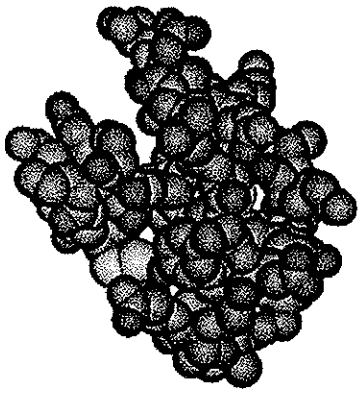


FIG 1

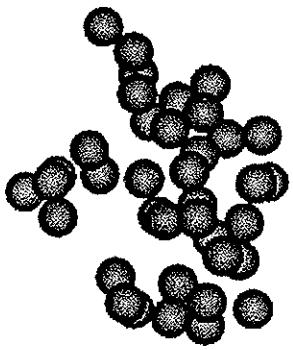
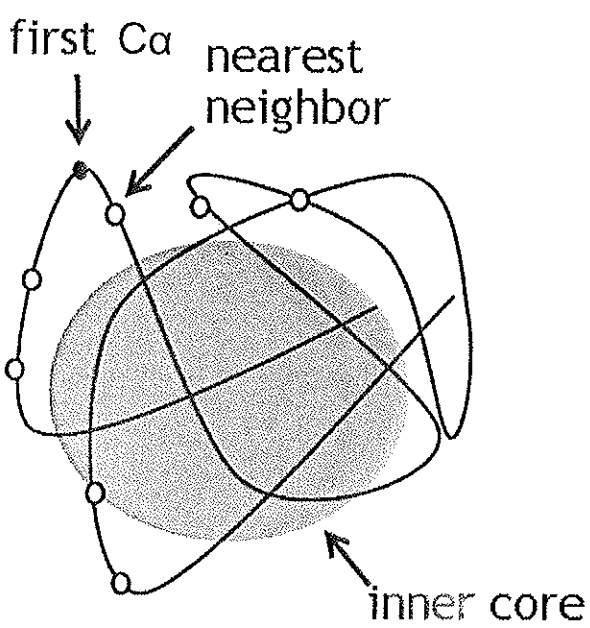
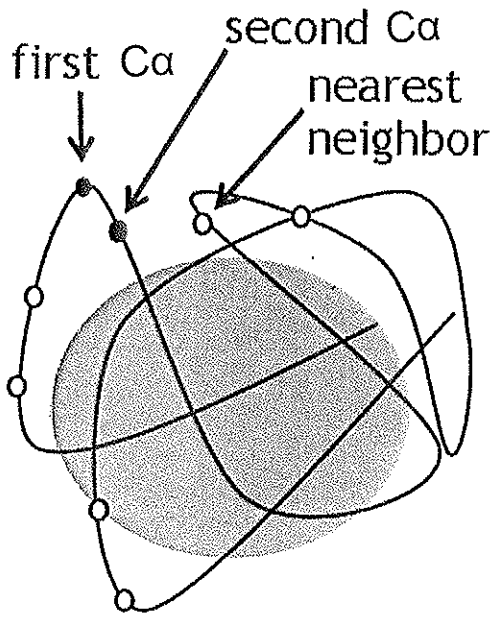


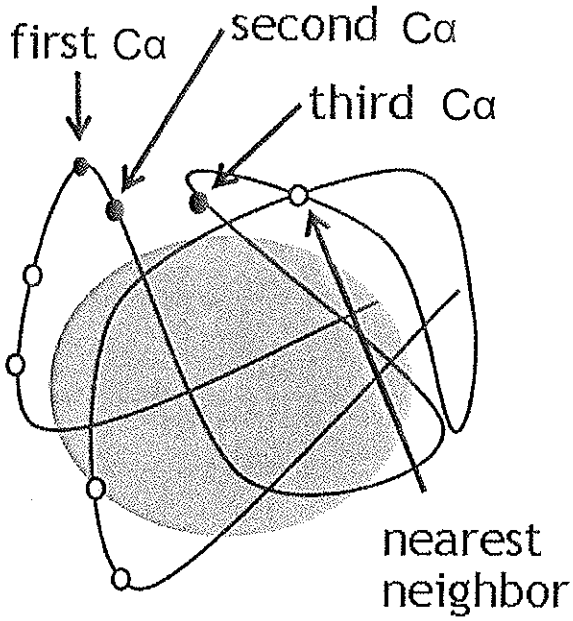
FIG 2


first Cα
nearest
neighbor
inner core

(a)


first Cα
second Cα
nearest
neighbor

(b)


first Cα
second Cα
third Cα
nearest
neighbor

(c)

FIG 3

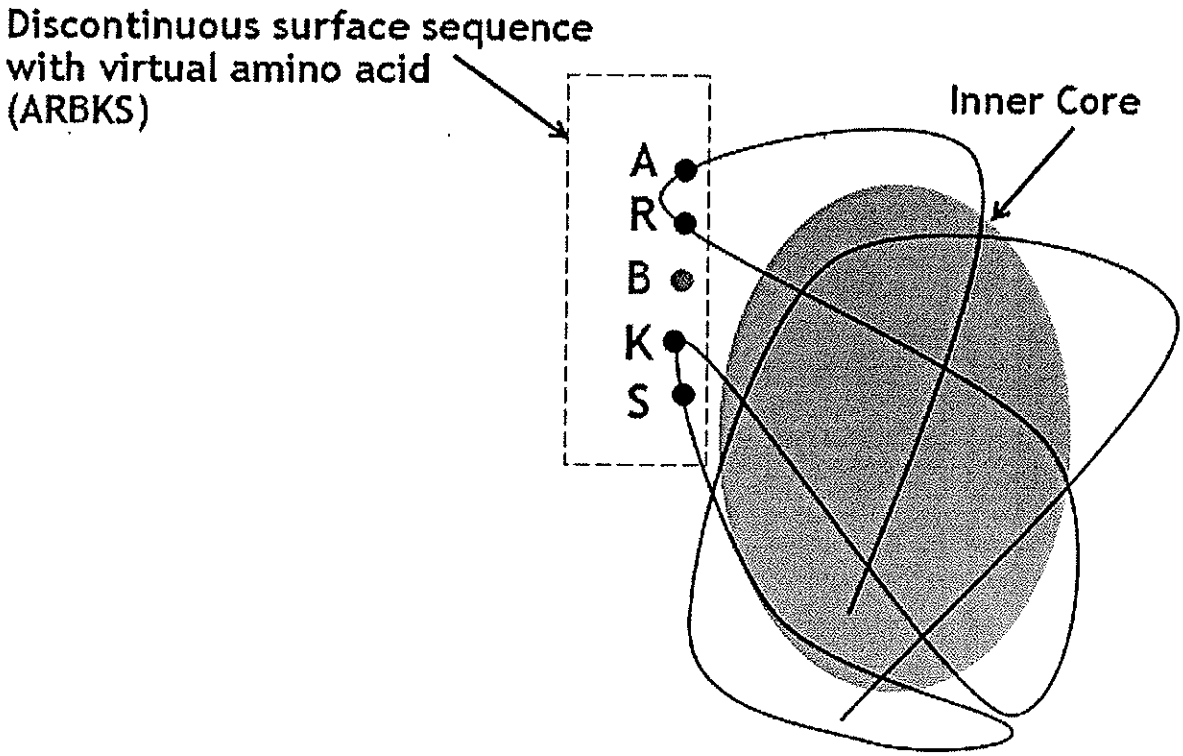

Discontinuous surface sequence
with virtual amino acid
(ARBKS)
A
R
B
K
S
Inner Core

FIG 4

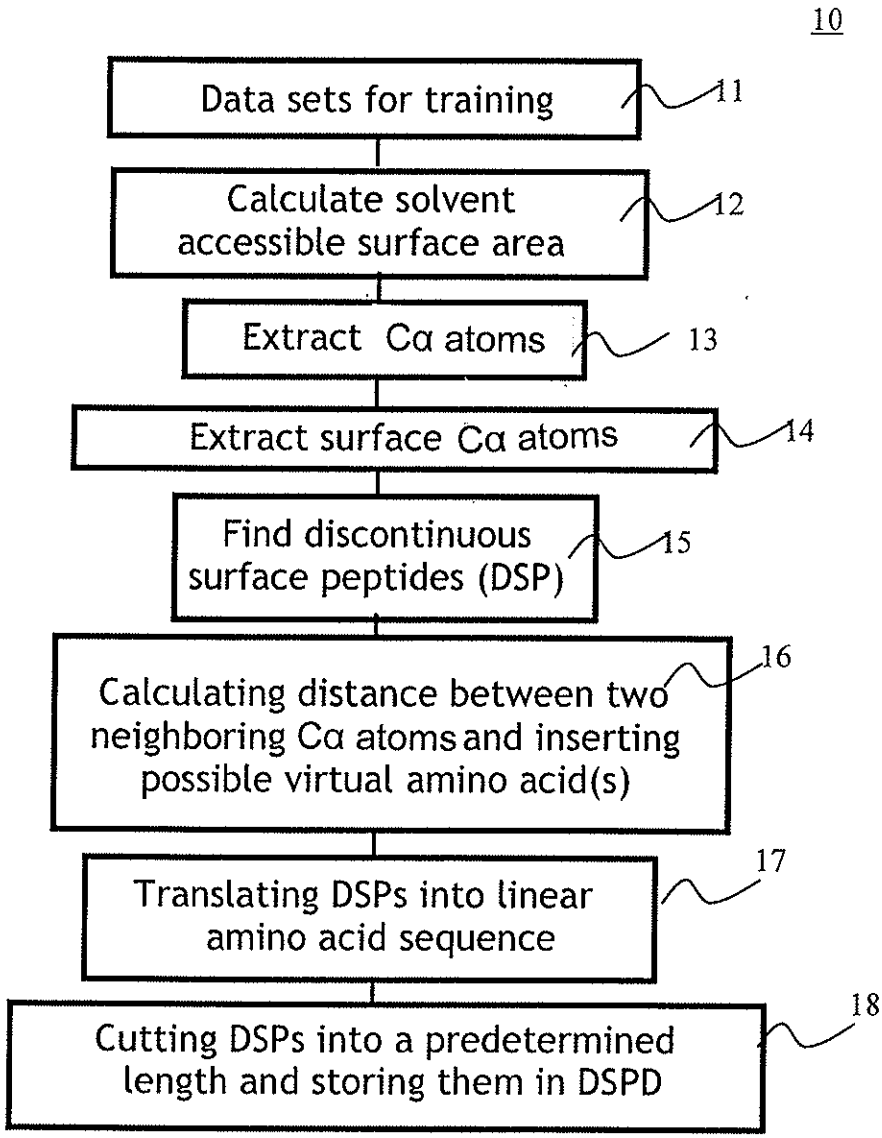

10
Data sets for training
11
Calculate solvent accessible surface area
12
Extract Cα atoms
13
Extract surface Cα atoms
14
Find discontinuous surface peptides (DSP)
15
Calculating distance between two neighboring Cα atoms and inserting possible virtual amino acid(s)
16
Translating DSPs into linear amino acid sequence
17
Cutting DSPs into a predetermined length and storing them in DSPD
18

FIG 5

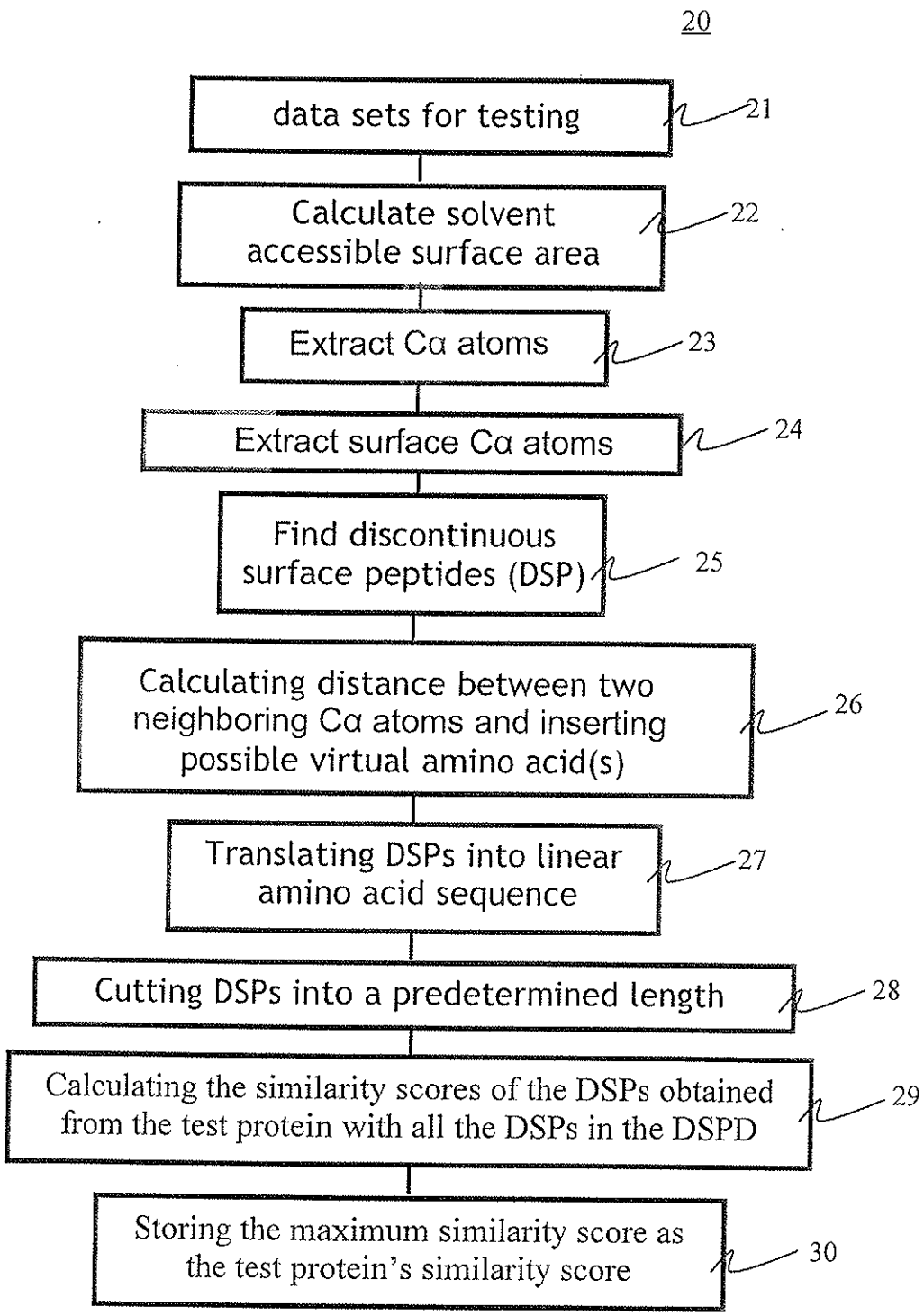

20
data sets for testing
21
Calculate solvent accessible surface area
22
Extract Cα atoms
23
Extract surface Cα atoms
24
Find discontinuous surface peptides (DSP)
25
Calculating distance between two neighboring Cα atoms and inserting possible virtual amino acid(s)
26
Translating DSPs into linear amino acid sequence
27
Cutting DSPs into a predetermined length
28
Calculating the similarity scores of the DSPs obtained from the test protein with all the DSPs in the DSPD
29
Storing the maximum similarity score as the test protein's similarity score
30

FIG 6

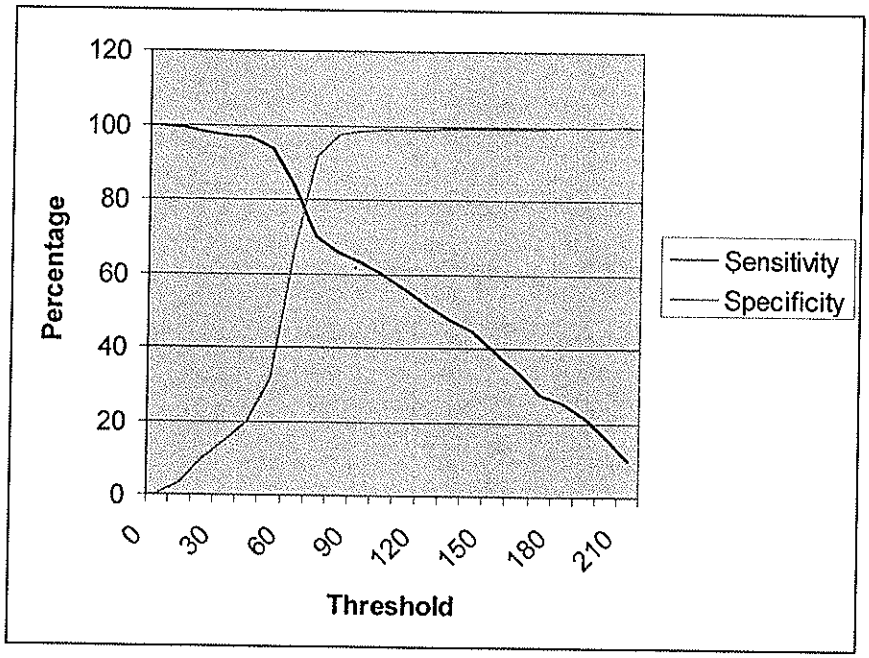


FIG 7

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V	B	Z	X	*
A	7	-3	-3	-3	-1	-2	-2	0	-3	-3	-3	-1	-2	-4	-1	2	0	-5	-4	-1	-3	-2	-1	-8
R	-3	9	-1	-3	-6	1	-1	-4	0	-5	-4	3	-3	-5	-3	-2	-2	-5	-4	-4	-2	0	-2	-8
N	-3	-1	9	2	-5	0	-1	-1	1	-6	-6	0	-4	-6	-4	1	0	-7	-4	-5	5	-1	-2	-8
D	-3	-3	2	10	-7	-1	2	-3	-2	-7	-7	-2	-6	-6	-3	-1	-2	-8	-6	-6	6	1	-3	-8
Q	-2	1	0	-1	-5	9	3	-4	1	-5	-4	2	-1	-5	-3	-1	-1	-4	-3	-4	-1	5	-2	-8
E	-2	-1	-1	2	-7	3	8	-4	0	-6	-6	1	-4	-6	-2	-1	-2	-6	-5	-4	1	6	-2	-8
G	0	-4	-1	-3	-6	-4	-4	9	-4	-7	-7	-3	-5	-6	-5	-1	-3	-6	-6	-6	-2	-4	-3	-8
H	-3	0	1	-2	-7	1	0	-4	12	-6	-5	-1	-4	-2	-4	-2	-3	-4	3	-5	-1	0	-2	-8
I	-3	-5	-6	-7	-2	-5	-6	-7	-6	7	2	-5	2	-1	-5	-4	-2	-5	-3	4	-6	-6	-2	-8
L	-3	-4	-6	-7	-3	-4	-6	-7	-5	2	6	-4	3	0	-5	-4	-3	-4	-2	1	-7	-5	-2	-8
K	-1	3	0	-2	-6	2	1	-3	-1	-5	-4	8	-3	-5	-2	-1	-1	-6	-4	-4	-1	1	-2	-8
M	-2	-3	-4	-6	-3	-1	-4	-5	-4	2	3	-3	9	0	-4	-3	-1	-3	-3	1	-5	-3	-2	-8
F	-4	-5	-6	-6	-4	-5	-6	-6	-2	-1	0	-5	0	10	-6	-4	-4	0	4	-2	-6	-6	-3	-8
P	-1	-3	-4	-3	-6	-3	-2	-5	-4	-5	-5	-2	-4	-6	12	-2	-3	-7	-6	-4	-4	-2	-3	-8
S	2	-2	1	-1	-2	-1	-1	-1	-2	-4	-4	-1	-3	-4	-2	7	2	-6	-3	-3	0	-1	-1	-8
T	0	-2	0	-2	-2	-1	-2	-3	-3	-2	-3	-1	-1	-4	-3	2	8	-5	-3	0	-1	-2	-1	-8
W	-5	-5	-7	-8	-5	-4	-6	-6	-4	-5	-4	-6	-3	0	-7	-6	-5	16	3	-5	-8	-5	-5	-8
Y	-4	-4	-4	-6	-5	-3	5	-6	3	-3	-2	-4	-3	4	-6	-3	-3	3	11	-3	-5	-4	-3	-8
V	-1	-4	-5	-6	-2	-4	-4	-6	-5	4	1	-4	1	-2	-4	-3	0	-5	-3	7	-6	-4	-2	-8
B	-3	-2	5	6	-6	-1	1	-2	-1	-6	-7	-1	-5	-6	-4	0	-1	-8	-5	-6	6	0	-3	-8
Z	-2	0	-1	1	-7	5	6	-4	0	-6	-5	1	-3	-6	-2	-1	-2	-5	-4	-4	0	6	-1	-8
X	-1	-2	-2	-3	-4	-2	-2	-3	-2	-2	-2	-2	-2	-3	-3	-1	-1	-5	-3	-2	-3	-1	-2	-8
*	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	-8	1

FIG 8