FCn4Ce1	-0.81	-0.41	-0.07	0.84	-0.71	-0.61	1.86	1.1	-0.8	-0.87	-1.38	-0.2	0.34	-1.59	-0.81	-0.41	0.16	-2.36	-0.79	0.34
α -helix(A)	FCn4Ce2	0.32	0.83	-0.44	0.71	-0.61	1.1	-2.04	-1.52	-0.8	-0.79	0.71	0.65	0.56	-0.11	-0.44	-0.44	0.32	-0.4	0.32	-0.2
α -helix(A)	FCn4Ce3	0.41	-1.12	0.65	-0.87	0.32	0.2	0.83	-0.8	-1.52	-0.87	-0.2	0.34	0.16	-0.41	0.32	0.6	0.01	-1.25	0.01	0.58
α -helix(A)	FCn4Ce4	0.83	-0.8	-2.36	0.16	0.01	-1.12	-1.52	1.18	-0.33	0.09	0.74	-0.54	0.49	-0.44	0.01	0.84	-0.71	0.16	-0.79	-2.04
α -helix(A)	FCn4Ce5	-0.4	-1.52	-0.2	-0.41	-0.79	-0.41	0.56	0.83	0.09	-0.58	-0.8	-2.36	0.56	1.62	-0.44	-2.63	-0.81	-0.41	-0.87	0.83
α -helix(A)	FCn4Ce6	-1.23	-0.33	-0.58	-0.8	-0.58	0.01	-0.37	0.01	-1.25	-2.04	-1.52	-0.8	-2.36	0.16	1.18	-0.33	-0.44	-0.44	0.16	-1.52
α -helix(A)	FCn4Ce7	-1.01	0.09	-2.04	-1.52	-2.04	0.41	0.04	-0.79	-1.66	0.83	-0.8	-1.52	-0.2	-0.41	-0.2	0.09	-0.82	-0.8	-1.32	-1.52
α -helix(A)	FG	0.49	-0.71	0.16	-0.2	-0.82	0.84	0.56	1.23	-0.44	-1.23	0.49	-0.71	0.16	-0.44	-1.23	-0.4	0.32	0.83	1.63	0.16

Table 3 The environment-specific substitution tables (β -sheet).The environment scores for each amino acid. Large negative values indicate a strong preference for the particular environment whereas large positive values indicate an aversion.

		W	F	Y	L	I	V	M	A	G	P	C	T	S	Q	N	E	D	H	K	R	
β -sheet (B)	ECn1Ce1	1.11	1.28	0.27	1.3	1.11	0.74	1.26	-0.77	-2.22	-1.56	-0.43	-1.72	-2.43	-1.38	-1.76	-2.15	-2.48	-0.34	-1.37	-1.8	
β -sheet (B)	ECn1Ce2	0.92	0.96	0.17	1.07	1.5	1.18	0.51	-1.05	-2.35	-0.77	-0.45	-1.27	-2.56	-2.03	-2.18	-1.59	-1.8	-2.26	-3.04	-1.52	
β -sheet (B)	ECn1Ce3	0.96	1.4	0.52	1.06	0.93	-1.48	0.91	-0.54	-2.78	0.59	-0.59	-1.41	-2.99	-0.84	-2.61	-2.01	-2.63	-0.61	-2.78	-2.35	
β -sheet (B)	ECn1Ce4	1.01	0.87	0.86	0.71	0.55	0.41	1.02	-0.65	-2.04	-0.97	0.15	-0.67	-1.33	0.16	-0.48	-0.38	-0.8	0.82	-0.94	-0.11	
β -sheet (B)	ECn1Ce5	0.83	1.32	1.3	0.36	1.07	0.71	0.49	-1.52	-2.22	-0.86	-0.72	-1.14	-0.82	-0.79	-0.26	-0.2	-2.08	-0.05	-0.83	-0.41	
β -sheet (B)	ECn1Ce6	1.62	1.4	1.14	0.77	0.81	0.66	-0.15	-0.81	-1.71	-0.07	-0.62	-1.03	-1.23	-0.87	-0.56	-1.13	-1.7	0.54	2.12	-0.44	
β -sheet (B)	ECn1Ce7	0.86	-0.22	0.5	0.16	0.02	-0.29	0.87	-0.44	-1.09	-1.11	-1.38	-0.69	-1.01	0.16	-0.07	0.09	-0.43	0.61	0.56	1.1	
β -sheet (B)	ECn2Ce1	0.07	0.37	1.09	0.14	0.26	0.16	-0.68	-1.08	-2.29	-0.01	-0.79	-0.1	-0.71	0.52	-0.33	-0.42	-0.76	0.8	0.35	0.84	
β -sheet (B)	ECn2Ce2	1.12	0.71	1.25	0.29	-0.54	-0.4	0.23	-0.87	-0.61	-0.11	-0.98	-0.48	-0.61	0.1	0.09	-0.46	-0.83	1.04	0.08	0.71	
β -sheet (B)	ECn2Ce3	-1.26	-1.81	-1.7	-1.37	-2.36	-1.25	-0.9	0.44	0.63	0.05	-0.17	-0.2	0.16	0.29	0.32	0.6	0.44	-0.06	0.07	-0.2	
β -sheet (B)	ECn2Ce4	0.81	-0.83	-0.03	-1.6	-1.39	-1.66	-0.62	0.14	1.75	-0.88	-0.04	-0.17	0.65	-0.12	0.01	-0.37	-0.3	-0.76	-1.54	-1.12	
β -sheet (B)	ECn2Ce5	-2.06	-1.63	-1.04	-1.14	-1.63	-0.8	-1.3	0.16	1.1	0.25	-0.35	0.08	0.34	-0.03	0.41	0.04	0.23	-0.41	-0.1	-0.41	
β -sheet (B)	ECn2Ce6	-1.29	0.07	1.81	1.81	4.32	-0.83	3.63	-1.24	0.49	-0.85	-1.24	3.63	0.12	0.69	-1.55	-0.1	-1.63	0.12	0.47	0.3	
β -sheet (B)	ECn2Ce7	-0.9	-0.35	2.33	2.33	1.37	-0.5	1.51	-0.26	0.74	-0.28	-0.26	1.51	0.56	-0.19	0.17	0.65	0.13	0.56	0.79	0.76	
β -sheet (B)	ECn3Ce1	0.41	0.32	1.64	1.64	1.67	-1	-1.02	0.57	0.58	0.5	0.75	-1.02	-1.18	0.27	-0.23	0.11	-1.71	0.59	0.74	0.57	
β -sheet (B)	ECn3Ce2	-0.5	-0.27	0.38	0.38	0.94	0.39	0.61	-1.05	-0.47	0.56	-1.05	0.61	1.06	0.87	0.34	0.08	-0.11	-1.3	1.32	0.89	0.63
β -sheet (B)	ECn3Ce3	0.42	0.06	-0.12	-0.12	1.32	-0.64	0.89	-0.28	0.41	0.74	-0.28	0.89	0.11	0.54	-1.62	0.95	0.91	-0.74	-0.46	1.1	
β -sheet (B)	ECn3Ce4	-0.26	-0.36	0.65	0.65	0.55	0.71	1.26	-0.24	-1.06	-0.95	-0.24	1.26	0.93	0.14	0.15	0.12	-1.59	0.7	1.18	0.47	
β -sheet (B)	ECn3Ce5	-0.33	-0.16	0.09	0.09	0.48	0.83	-0.76	-0.54	0.63	0.08	-0.54	-0.76	0.65	0.23	0.48	-0.11	-1.3	1.32	0.89	0.63	
β -sheet (B)	ECn3Ce6	0.36	1.16	0.73	0.73	1.22	0.16	0.14	-0.49	-0.25	0.35	-0.49	0.14	0.47	2.23	-1.13	0.6	-1.01	0.11	0.12	0.6	
β -sheet (B)	ECn3Ce7	0.13	0.83	0.82	-0.06	0.62	-1.88	-0.23	0.23	-0.57	-0.08	0.23	-0.23	0.43	1.15	-1.64	0.57	-1.64	0.35	0.75	0.73	
β -sheet (B)	ECn4Ce1	-0.4	-0.05	2.23	-0.13	0.27	0.5	-0.15	0.39	-1.17	-0.99	2.38	-0.15	0.72	0.74	-0.97	1.02	-1.26	-0.45	-0.14	-0.77	
β -sheet (B)	ECn4Ce2	1.05	-0.6	1.74	0.82	0.32	0.56	0.14	2.38	-0.1	-0.32	1.82	-1.91	-0.31	1.59	-0.39	-0.5	-1.13	-0.25	0.32	0.27	
β -sheet (B)	ECn4Ce3	-0.29	-0.79	0.67	1.24	-0.42	0.93	1.32	1.82	-0.48	-0.34	1.6	-1.35	-0.18	1.24	0.32	0.64	-0.3	1.06	1.99	0.2	
β -sheet (B)	ECn4Ce4	-0.31	-0.01	0.8	0.85	0.25	0.78	-0.75	1.6	-1.3	-0.48	1.32	-0.67	-0.38	0.81	-1.26	-0.47	-1.88	0.29	0.83	0.48	
β -sheet (B)	ECn4Ce5	0.8	0.49	1.53	0.91	-1.48	-0.5	-1.91	1.32	-1.61	-0.62	-1.02	-1.04	-0.46	-1.04	-1.16	-0.32	-1.84	-0.83	0.03	0.62	
β -sheet (B)	ECn4Ce6	0.67	-0.66	2.34	-0.66	1.18	0.6	-1.35	-1.02	-1.35	0.13	0.55	-1.35	-0.87	-0.68	-1.2	0.42	-0.57	-0.44	0.11	-0.41	
β -sheet (B)	ECn4Ce7	2.35	-0.28	1.31	-0.46	-1.78	1.43	-0.67	0.32	-0.77	-0.32	0.74	-1.97	-0.57	-0.56	-0.34	-0.28	0.23	-0.03	0.79	-0.6	
β -sheet (B)	EG	0.32	0.6	0.44	-0.06	0.07	-1.29	0.07	1.81	1.81	4.32	-0.83	0.34	-0.03	0.41	0.04	0.23	-0.41	-0.1	-0.41	1.51	
β -sheet (B)	FCn1Ce1	0.01	-0.37	-0.3	-0.76	-1.54	-0.9	-0.35	2.33	2.33	1.37	-0.5	0.12	0.69	-1.55	-0.1	-1.63	0.12	0.47	0.3	-1.02	
β -sheet (B)	FCn1Ce2	0.41	0.04	0.23	-0.41	-0.1	0.41	0.32	1.64	1.64	1.67	-1	0.56	-0.19	0.17	0.65	0.13	0.56	0.79	0.76	0.61	
β -sheet (B)	FCn1Ce3	-1.55	-0.1	-1.63	0.12	0.47	-0.5	-0.27	0.38	0.38	0.94	0.39	-0.18	0.27	-0.23	0.11	-1.71	0.59	0.74	0.57	0.89	
β -sheet (B)	FCn1Ce4	0.17	0.65	0.13	0.56	0.79	0.42	0.06	-0.12	1.32	-0.64	1.06	0.87	0.34	0.08	-0.11	0.65	-0.25	0.46	1.26	0.6	
β -sheet (B)	FCn1Ce5	0.74	-0.28	-0.26	1.51	0.56	-0.26	-0.36	0.65	0.65	0.55	0.71	0.11	0.54	-1.62	0.95	0.91	-0.74	-0.46	1.1	-0.76	
β -sheet (B)	FCn1Ce6	0.58	0.5	0.75	-1.02	-0.18	-0.67	-0.38	0.27	0.5	-0.15	0.59	0.93	0.14	0.15	0.12	-1.59	0.7	1.18	0.47	0.14	
β -sheet (B)	FCn1Ce7	-0.47	0.56	-1.05	0.61	1.06	-1.04	-0.46	0.32	0.56	0.14	2.38	0.65	0.23	0.48	-0.11	-1.3	1.32	0.89	0.63	-0.23	
β -sheet (B)	FCn2Ce1	0.41	0.74	-0.28	0.89	0.11	-1.35	-0.87	-0.42	0.93	1.32	1.82	-0.48	-0.34	1.6	-1.35	-0.18	1.07	0.71	0.49	-0.15	
β -sheet (B)	FCn2Ce2	-1.06	-0.95	-0.24	1.26	0.93	-1.97	-0.57	0.25	0.78	-0.75	1.6	-1.3	-0.48	1.32	-0.67	-0.38	0.81	0.66	-0.15	-0.81	
β -sheet (B)	FCn2Ce3	0.63	0.08	-0.54	-0.76	0.65	0.34	-0.03	-1.48	-0.5	-1.91	1.32	-1.61	-0.62	-1.02	-1.04	-0.46	0.02	-0.29	0.87	-0.44	
β -sheet (B)	FCn2Ce4	-0.4	0.23	-0.87	-0.61	-0.11	0.12	0.69	1.18	0.6	-1.35	-1.02	-1.35	0.13	0.55	-1.35	-0.87	0.26	0.16	-0.68	-1.08	
β -sheet (B)	FCn2Ce5	-1.25	-0.9	0.44	0.63	0.05	0.56	-0.19	-1.78	1.43	-0.67	0.32	-0.77	-0.32	0.74	-1.97	-0.57	-0.54	-0.4	0.23	-0.87	
β -sheet (B)	FCn2Ce6	-1.66	-0.62	0.14	1.75	-0.88	-0.04	-0.79	-0.1	-0.71	0.52	-0.47	-1.88	0.55	0.41	1.02	-0.65	-2.36	-1.25	-0.9	0.44	
β -sheet (B)	FCn2Ce7	-0.8	-1.3	0.16	1.1	0.25	-0.35	-0.98	-0.48	-0.61	0.1	-0.32	-1.84	1.07	0.71	0.49	-1.52	-1.39	-1.66	-0.62	0.14	
β -sheet (B)	FCn3Ce1	-0.83	3.63	-1.24	0.49	-0.85	-1.24	-0.17	-0.2	0.16	0.29	-0.42	-0.57	0.81	0.66	-0.15	-0.04	-0.79	-0.1	-0.71	-0.56	
β -sheet (B)	FCn3Ce2	-0.5	1.51	-0.26	0.74	-0.28	-0.26	-0.04	-0.17	0.65	-0.12	-0.28	0.23	0.02	-0.29	0.87	-0.35	-0.98	-0.48	-0.61	0.41	
β -sheet (B)	FCn3Ce3	-1	-1.02	0.57	0.58	0.5	0.75	-0.35	0.08	0.34	-0.03	0.23	-0.41	0.26	0.16	-0.68	-1.24	-0.17	-0.2	0.16	-1.55	
β -sheet (B)	FCn3Ce4	0.39	0.61	-1.05	-0.47	0.56	-1.05	-1.24	3.63	0.12	0.69	-1.63	0.12	-0.54	-0.4	0.23	-0.26	-0.04	-0.17	0.65	0.17	
β -sheet (B)	FCn3Ce5	-0.64	0.89	-0.28	0.41	0.74	-0.28	-0.26	1.51	0.56	-0.19	0.13	0.56	-2.36	-1.25	-0.9	0.44	0.34	-0.03	-1.48	-0.23	
β -sheet (B)	FCn3Ce6	0.71	1.26	-0.24	-1.06	-0.95	-0.24	-0.64	1.06	0.87	0.34	0.08	-0.11	0.5	0.75	-1.02	-0.18	0.12	0.69	1.18	-0.04	
β -sheet (B)	FCn3Ce7	-0.46	-0.83	1.04	-0.87	-0.56	-1.13	-1.7	0.54	-2.12	0.61	-1.05	-0.47	0.56	-1.05	0.61	1.06	0.56	-0.19	-1.78	-0.35	
β -sheet (B)	FCn4Ce1	0.6	0.44	-0.06	0.16	-0.07	0.09	-0.43	0.61	0.56	0.89	-0.28	0.41	0.74	-0.28	0.89	0.11	-0.04	-0.79	-0.1	-1.24	
β -sheet (B)	FCn4Ce2	-0.37	-0.3	-0.76	0.52	-0.33	-0.42	-0.76	0.8	0.35	1.26	-0.24	-1.06	-0.95	-0.24	1.26	0.93	-0.35	-0.98	-0.48	-0.26	
β -sheet (B)	FCn4Ce3	0.04	0.23	-0.41	0.1	0.09	-0.46	-0.83	1.04	0.08	-0.76	-0.54	0.63	0.08	-0.54	-0.76	0.65	-0.83	1.04	-0.87	-0.56	
β -sheet (B)	FCn4Ce4	-0.1	-1.63	0.12	0.29	0.32	0.6	0.44	-0.06	0.07	0.14	-0.49	-0.25	0.35	-0.49	0.14	0.47	0.44	-0.06	0.16	-0.07	
β -sheet (B)	FCn4Ce5	0.65	0.13	0.56	-0.12	0.01	-0.37	-0.3	-0.76	-1.54	-0.23	0.23	-0.57	-0.08	0.13	0.55	-1.35	-0.3	-0.76	0.52	-0.33	
β -sheet (B)	FCn4Ce6	0.11	-1.71	0.59	-0.03	0.41	0.04	0.23	-0.41	-0.1	-0.15	0.59	-1.17	-0.99	-0.32	0.74	-1.97	0.23	-0.41	0.1	0.09	
β -sheet (B)	FCn4Ce7	0.08	-0.11	0.65	-0.28	0.23	-0.03	0.65	0.23	0.48	-0.11	-1.3	1.32	0.89	4.32	-0.83	0.34	-0.71	0.52	-0.47	-1.88	
β -sheet (B)	FG	0.95	0.91	-0.74	0.23	-0.41	-0.1	0.47	2.23	-1.13	0.6	-1.01	0.11	0.12	4.32	-0.83	0.34	-0.03	0.41	0.04	0.23	

Table 4 The environment-specific substitution tables (3_{10} -helix).The environment scores for each amino acid. Large negative values indicate a strong preference for the particular environment whereas large positive values indicate an aversion.

		W	F	Y	L	I	V	M	A	G	P	C	T	S	Q	N	E	D	H	K	R
3_{10} -helix (C)	ECn1Ce1	0.65	0.65	0.55	0.71	1.26	0.74	1.26	-0.77	-2.22	-1.56	-0.43	-1.72	-2.43	-1.38	-1.76	-2.15	-2.48	-0.34	-1.37	-1.8
3_{10} -helix (C)	ECn1Ce2	0.09	0.09	0.48	0.83	-0.76	1.18	0.51	-1.05	-2.35	-0.77	-0.45	-1.27	-2.56	-2.03	-2.18	-1.59	-1.8	-2.26	-3.04	-1.52
3_{10} -helix (C)	ECn1Ce3	0.73	0.73	1.22	0.16	0.14	1	0.91	-0.54	-2.78	0.59	-0.59	-1.41	-2.99	-0.84	-2.61	-2.01	-2.63	-0.61	-2.78	-2.35
3_{10} -helix (C)	ECn1Ce4	0.82	-0.06	0.62	-0.81	-1.71	-0.07	0.62	-1.03	-0.87	0.15	-0.67	-1.33	0.16	-0.48	-0.58	-0.8	0.82	-0.94	-0.11	
3_{10} -helix (C)	ECn1Ce5	2.23	-0.13	0.27	-0.44	-1.09	-1.11	-1.38	-0.69	-1.01	0.16	-0.72	-1.14	-0.82	-0.79	-0.26	-0.2	-2.08	-0.05	-0.83	-0.41
3_{10} -helix (C)	ECn1Ce6	1.74	0.82	0.32	-1.08	-2.29	-0.01	-0.79	-0.1	-0.71	0.52	-0.62	-1.03	-1.23	-0.87	-0.56	-1.13	-1.7	0.54	-2.12	-0.44
3_{10} -helix (C)	ECn1Ce7	0.67	1.24	-0.42	-0.87	-0.61	-0.11	-0.98	-0.48	-0.61	0.1	-1.38	-0.69	-1.01	0.16	-0.07	0.09	-0.43	0.61	0.56	1.1
3_{10} -helix (C)	ECn2Ce1	0.8	0.85	0.25	0.44	0.63	0.05	-0.17	-0.2	0.16	0.29	-0.79	-0.1	-0.71	0.52	-0.33	-0.42	-0.76	0.8	0.35	0.84
3_{10} -helix (C)	ECn2Ce2	1.12	0.71	1.25	0.14	1.75	-0.88	-0.04	-0.17	0.65	-0.12	-0.98	-0.48	-0.61	0.1	0.09	-0.46	-0.83	1.04	0.08	0.71
3_{10} -helix (C)	ECn2Ce3	-1.26	-1.81	-1.7	0.16	1.1	0.25	-0.35	0.08	0.34	-0.03	-0.17	-0.2	0.16	0.29	0.32	0.6	0.44	-0.06	0.07	-0.2
3_{10} -helix (C)	ECn2Ce4	0.81	-0.83	-0.03	-1.24	0.49	-0.85	-1.24	3.63	0.12	0.69	-0.04	-0.17	0.65	-0.12	0.01	-0.37	-0.3	-0.76	-1.54	-1.12
3_{10} -helix (C)	ECn2Ce5	-2.06	-1.63	-0.42	0.93	1.32	1.82	-0.48	-0.34	0.56	-0.19	-0.35	0.08	0.34	-0.98	-0.48	-0.61	0.1	-0.32	-1.84	-0.41
3_{10} -helix (C)	ECn2Ce6	-0.11	-0.98	-0.48	0.78	-0.75	1.6	-1.3	-0.48	-0.18	0.27	-1.52	0.84	-0.71	0.16	-0.82	0.16	0.29	1.96	0.93	1.32
3_{10} -helix (C)	ECn2Ce7	0.05	-0.17	-0.2	-0.5	-1.91	1.32	-1.61	-0.62	1.06	0.87	-0.8	-2.63	-0.81	-0.41	-0.07	0.65	-0.12	0.84	0.78	-0.75
3_{10} -helix (C)	ECn3Ce1	-0.88	-0.04	-0.17	0.6	-1.35	-1.02	-1.35	0.13	0.11	0.54	1.1	-0.33	-0.44	-0.44	-1.25	0.34	-0.03	0.84	-0.5	-1.91
3_{10} -helix (C)	ECn3Ce2	0.25	-0.35	0.08	1.43	-0.67	0.32	-0.77	-0.32	0.02	-0.29	0.87	-0.35	-0.98	-0.8	-2.36	0.12	0.69	0.71	0.6	-1.35
3_{10} -helix (C)	ECn3Ce3	-0.85	-1.24	3.63	-1.29	0.07	1.81	1.81	4.32	0.26	0.16	-0.68	-1.24	-0.17	1.64	-1.01	0.56	-0.19	-0.2	1.43	-0.67
3_{10} -helix (C)	ECn3Ce4	-0.28	-0.26	1.51	-0.9	-0.35	2.33	2.33	1.37	-0.54	-0.4	0.23	-0.26	-0.04	0.83	-0.8	0.87	0.34	0.74	-0.71	0.52
3_{10} -helix (C)	ECn3Ce5	0.5	0.75	-1.02	-1.13	-0.81	1.64	1.64	1.67	-2.36	-1.25	-0.9	0.44	0.34	-0.4	-1.52	-2.12	0.61	-0.8	-0.61	0.1
3_{10} -helix (C)	ECn3Ce6	0.56	-1.05	0.61	0.09	-1.01	0.38	0.38	0.94	0.5	0.75	-1.02	-0.18	0.12	-1.25	-0.81	0.56	0.89	-1.52	0.16	0.29
3_{10} -helix (C)	ECn3Ce7	0.74	-0.28	0.89	-0.33	-0.71	0.41	0.04	0.32	0.56	-1.05	0.61	1.06	0.56	0.84	0.09	0.35	1.26	-0.8	0.65	-0.12
3_{10} -helix (C)	ECn4Ce1	-0.95	-0.24	1.26	0.09	-0.61	0.74	-0.54	0.01	0.74	-0.28	0.89	0.11	-0.04	0.71	-0.61	1.32	1.82	1.39	0.34	-0.03
3_{10} -helix (C)	ECn4Ce2	-0.03	0.41	0.04	0.32	0.16	-0.8	-2.36	-0.79	-0.95	-0.24	1.26	0.93	-0.35	0.16	0.84	-0.75	1.6	-1.3	0.12	0.69
3_{10} -helix (C)	ECn4Ce3	-0.5	0.74	-0.54	0.01	0.65	-1.52	-2.22	-0.87	0.08	-0.54	-0.76	0.65	-0.83	-1.48	-0.5	-1.91	1.32	-1.61	0.56	-0.19
3_{10} -helix (C)	ECn4Ce4	-0.58	-0.8	-2.36	-0.79	0.34	-0.8	-2.36	0.16	0.35	-0.49	0.14	0.47	0.44	1.18	0.6	-1.35	-1.02	-1.35	0.87	0.34
3_{10} -helix (C)	ECn4Ce5	0.09	-1.52	-2.22	-0.87	0.41	1.73	-1.01	-0.41	-0.08	0.13	0.55	-1.35	-0.3	-1.78	1.43	-0.67	0.32	-0.77	-0.18	0.27
3_{10} -helix (C)	ECn4Ce6	-0.03	0.32	0.6	-0.33	-0.71	0.84	-0.71	-0.44	-0.44	-1.25	-0.41	0.34	-0.2	0.07	-1.29	0.07	1.81	1.81	1.06	0.87
3_{10} -helix (C)	ECn4Ce7	-0.5	0.01	-0.37	0.09	-0.61	-0.41	0.34	-0.82	-0.8	-2.36	-0.58	-0.8	-1.59	-1.54	-0.9	-0.35	2.33	2.33	0.11	0.54
3_{10} -helix (C)	EG	-0.61	1.45	-1.01	0.09	-0.58	-1.59	0.41	1.21	1.64	-1.01	-0.41	1.78	-1.01	-0.1	0.41	0.32	1.64	1.64	0.93	0.14
3_{10} -helix (C)	FCn1Ce1	0.16	0.84	-0.71	0.52	-0.2	0.56	0.83	0.84	0.83	-0.8	-2.36	0.16	0.71	0.52	-1.33	0.16	0.84	-0.71	0.65	0.23
3_{10} -helix (C)	FCn1Ce2	0.65	0.71	-0.61	0.01	-0.71	0.74	-0.54	0.71	-0.4	-1.52	-0.2	-0.41	-2.01	-0.8	-0.82	-0.79	0.71	0.61	-0.48	0.34
3_{10} -helix (C)	FCn1Ce3	0.34	-0.2	0.16	-0.79	-0.81	-0.8	-2.36	-0.2	-1.25	-0.81	-1.12	-0.44	0.16	-0.33	-1.23	-0.87	-0.2	0.16	-1.3	-0.48
3_{10} -helix (C)	FCn1Ce4	0.41	-1.12	0.65	-0.87	0.87	0.34	0.08	-0.11	0.5	0.75	-0.82	-2.01	-1.13	-1.01	-1.01	0.16	-1.12	0.65	-1.61	0.62
3_{10} -helix (C)	FCn1Ce5	0.83	-0.8	-2.36	0.16	-2.12	0.61	-1.05	-0.47	0.56	-1.05	1.54	0.16	0.84	0.16	-0.71	0.52	0.8	-2.36	-1.35	0.13
3_{10} -helix (C)	FCn1Ce6	-0.4	-1.52	-0.2	-0.41	0.56	0.89	-0.28	0.41	0.74	-0.28	-0.71	-1.01	-0.41	-0.4	-0.81	-0.2	-1.52	-0.2	-0.77	-0.32
3_{10} -helix (C)	FCn1Ce7	0.16	0.29	0.41	0.16	0.35	1.26	-0.24	-1.06	-0.95	-0.24	-0.61	-0.71	-0.44	-1.23	-0.44	-1.13	-0.81	-1.12	-1.88	0.55
3_{10} -helix (C)	FCn2Ce1	-0.11	-1.33	-0.48	-0.41	0.08	-0.76	-0.54	0.63	0.08	-0.54	0.16	-0.61	1.15	-1.01	-0.82	0.09	-1.01	-0.41	-1.84	1.07
3_{10} -helix (C)	FCn2Ce2	-0.41	-0.82	-0.07	-0.44	0.07	0.14	-0.49	-0.25	0.35	-0.49	0.65	-1.52	-0.2	-0.41	0.6	-0.33	-0.71	-0.44	-0.57	0.81
3_{10} -helix (C)	FCn2Ce3	-0.44	-1.23	-0.48	1.96	-1.54	-0.23	0.23	-0.57	-0.08	0.13	0.65	-0.87	-1.38	-2.01	-0.37	0.09	-0.61	1.45	0.74	-0.28
3_{10} -helix (C)	FCn2Ce4	1.65	-1.01	-1.59	0.84	-0.1	-0.15	0.59	-1.17	-0.99	-0.32	-2.36	0.16	0.16	0.16	-0.33	-0.58	-0.8	-0.87	0.58	0.5
3_{10} -helix (C)	FCn2Ce5	0.16	0.01	-1.25	0.84	0.48	-0.11	-1.3	1.32	0.89	4.32	-0.2	-0.41	-0.82	-0.79	0.09	-2.04	-1.52	0.84	-0.47	0.56
3_{10} -helix (C)	FCn2Ce6	-0.41	-0.79	-1.66	0.71	1.02	-0.65	-2.04	-0.97	0.15	-0.67	-1.12	-0.44	-1.23	-0.87	-0.82	0.83	-0.8	-2.63	0.41	0.74
3_{10} -helix (C)	FCn2Ce7	-0.8	-0.87	-1.38	-0.2	0.49	-1.52	-2.22	-0.86	-0.72	-1.14	-0.41	-0.82	0.09	0.09	-1.23	-1.52	1.1	-0.33	-1.06	-0.95
3_{10} -helix (C)	FCn3Ce1	-1.52	0.84	-0.71	0.74	-0.15	-0.81	-1.71	-0.07	-0.62	-1.03	-0.44	-1.23	-0.33	-0.58	0.09	-0.81	0.84	0.09	-0.87	-0.56
3_{10} -helix (C)	FCn3Ce2	-0.8	-2.63	-0.81	-0.8	0.87	-0.44	-1.09	-1.11	-1.38	-0.69	1.45	-1.01	0.09	-2.04	-0.33	-0.44	0.71	-0.61	0.16	-0.07
3_{10} -helix (C)	FCn3Ce3	1.85	-0.33	-0.44	-1.52	-0.68	-1.08	-2.29	-0.01	-0.79	-0.1	0.84	-0.71	0.32	0.83	0.09	-0.82	-0.2	0.16	0.52	-0.33
3_{10} -helix (C)	FCn3Ce4	0.84	0.09	-0.82	-0.8	0.23	-0.87	-0.61	-0.11	-0.98	-1.62	0.95	0.91	-0.74	-0.46	0.32	-0.07	-1.12	0.65	0.1	0.09
3_{10} -helix (C)	FCn3Ce5	0.71	-0.61	1.84	1.39	-0.9	0.44	0.63	0.05	-0.17	0.15	0.12	-1.59	0.7	1.18	0.01	-1.25	-0.41	0.34	0.29	0.32
3_{10} -helix (C)	FCn3Ce6	0.09	-0.82	-0.8	-2.36	-2.04	0.14	1.75	-0.88	-0.04	0.48	-0.11	-1.3	1.32	0.89	-0.79	-0.58	-1.59	0.41	-0.12	0.01
3_{10} -helix (C)	FCn3Ce7	-0.61	1.84	1.39	-1.01	0.34	0.16	1.1	0.25	-0.35	-1.13	0.6	-1.01	0.11	0.12	-0.87	-0.2	0.56	0.83	-0.03	0.41
3_{10} -helix (C)	FCn4Ce1	0.16	0.84	0.83	-0.8	0.41	-1.24	0.49	-0.85	-1.24	-1.64	0.57	-1.64	0.35	0.75	0.84	-0.71	0.74	-0.54	-1.52	0.84
3_{10} -helix (C)	FCn4Ce2	0.65	0.71	-0.4	-1.52	-0.48	-0.26	0.74	-0.28	-0.26	-0.97	1.02	-1.26	-0.45	-0.44	-1.23	-0.87	-0.2	0.34	-0.8	-2.63
3_{10} -helix (C)	FCn4Ce3	0.34	-0.2	-1.25	-0.81	-0.07	0.57	0.58	0.5	0.75	-0.39	-0.5	-1.13	-0.25	-0.82	0.09	0.09	0.74	-0.54	1.1	-0.33
3_{10} -helix (C)	FCn4Ce4	-0.8	-1.59	-0.44	-1.23	-0.48	-1.05	-0.47	0.56	-1.05	0.32	0.64	-0.3	1.06	-1.23	-0.33	-0.58	-0.8	-2.36	0.84	0.09
3_{10} -helix (C)	FCn4Ce5	1.78	-1.01	1.94	-1.01	-1.59	-0.28	0.41	0.74	-0.28	-1.26	-0.47	-1.88	0.29	-1.01	0.09	-2.04	-1.52	-0.8	0.71	-0.61
3_{10} -helix (C)	FCn4Ce6	0.16	-0.71	0.52	-0.33	-0.44	-0.24	-1.06	-0.95	-0.24	-1.16	-0.32	-1.84	-0.83	-0.71	0.32	0.83	-0.8	-1.52	-0.2	0.16
3_{10} -helix (C)	FCn4Ce7	-0.41	-2.01	-0.8	-1.59	-1.8	-0.54	0.63	0.08	-0.54	-1.2	-0.42	-0.57	-0.44	-0.61	0.01	-1.52	1.15	-0.33	-2.63	0.23
3_{10} -helix (C)	FG	-0.44	0.16	-0.33	-0.87	-0.2	-0.49	-0.25	0.35	-											

Table5 The environment-specific substitution tables (random coil).The environment scores for each amino acid. Large negative values indicate a strong preference for the particular environment whereas large positive values indicate an aversion.

		W	F	Y	L	I	V	M	A	G	P	C	T	S	Q	N	E	D	H	K	R
random coil (D)	ECn1Ce1	-0.94	0.69	1.18	0.6	-1.35	-1.02	0.84	-0.71	0.32	0.83	-0.8	0.16	-0.41	-0.82	1.1	-1.01	0.09	-0.82	-0.2	0.16
random coil (D)	ECn1Ce2	0.56	-0.19	-1.78	1.43	-0.67	0.32	0.71	-0.61	0.01	-1.52	1.15	0.49	-0.44	-1.23	0.84	-0.71	0.52	-0.07	-1.12	0.65
random coil (D)	ECn1Ce3	0.59	-0.79	-0.1	-0.71	0.52	-0.47	-0.2	0.16	-0.79	0.56	0.83	0.56	1.65	-1.01	0.71	-0.61	0.01	-1.25	-0.41	0.34
random coil (D)	ECn1Ce4	0.66	-0.98	-0.48	-0.61	0.1	-0.32	-1.12	0.65	0.87	0.32	-0.4	-2.36	0.16	0.01	-0.2	0.16	0.79	0.58	-1.59	0.41
random coil (D)	ECn1Ce5	-1.75	-0.17	-0.2	0.16	0.29	-0.42	-0.8	-2.36	0.16	0.01	-1.25	-0.2	-0.41	-0.79	-1.12	0.65	-0.87	-0.2	0.56	0.83
random coil (D)	ECn1Ce6	0.33	-0.04	-0.17	0.65	-0.12	-0.28	-1.52	-0.2	-0.41	-0.79	-1.66	-0.58	-0.8	-0.87	-2.04	-1.52	0.84	-0.71	0.74	-0.54
random coil (D)	ECn1Ce7	0.82	-0.35	0.08	0.34	-0.03	0.23	-0.33	-0.58	-0.8	-0.87	-1.38	-2.04	-1.52	0.84	-1.35	0.13	-0.33	-0.58	-0.8	-2.36
random coil (D)	ECn2Ce1	-0.4	-1.24	3.63	0.12	0.69	-1.63	0.09	-2.04	4.32	-0.83	0.34	-0.03	0.41	-2.63	-0.77	-0.32	0.09	-2.04	-0.12	1.32
random coil (D)	ECn2Ce2	0.75	-0.26	1.51	0.56	-0.19	0.13	0.12	-1.59	1.37	-0.5	0.12	0.69	1.85	-0.33	1.81	4.32	0.32	0.83	0.65	0.55
random coil (D)	ECn2Ce3	-1.99	-0.64	1.06	0.87	0.34	0.08	-0.11	-1.3	1.67	-1.22	0.56	-0.19	0.84	0.09	2.33	1.37	0.01	-1.52	0.5	-0.15
random coil (D)	ECn2Ce4	-2.02	0.12	0.69	-1.55	-0.1	-1.63	0.6	-1.01	0.94	0.39	-0.18	0.27	0.71	-0.61	1.64	1.67	-2.04	-1.52	0.56	0.14
random coil (D)	ECn2Ce5	0.6	0.56	-0.19	0.17	0.65	0.13	0.57	-1.64	1.32	-0.64	1.06	0.87	-0.2	0.16	0.38	0.94	0.83	-0.8	0.93	1.32
random coil (D)	ECn2Ce6	-1.54	-0.18	0.27	-0.23	0.11	-1.71	1.02	-1.26	0.55	0.71	0.11	0.54	-1.12	0.65	0.04	0.32	-1.52	1.1	0.78	-0.75
random coil (D)	ECn2Ce7	-2.12	1.06	0.87	0.34	0.08	-0.11	-0.5	-1.13	-0.15	0.59	0.93	0.14	-0.41	0.34	-0.54	0.01	-0.11	-1.33	-0.5	-1.91
random coil (D)	ECn3Ce1	-1.74	0.11	0.54	-1.62	0.95	0.91	0.64	-0.3	1.14	2.38	0.65	0.23	-1.02	-0.18	-2.36	-0.79	-0.41	-0.82	0.6	-1.35
random coil (D)	ECn3Ce2	-0.89	0.93	0.14	0.15	0.12	-1.59	-0.47	-1.88	0.04	0.32	0.56	-1.05	0.61	1.06	0.56	0.16	-0.44	-1.23	1.43	-0.67
random coil (D)	ECn3Ce3	0.12	0.65	0.23	0.48	-0.11	-1.3	-0.32	-1.84	-0.54	0.01	0.74	-0.28	0.89	0.11	-0.04	-0.41	0.09	-2.12	-0.71	0.52
random coil (D)	ECn3Ce4	0.98	-0.48	-0.34	1.6	-1.35	-0.18	-0.42	-0.57	-2.36	-0.79	-1.95	-0.24	1.26	0.93	-0.35	-0.16	-0.82	0.83	-0.61	0.1
random coil (D)	ECn3Ce5	0.69	0.65	0.34	-0.03	-1.48	-0.5	-0.28	0.23	-2.22	-0.87	0.08	-0.54	0.76	0.65	0.83	-1.48	-1.23	-1.52	0.16	0.29
random coil (D)	ECn3Ce6	0.9	-0.11	0.12	0.69	1.18	0.6	-0.5	-0.27	-2.36	0.16	0.35	1.64	1.64	1.67	-1.43	0.56	0.09	-0.81	0.65	-0.12
random coil (D)	ECn3Ce7	-0.83	0.05	0.56	-0.19	-1.78	1.43	0.42	0.06	0.12	-0.12	1.32	0.38	0.38	0.94	0.39	-0.18	-1.01	0.09	0.34	-0.03
random coil (D)	ECn4Ce1	1.33	-0.88	-0.04	-0.79	-0.1	-0.71	-0.26	-0.36	0.65	0.65	0.55	-0.12	-0.12	1.32	-0.64	1.06	-0.71	0.32	0.12	0.69
random coil (D)	ECn4Ce2	-0.22	0.25	-0.35	-0.98	-0.48	-0.61	-0.67	-0.38	0.27	0.5	-0.15	0.65	0.65	0.55	0.71	0.11	-0.61	0.01	0.84	0.83
random coil (D)	ECn4Ce3	-0.63	-0.85	-1.24	-0.17	1.34	0.16	-1.04	-0.46	0.32	0.56	0.14	0.27	0.5	-0.15	0.59	0.93	0.16	-0.79	0.71	-0.4
random coil (D)	ECn4Ce4	0.5	-0.28	-0.26	-0.04	-0.17	0.65	-1.35	-0.87	-0.42	0.93	1.32	0.32	0.56	0.14	2.38	0.65	0.65	-0.87	-0.2	-1.25
random coil (D)	ECn4Ce5	-0.41	0.5	0.75	-0.35	0.08	0.34	-1.97	-0.57	0.25	0.78	0.75	-0.28	0.26	1.51	0.56	0.19	0.13	0.56	-0.11	0.5
random coil (D)	ECn4Ce6	-0.21	0.56	-1.05	-1.24	3.63	0.12	0.34	-0.03	-1.48	-1.78	1.43	-0.24	-0.64	1.06	0.87	0.34	0.08	-0.11	-0.47	0.56
random coil (D)	ECn4Ce7	0.23	-0.48	-0.34	1.6	-1.35	-0.18	1.07	-0.67	0.32	-0.1	-0.71	-1.13	-1.7	0.54	-2.12	0.61	-1.05	-0.47	0.41	0.74
random coil (D)	EG	1.18	-1.3	-0.48	1.32	-0.67	-0.38	0.81	0.52	-0.47	-0.48	-0.61	0.1	-0.32	-1.84	1.07	0.71	-0.44	-1.25	-1.06	-0.95
random coil (D)	FCn1Ce1	0.2	-1.61	0.62	-1.02	-1.04	0.46	0.02	0.1	0.32	0.2	0.16	0.29	-0.42	0.57	0.81	0.66	-0.8	-2.36	0.63	0.88
random coil (D)	FCn1Ce2	0.54	-1.35	0.13	0.55	-1.35	-0.87	0.26	0.29	-0.42	-0.17	0.65	-0.12	-0.28	0.23	0.02	-0.29	1.64	-1.01	-0.25	0.35
random coil (D)	FCn1Ce3	-0.04	-0.77	-0.32	0.74	-1.97	-0.57	-0.54	-0.12	0.28	0.08	0.34	0.03	0.23	0.41	0.26	0.16	0.83	-0.8	-0.57	-0.08
random coil (D)	FCn1Ce4	1.99	0.38	0.94	0.39	0.61	-1.05	-0.47	-0.03	0.23	3.63	0.12	0.69	-1.63	0.12	-0.54	-0.4	-0.4	-1.52	-1.17	-0.99
random coil (D)	FCn1Ce5	0.24	-0.12	1.32	-0.64	0.89	-0.28	0.41	0.69	-1.63	1.51	0.56	-0.19	0.13	0.56	-2.36	-1.25	-1.25	-0.81	1.32	0.89
random coil (D)	FCn1Ce6	0.83	0.65	0.55	0.71	1.26	-0.24	-1.06	-0.19	0.13	1.06	0.87	0.34	0.08	0.11	0.5	0.75	0.84	0.09	-0.97	0.15
random coil (D)	FCn1Ce7	0.47	0.09	0.48	0.83	-0.76	-0.54	0.63	0.34	0.08	0.54	-2.12	0.61	-1.05	-0.47	0.56	-1.05	0.71	-0.61	-0.86	-0.72
random coil (D)	FCn2Ce1	0.12	0.73	1.22	0.16	0.14	-0.49	-0.25	0.35	0.16	0.84	-0.71	-0.11	0.65	-0.25	0.46	-1.04	-0.46	0.32	-0.07	-0.62
random coil (D)	FCn2Ce2	-0.25	-0.06	0.62	-1.88	-0.23	0.23	-0.57	-0.08	0.65	0.71	-0.61	0.91	-0.74	0.46	1.1	-1.35	-0.87	0.27	-1.11	-1.38
random coil (D)	FCn2Ce3	0.75	-0.13	0.27	0.5	-0.15	0.59	-1.17	-0.99	0.34	-0.2	0.16	-1.59	0.7	1.18	0.47	-1.97	-0.57	0.06	-0.01	-0.79
random coil (D)	FCn2Ce4	0.65	0.82	0.32	0.56	0.14	2.38	-0.1	-0.32	0.41	-1.12	0.65	-1.3	1.32	0.89	0.63	1.02	-0.65	-0.36	-1.52	-2.22
random coil (D)	FCn2Ce5	1.75	1.24	-0.42	0.93	1.32	1.82	-0.48	-0.34	0.83	-0.8	-2.36	-0.18	1.07	0.71	0.49	0.49	-1.52	-0.38	-0.8	-2.36
random coil (D)	FCn2Ce6	0.05	0.85	0.25	0.78	-0.75	1.6	-1.3	-0.48	-0.4	-1.52	-0.2	-0.38	0.81	0.66	-0.15	-0.15	-0.04	-0.46	1.73	-1.01
random coil (D)	FCn2Ce7	-0.65	-2.18	-1.59	-1.8	-1.91	1.32	-1.61	0.62	-1.23	0.33	0.58	0.46	0.02	0.29	0.87	0.87	-0.35	0.87	0.84	0.71
random coil (D)	FCn3Ce1	-0.03	-2.61	-2.01	-2.63	-0.44	0.71	-0.61	1.86	-1.01	0.09	-2.04	-0.87	0.26	0.16	-0.68	-0.68	-1.24	-0.57	0.71	-0.61
random coil (D)	FCn3Ce2	0.74	-0.71	0.52	-0.33	-0.87	-1.38	-0.88	-0.04	-0.17	-0.64	-0.44	-0.57	-0.54	-0.4	0.23	0.23	-0.26	-0.03	-0.2	-1.52
random coil (D)	FCn3Ce3	1.32	-2.01	-1.13	-1.01	0.84	-0.71	0.25	-0.35	0.08	-2.36	-0.82	0.09	-1.01	0.41	-0.82	0.9	0.44	0.69	-1.13	-0.81
random coil (D)	FCn3Ce4	0.11	0.16	0.84	0.16	-2.63	-0.81	-0.85	-1.24	3.63	-0.87	0.6	-0.33	-0.71	-0.44	-1.23	-1.02	-0.18	-0.19	0.09	-1.01
random coil (D)	FCn3Ce5	0.35	-0.79	0.71	0.65	-0.33	-0.44	1.82	-0.48	-0.34	-0.58	-0.37	0.09	0.61	1.1	-1.01	0.61	1.06	-0.79	-0.33	-0.71
random coil (D)	FCn3Ce6	-0.45	-0.87	-0.2	0.34	-0.12	-0.98	-0.48	-0.61	0.1	0.41	0.04	0.32	0.16	0.84	-0.71	0.89	0.11	-0.98	0.09	-0.61
random coil (D)	FCn3Ce7	-0.25	0.09	0.74	-0.54	-0.03	-0.17	-0.2	0.16	0.29	0.74	-0.54	0.01	0.65	0.71	-0.61	1.26	0.93	-0.17	0.32	0.16
random coil (D)	FCn4Ce1	1.06	-0.58	-0.8	-2.36	0.69	-0.04	-0.17	0.65	-0.12	-0.8	-2.36	-0.79	0.34	-0.2	0.16	0.76	0.65	-2.63	0.01	0.65
random coil (D)	FCn4Ce2	0.29	-2.04	-1.52	-0.8	-0.19	-0.35	0.08	0.34	-0.98	-1.52	-2.22	-0.87	0.41	-1.12	0.65	0.14	0.47	0.74	-0.79	0.34
random coil (D)	FCn4Ce3	-0.41	0.83	-0.8	-1.52	0.27	-1.52	0.84	-0.71	0.16	0.32	0.6	-0.33	-0.71	-2.04	-1.52	0.55	-1.35	0.58	-0.87	0.41
random coil (D)	FCn4Ce4	0.45	-1.52	1.15	-0.33	0.87	-0.8	-2.63	-0.81	-0.41	0.01	-0.37	0.09	-0.61	0.83	-0.8	0.74	-1.97	-0.47	-0.33	-0.71
random coil (D)	FCn4Ce5	0.3	0.56	0.83	0.09	0.54	1.1	-0.33	-0.44	-0.44	0.71	0.65	0.56	-0.11	-0.87	0.41	-0.83	0.34	-0.82	0.83	-0.8
random coil (D)	FCn4Ce6	-0.28	0.32	-0.4	0.32	0.84	-0.71	0.74	-0.54	0.32	0.16	0.84	-0.44	-1.13	-0.81	-1.12	-1.88	-0.37	0.09	-0.61	1.45
random coil (D)	FCn4Ce7	-0.07	0.01	-1.25	0.01	-2.63	-0.81	-0.8	-2.36	0.01	0.65	0.71	-0.82	0.09	-1.01	-0.41	-1.84	-0.33	-0.58	-0.8	-0.87
random coil (D)	FG	-0.18	-0.79	-1.66	-0.79	-0.33	-0.44	-1.52	-2.22	-0.79	0.34	-0.2	0.6	-0.33	-0.71	-0.44	-0.57	0.09	-2.04	-1.52	0.84

Table 6 Evaluation of the method and other published potentials on decoy sets from ProStar

109 structure-decoy pairs in three different set

Decoy set	Size	TUNE	RKBP	KBP	CDF	RDHSP
ifu	24	31	22	32	21	23
misfold	41	24	24	24	19	39
asilomar	44	31	35	39	35	42
Total	109	86	81	93	75	94
		79.0%	74.3%	85.3%	68.8%	86.24%

TUNE : Threading using neural network, Kuang Lin (2002)

RKBP : The residue contact potential from Lu and Skolnick (2001)

KBP : The atomic potential from Lu and Skolnick (2001)

CDF : The residue-based potential from Samudrala and Moulton (1998)



Table 7 Performance of different methods for fold recognition in Lindahl and Elofsson recognition benchmark

Method	Family only		Superfamily only		Fold only	
	TOP1	TOP5	TOP1	TOP5	TOP1	TOP5
THREADER	49.2%	58.9%	10.8%	24.7%	14.6%	37.7%
HMMER-PSIBLAST	67.7%	73.5%	20.7%	31.3%	4.4%	14.6%
SAMT98-PSIBLAST	70.1%	75.4%	28.3%	38.9%	3.4%	18.7%
BLASTLINK	74.6%	78.9%	29.3%	40.6%	6.9%	16.5%
SSEARCH	68.6%	75.5%	20.7%	32.5%	5.6%	15.6%
PSI-BLAST	71.2%	72.3%	27.4%	27.9%	4.0%	4.7%
FUGUE	82.2%	85.8%	41.9%	53.2%	12.5%	26.8%
RDHSP	75.8%	80.6%	45.6%	62.5%	22.6%	39.7%

Table 8 Performance of different term on decoy sets from Prostar

109 structure-decoy pairs in three different set

Decoy set	Size	S	S+ASA	S+ASA+Ce	S+ASA+Cn	S+ASA+Ce+Cn
<i>misfold</i>	24	9	11	15	13	22
<i>asilomar</i>	41	19	20	28	22	34
<i>ifu</i>	44	15	24	30	27	32
Total	109	43	55	73	62	88
			50.5%	67.0%	56.9%	81.0%

S : Secondary structure ASA : Residue Accessibility
Cn : Contact residue numbers Ce : Contact energy

Reference

1. Overington, J., Johnson, M. S., Sali, A. & Blundell, T. L. (1990). Tertiary structural constraints on protein evolutionary diversity: templates, key residues and structure prediction. *Proc. Roy. Soc. London s-er. B*, 241, 132-145.
2. Gonnet, G. H., Cohen, M. A. & Benner, S. A. (1992). Exhaustive matching of the entire protein sequence database. *Science*, 256, 1443-1445.
3. Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. (1978). A model for evolutionary change. In *Atlas of Protein Sequence and Structure* (Dayhoff, M. O., ed), vol. 5, pp. 345- 358, National Biomedical Research Foundation, Washington, DC.
4. Jones, D. T., Taylor, W. R. & Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275-282.
5. Henikoff, S. & HENIKOFF, J. G. (1992). Amino acid substitution matrices from protein blocks. *Proc. Natl Acad. Sci. USA*, 89, 10915-10919.
7. Mehta, P. K., Heringa, J. & Argos, P. (1995). A simple and fast approach to prediction of protein secondary structure from multiply aligned sequences with accuracy above 70 %. *Protein Sci.* 4, 2517-2525.
8. Berman, H. m., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. & Bourne, P.E. (2000) *Nucleic Acids Res.* 28, 235-242.
9. Hobohm, U. & Sander, C. (1994) *Protein Sci.* 3, 522-524
10. Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995) *J.Mol. Biol.* 247, 536-540.
11. Fischer D, Eisenberg D. Protein fold recognition using sequence-derived predictions, *Protein Sci* 1996;5:937-955.

12. Shi et al. (2001) FUGUE: Sequence-Structure Homology recognition using environment-specific substitution tables and structure-dependent gap Penalties. *J Mol Biol.* 310:243-257.
13. Iksoo et al. (2001) Protein threading by learning. *Proteins.* 98:14350-14355.
14. Williams et al. (2001) Sequence – Structure Homology Recognition by Iterative Alignment and Comparative Modeling. *Proteins* 5:92-97.
15. Overington, et al. (1992) Environment-specific amino acid substitution tables: Tertiary templates and prediction of protein folds. *Prot Sci* 1:216-226
16. Kenji, et al. (1998) JOY : protein sequence-structure representation and analysis. *Bioinformatics* 14:617-623.
17. Daniel J et al. (1999) Protein Fold Recognition Score Functions: Unusual Construction Strategies. *Proteins* 36:454-461.
18. Burkhard Rost et al. (1997) Protein Fold Recognition by Prediction-based Threading. *JMB* 270:471-480.
- Bowie et al. (1991) A method to identify protein sequences that fold into a known three-dimensional structure. *Science*, 253:164-169.
20. Proteins: Structure, Function, and Genetics , Volume 40, Issue 3 , Pages 343 - 354
21. D. Kim, D. Xu, J. Guo, K. Ellrott, Y. Xu, Prospect II: Protein Structure Prediction Program for the Genome-Scale Application, Protein Engineering, , vol. 16 no. 9 pp. 641-650, 2003.