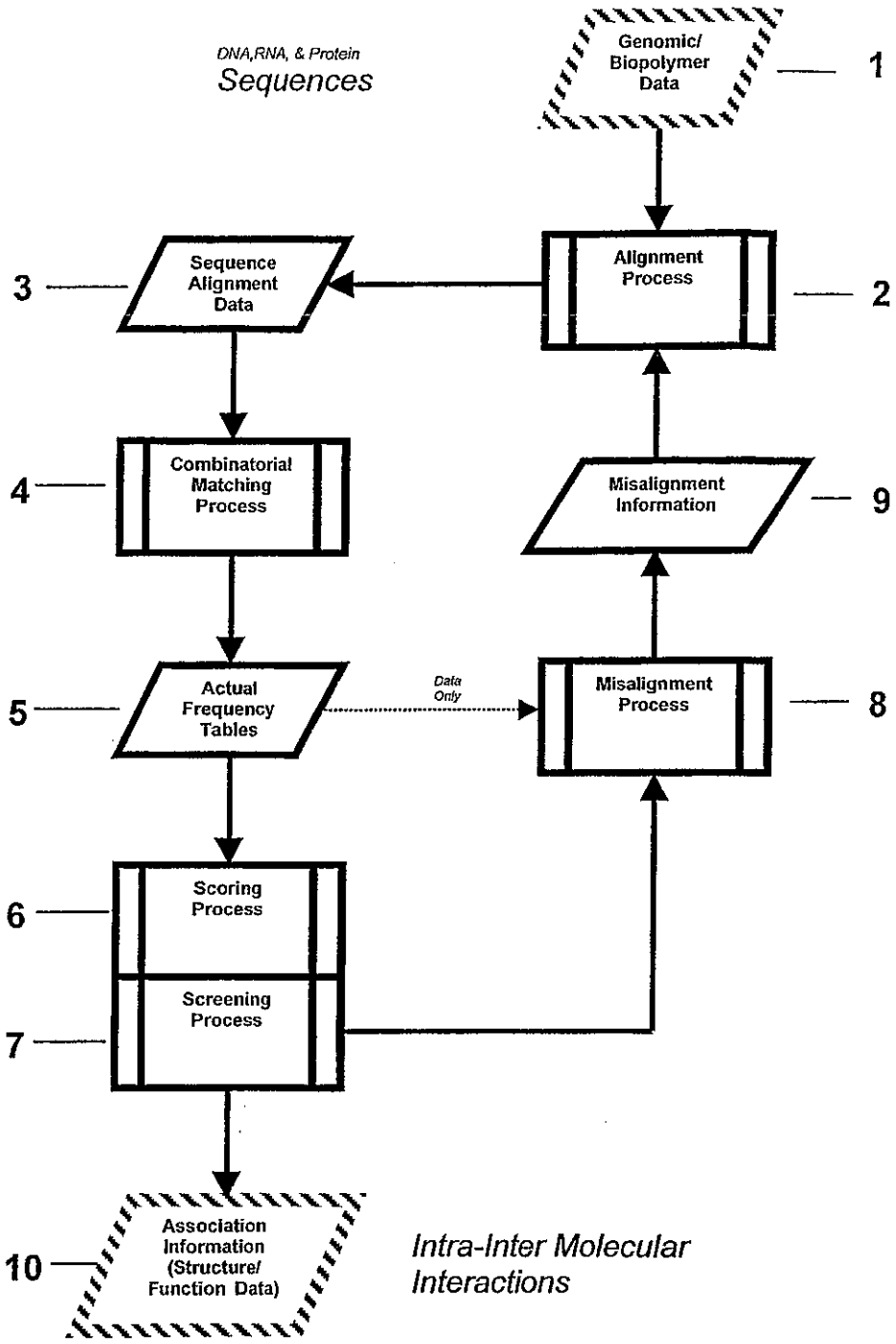

DNA,RNA, & Protein
Sequences
Genomic/
Biopolymer
Data
1
Alignment
Process
2
Sequence
Alignment
Data
3
Combinatorial
Matching
Process
4
Misalignment
Information
9
Actual
Frequency
Tables
5
Data
Only
Misalignment
Process
8
Scoring
Process
6
Screening
Process
7
Association
Information
(Structure/
Function Data)
10
Intra-Inter Molecular
Interactions

Fig. 1

	Position A	Position B	Position C	Position D	Position E	Position F	Position G	Position H	Position I	Position J	Position K	Position L	Position M	Position N	Position O	Position P	Position Q	Position R	Position S	Position T	Position U	Position V	Position W	Position X	Position Y	Position Z	Position Aa	Position Ab
Sequence 1	q	δ	g	π	ρ	ρ	ρ	g	π	g	g	g	δ	δ	-	π	ρ	ρ	g	π	δ	g	ρ	δ	π	δ	g	ρ
Sequence 2	q	δ	g	π	ρ	ρ	ρ	g	π	g	g	g	δ	δ	-	π	ρ	ρ	g	π	δ	g	ρ	δ	π	δ	g	ρ
Sequence 3	ρ	δ	g	π	π	g	ρ	g	x	g	δ	g	δ	δ	-	π	π	ρ	w	π	g	π	ρ	δ	π	g	g	ρ
Sequence 4	ρ	δ	g	π	π	g	ρ	g	δ	ρ	g	g	δ	δ	-	π	π	ρ	ρ	π	π	g	π	δ	π	g	g	ρ
Sequence 5	ρ	δ	g	π	π	g	ρ	g	δ	ρ	g	g	δ	δ	-	π	π	ρ	ρ	π	π	g	π	δ	π	g	g	ρ
Sequence 6	ρ	δ	g	π	π	g	ρ	g	π	g	δ	g	δ	δ	-	π	ρ	ρ	g	π	δ	g	ρ	g	π	g	g	ρ
Sequence 7	ρ	δ	g	π	π	g	ρ	g	s	g	δ	g	δ	δ	-	π	ρ	ρ	π	π	δ	g	ρ	g	π	g	g	ρ
Sequence 8	ρ	δ	g	π	π	g	ρ	g	π	g	δ	g	δ	δ	x	π	ρ	ρ	g	π	δ	g	ρ	g	π	g	g	ρ

Fig. 2

IxS

	g	w	ρ	π
π	4	0	0	0
x	0	1	0	0
δ	0	0	2	0
s	0	0	0	1

301

Fig. 3(a)

IxS

	g	w	ρ	π
π	2	0.5	1	0.5
x	0.5	0.125	.25	0.125
δ	1	.25	.5	.25
s	.5	0.125	.25	0.125

302

Fig. 3(b)

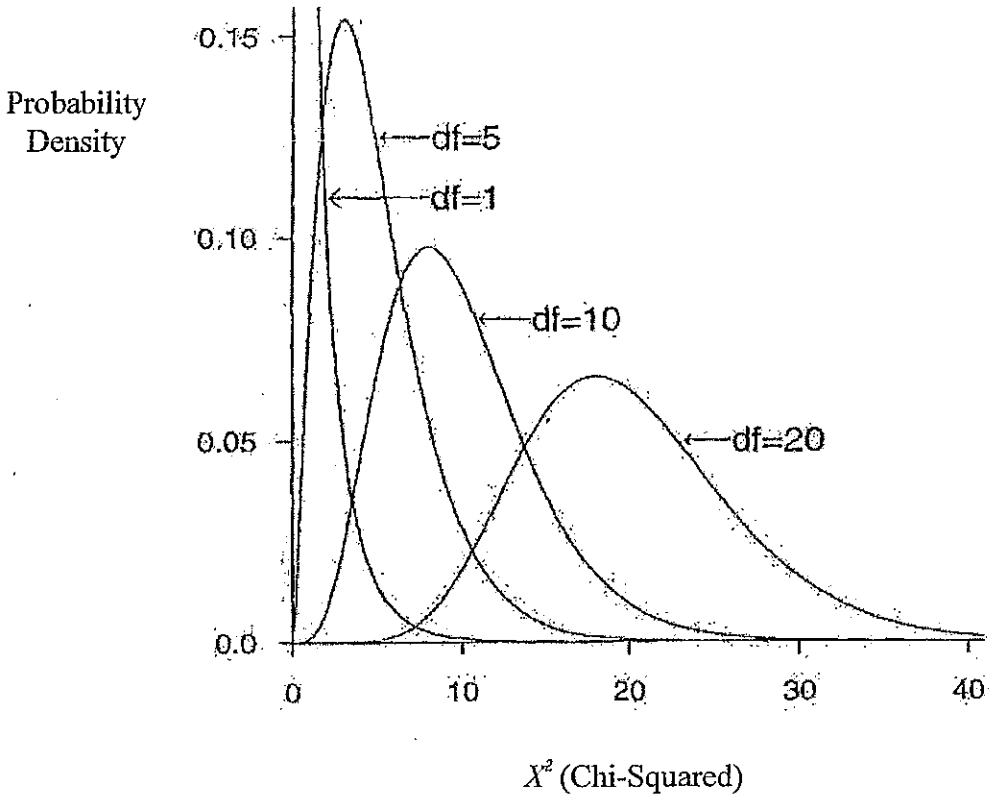

Probability Density
0.15
0.10
0.05
0.0
df=5
df=1
df=10
df=20
0
10
20
30
40
X^2 (Chi-Squared)

Fig. 4

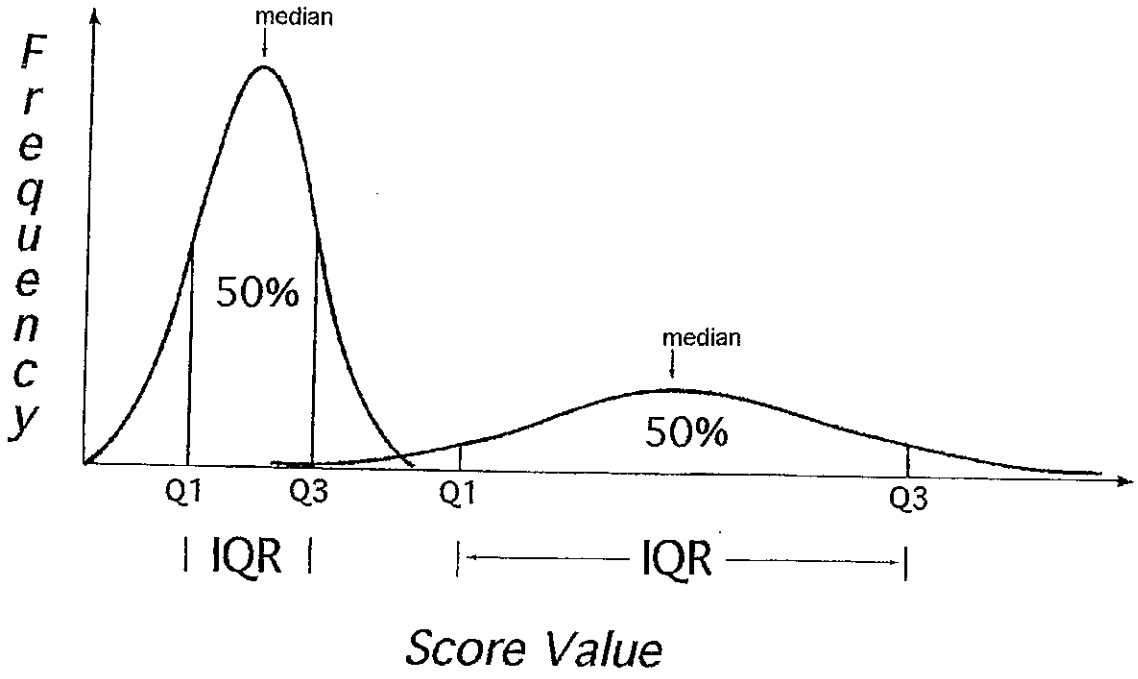


Fig. 5

Positions
2 x 10

	C	T	G
G	83	24	0
A	0	18	***0***
C	0	***0***	2

61

Fig. 6(a)

Positions
2 x 10

	C	T	G
G	6.69172	-5.89562	-3.29715
A	-6.28981	6.51493	***-0.57928***
C	-1.95783	***-1.00202***	11.26943

62

Fig. 6(b)

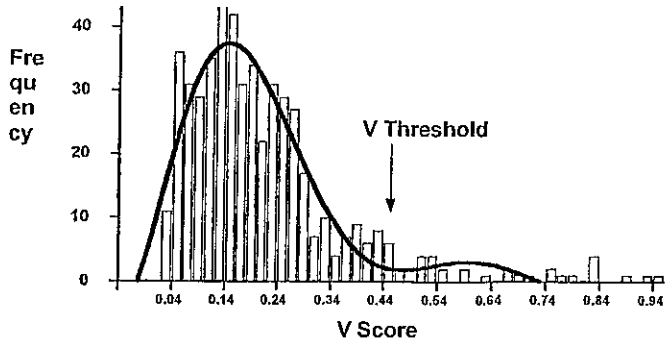


Fig. 7(a)

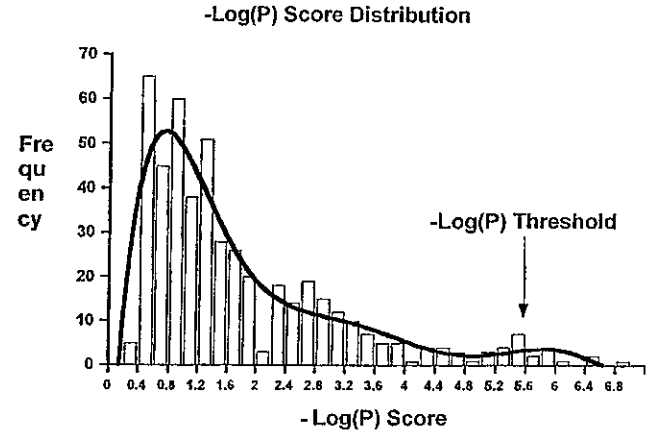


Fig. 7(b)

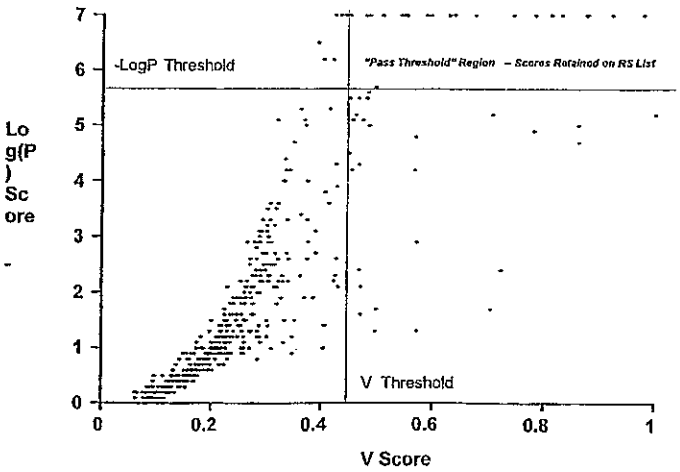


Fig. 7(c)

SHEVEK Predicted Structure

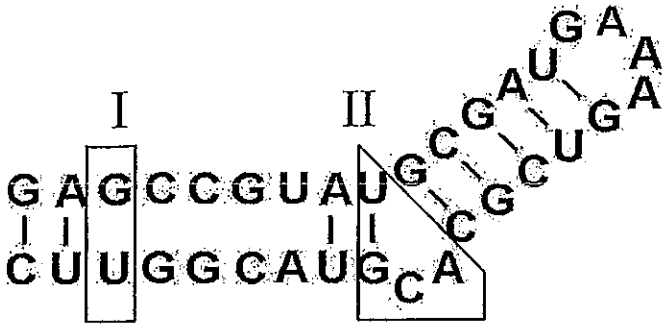


Fig. 8(a)

Biochemically Predicted Structure

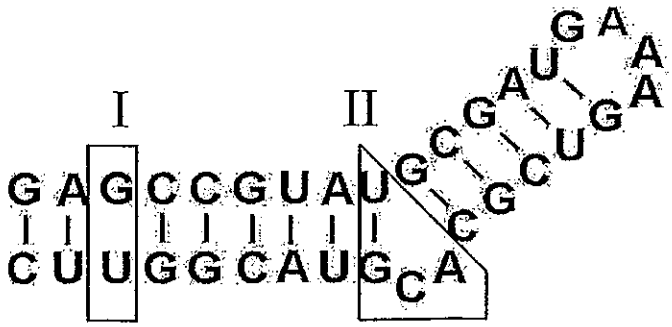


Fig. 8(b)

		Shevek Predictions						
Pos1	Pos2	Crm V	(-1)logP	DF		Corresponding Positions		
2	93	0.83011	6	4		2-C	71-G	
5	90	0.79957	6	12		5-A	68-U	
6	89	0.96785	6	12		6-U	67-A	
7	88	0.87933	6	4		7-U	66-A	
10	48	0.94944	6	2		10-m2G	45-G	
71	87	0.95895	6	8		49-m5C	65-G	
74	84	1	6	4		52-U	62-A	

Fig. 9(a)



Fig. 9(b)

		Shevek Predictions						
Pos1	Pos2	Crm V	(-1)logP	DF		Corresponding Positions		
30	121	0.88097	8	14		22-Ala	84-Ile	
64	112	0.78824	7.69897	42		50-Ile	81-Pro	

Fig. 10(a)

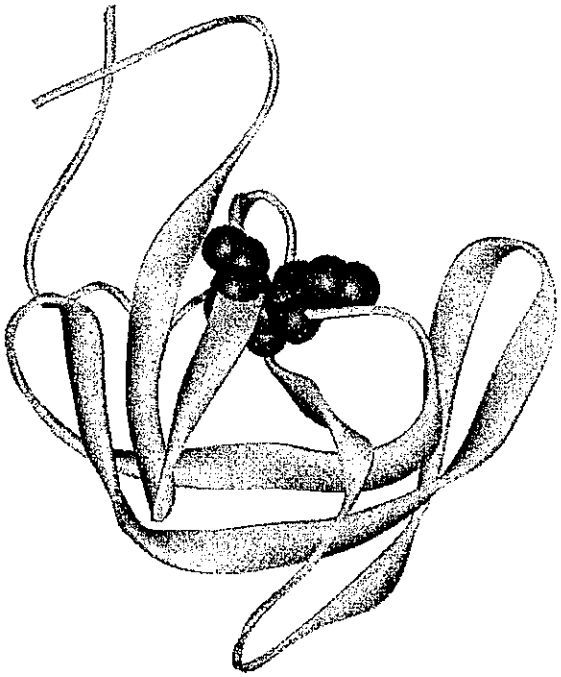


Fig. 10(b)



Fig. 10(c)

		Shevek Predictions						
Pos1	Pos2	Crm V	(-1)logP	DF		Corresponding Positions		
18	69	0.7102	7	36		19-Ala	66-Leu	
87	143	0.96041	7	14		84-Ser	139-Lys	

Fig. 11(a)

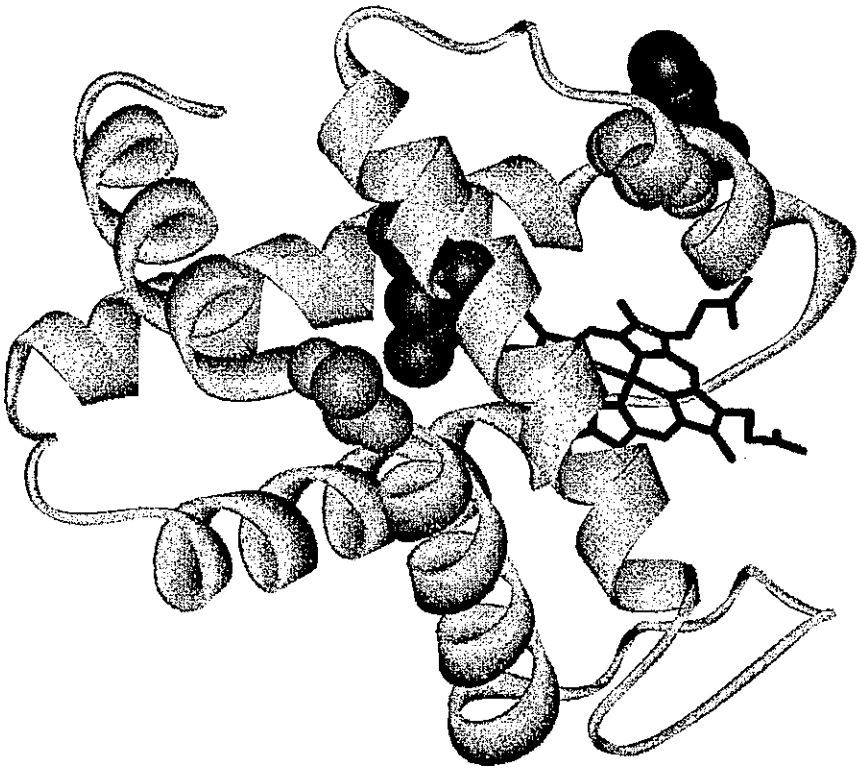
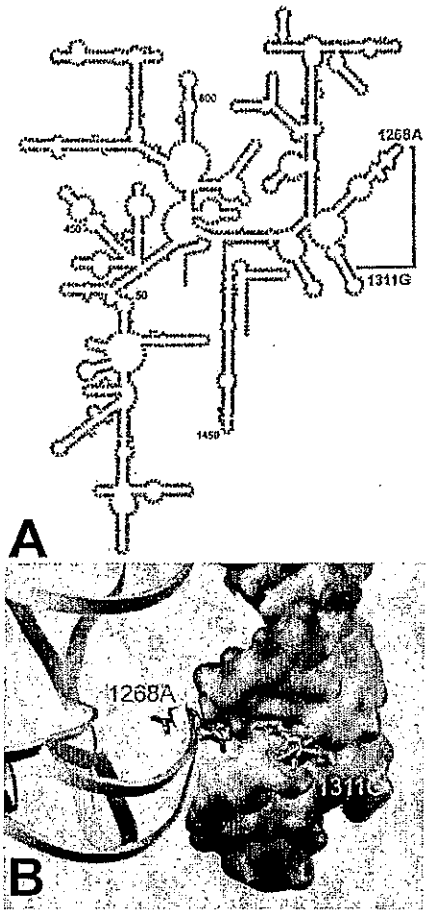
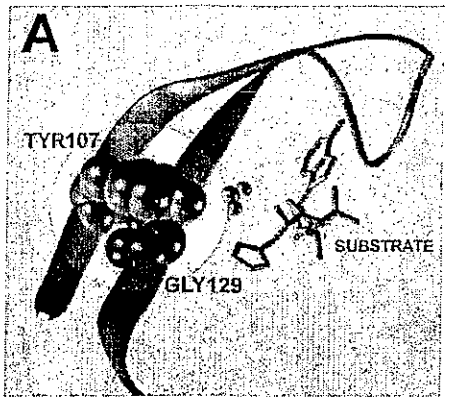


Fig. 11(b)


1268A
1311G
1450
A
1268A
1311G
B

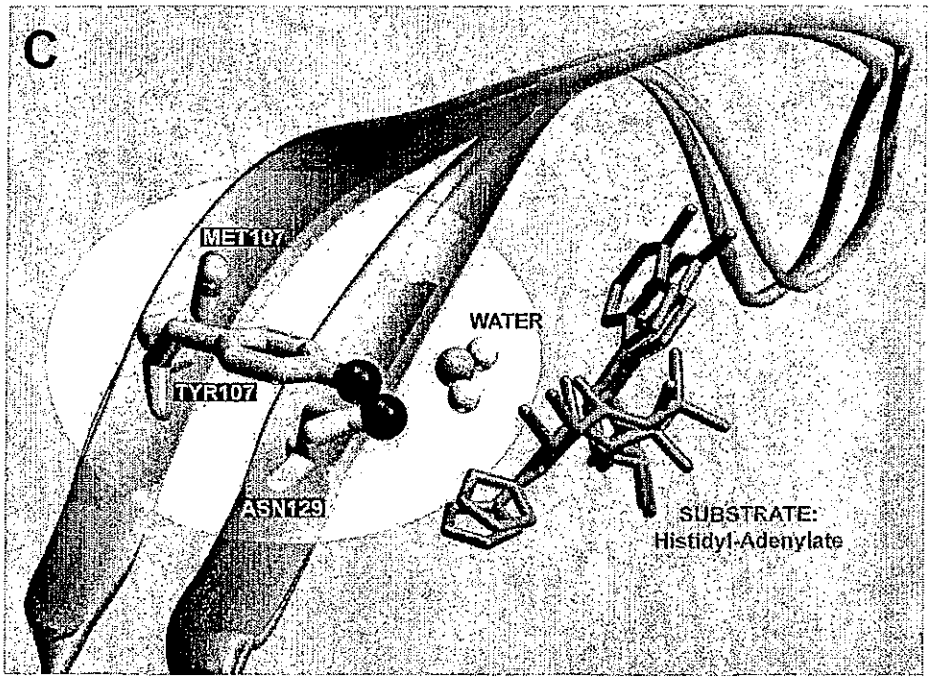


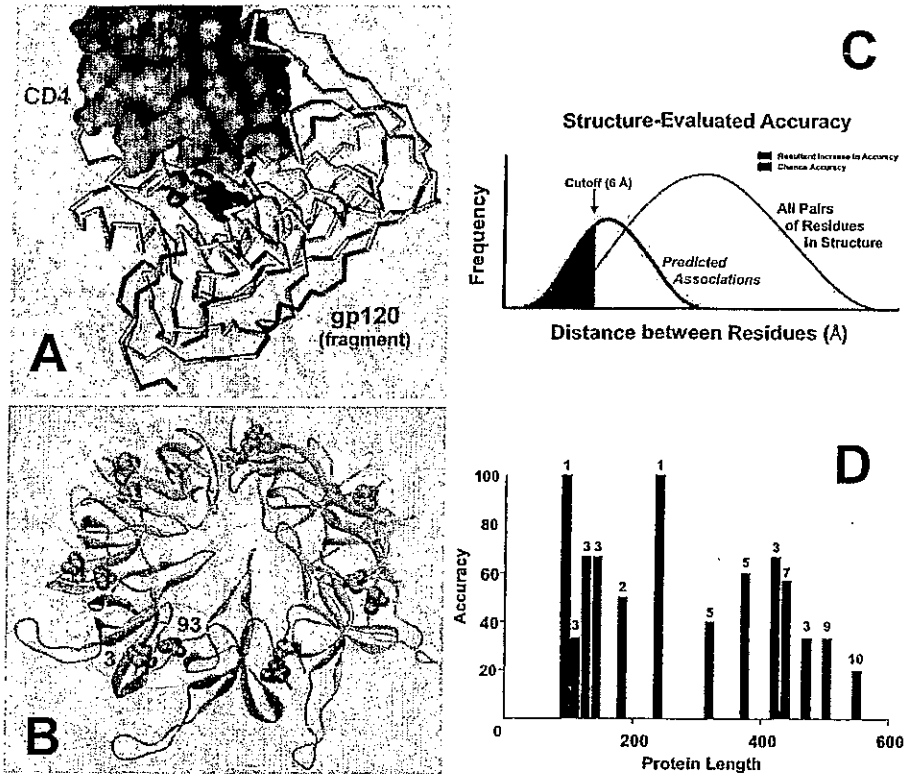
B

	Y	C	T	S	Q	H	M
G	27	0	0	0	0	0	0
S	0	1	1	0	0	0	0
N	0	0	0	2	0	0	1
D	0	0	1	0	2	3	0

Pos. 129

Pos. 107




CD4
gp120
(fragment)
A
B
93
C
Structure-Evaluated Accuracy
Cutoff (6 Å)
All Pairs
of Residues
In Structure
Predicted
Associations
Frequency
Distance between Residues (Å)
D
Accuracy
Protein Length
0
20
40
60
80
100
200
400
600