Journal of Biomolecular Structure & Dynamics, ISSN 0739-1102
Volume 21, Issue Number 2, (2003)

About Factors Providing the Fast Protein-Protein Recognition in Processes of Complex Formation

http://www.jbsdonline.com

L. N. Drozdov-Tikhomirov[1,*]
D. M. Linde[2]
V. V. Poroikov[2]
A. A. Alexandrov[1]
G. I. Skurida[1]
P. V. Kovalev[1]
V. Yu. Potapov[2]

[1]Institute of Molecular Genetics
Russian Academy of Sciences
Moscow, 123182 Russia
[2]Institute of Biomedical Chemistry
Russian Academy of Medical Sciences
Moscow, 110121 Russia

Abstracts

A package of programs for the examination of areas of subunit contacts (interface) in protein-protein (PP) complexes has been created and used for a detailed study of amino acid (AA) composition and interface structure in a large number of PP complexes from Brookhaven database (PBD). It appeared that in about 75% of the complexes, the AA composition of the subunit surface is not important. This suggests that, along with the surface AA composition, interactions between AA from the inner parts of protein globules may play a significant role in PP recognition. Such interactions between relatively distant AA residues can only be of electrostatic nature and contribute to the total electric field of the protein molecule. The configuration of the electric field itself appears to determine the PP recognition. The total electric field created by protein molecules can be calculated as a result of superimposition of the fields created by the protein multipole (i.e. by the totality of partial electric charges assigned to each atom of the molecule).

We performed preliminary calculations for the distant electrostatic interaction of ribonuclease subunits in a vacuum. The results reveal that the effect of the electric fields of the protein multipole is strong enough to orient protein molecules prior to their Brown collision.

Key words: Complex formation, Physical mechanism, Contacts, Contact pair frequency, Electrostatic interaction, Partial charges, Multipole.

Introduction

The experimentally observed, highly specific, high-rate formation of a PP complex from a pair of protein subunits gives rise to a number of questions related to the physical nature of this process. Their sum is known as a problem of PP recognition. It covers at least three fundamental questions:

1. What factors and forces determine the structure and stability of a PP complex?

2. What factors and forces determine the high rate of a PP complex formation?

3. Is there any structural code (certain AA residues) for interface area?

The answer to the first question is provided by a large number of detailed studies of the structure of PP complexes (1-18), which show that the main forces responsible for complex stability are the interactions of AA residues, located on the surfaces of the globules, driven by hydrophobic forces (4, 6, 8, 11), hydrogen bonds (15) and saline bridges (15, 17) formation, electrostatic forces (2, 8, 9, 15), and by hydratation of globule surfaces by water molecules (5, 7, 14, 15). Structural studies also show that the specificity of binding manifested by the selection of certain

*Phone: (095) 196-0201
Fax: (095) 196-0221
Email: drozdov@img.ras.ru

mutual orientations out of the large number of possible orientations in Brown collisions is determined by the existence of geometrical complementarity of the regions of surface areas responsible for the contact (13). Geometrical complementarity allows the atoms of the subunits to come together tightly, to the distance where the formation of the H-bonds, saline bridges, and other short-range interactions responsible for the high stability of association is possible.

The physical nature of the mechanism responsible for the high rate of association of protein subunits into a complex is much less clear. The measured rates of PP complex formation (~10^6 - 10^9 M^{-1}/sec^{-1}) (19) are by many orders of magnitude higher than the rate (~10 M^{-1}/sec^{-1}) (21), which should be observed in the absence of the long distance (non contact) interactions orienting the subunits prior to Brown collision. This shows unequivocally the existence of relatively strong interactions between subunits before their Brown collision. It is believed that this mutual orientation is driven by electrostatic forces (19-22). However, other suggestions have been put forward as well. S. Northrap and H. Ericson (23) calculated the rate of association from molecular dynamics of Brown collision, and they came to the conclusion that rates of ~10^6 M^{-1}/sec^{-1} could be obtained, even using a pure diffusion model, if one takes into account the effect of the hydratation sphere, which does not allow the subunits to come apart, and therefore prolongs the time they stay together.

The electrostatic theory is supported by the studies of the interface structure of the complexes (6, 10, 12, 16, 19-22) and by genetic experiments with substitutions of the polar and charged AA in the subunits of complexes by non-polar alanine residues (24). These studies show that many interfaces include polar and charged residues and their substitution by alanine, as a rule, destabilizes the complex and significantly decreases the complex formation rate. Statistically, this mechanism may work for a significant proportion of complexes, but it does not explain the complex formation in the absence of the charged and polar AAs in the interface.

The capture of the collided subunits by the hydratation sphere, which does not allow them to come apart and prolongs the time of their tight contact, may also drive the high-rate complex formation. However, the weak spot of this hypothesis is a poorly justified selection of numerical geometrical criterion of collision efficiency. If, instead of the criterion of tolerable inaccuracy of collision proposed by Northrap and Ericson (±2 Å) (23), one uses a more stringent value (for instance, ±0.2 Å), the complex formation rate calculated by their method appears to be 4-5 orders of magnitude less than the experimental one. As there is no way to predict the real requirements to the accuracy of subunits coupling, the results of Northrap and Ericson, unfortunately, could not be considered a convincing argument for the universal nature of hydrophobic mechanism of PP recognition.

The results that were apparently contradictory to the electrostatic hypothesis were obtained by various authors (25-27) in theoretical calculations of the contribution of electrostatic interaction of the surface polar and charged AA to the free energy of subunits binding in the complex. Most calculations show that the presence of charged and polar AA in the interface leads to complex destabilization. At the same time, in certain cases (28) (in particular in the case of hyperthermofilic proteins (29)), the calculations show stabilizing influence of electrostatic interactions between the surface AA in the contact areas. It should be noted that the destabilizing effect does not exclude the possible significant role of surface AA in pre-orientation of subunits when they come closer to each other prior to Brown collision. Therefore, it cannot be considered an argument against the hypothesis explaining the high rate complex formation by electrostatic interaction of the interface AA.

We believe that the rich collection of data in PBD allows getting further information in order to understand the basic physical principles of PP recognition.

Experimental Procedures

Objects Selection

For this study, we used the atomic structures from the PDB database (www.rcsb.org/pdb/). 812 PP complexes were chosen, the atomic structures of which were obtained by X-ray and NMR analysis with resolution of no less than 3.0 Å (the precision of atomic coordinates determination are about ±0.3 Å). We realized that the sample did not represent all the known classes of PP complexes in comparable amounts. Any structure formed by two polypeptide chains linked by non-covalent bonds was treated as a complex. The list of complexes included in our sample is shown in Table I. Because of its large size, Table I is presented on our Internet site (http://obi.img.rus.ru/Humbio/proteins/table1.html) and also is available on request via mail or e-mail directly from the authors.

Table I

Please note

Table I has not been printed because of its large size.

Table I can be found on the authors Internet site: (http://obi.img.rus.ru/Humbio/proteins/table1.html)

It is also available on request via postal mail or email, directly from the authors.

Search for Amino Acid Contacts

We considered two AA to have a contact if the distance between the centers of any two atoms of these AA did not differ from the sum of Van-der-Vaalse radii of these atoms by more than ±0.3 Å (the error of atomic centers coordinates). The values of Van-der-Vaalse radii were taken from the tables presented in reference (30). The search for AA contacts in a given subunit of a complex was performed as follows: for each AA of the molecule, the distance between the center of each atom to the centers of all the atoms of all the amino acids of the counterpart molecule of the complex were calculated. Using the contact criterion in such a calculation allowed us to find all the amino acids contacting the AA of the counterpart subunit for each subunit of the complex.

The data on the contacts obtained for each complex of our sample were presented as a pair of tables, one table for each subunit of the complex. They show which AA of each part of the complex have contacts with which AA of the counterpart molecule.

A set of such tables stored in a computer have made up the major experimental database, which we used for an analysis of the contact areas of P-P complexes. Table II shows such a pair of tables for complex 1an9 (ab) PDB, as an example.

Table II

The list of contacts for homodimer 1an9 (oxidaze of D-amino acids). The contacts where the AA of one subunit touches only one AA of another subunit are shown in boldface. Direct electrostatic contacts (R≤ 0.3 A) are marked by "e", distant electrostatic contacts (0.3≤R≤3 A) are marked by "**d**".

The contacts formed AA residues of subunit **A**.			The contacts formed AA residues of subunit **B**.		
Sub A		Sub B	Sub B		Sub A
ALA 85	→	TRP 209	ARG 120	→	LEU 112, MET 110, VAL 111, d ASP 109
ARG 120	→	e ASP 109, MET 110	**ASN 86**	→	**LYS 271**
ASN 83	→	LYS 271	ASP 109	→	e ARG 120
ASN 86	→	**ASP 272**	**ASP 272**	→	**ASN 86**
ASP 109	→	d ARG 120	GLY 232	→	LYS 211
GLY 232	→	LYS 211	**LEU 233**	→	**PRO 208**
LEU 112	→	ARG 120	LYS 211	→	GLY 232, LYS 211
LYS 211	→	GLY 232, LYS 211	LYS 271	→	ASN 83, PRO 82
LYS 271	→	**ASN 86**	MET 110	→	ARG 120
MET 110	→	ARG 120	PHE 133	→	PHE 133
PHE 133	→	PHE 133, SER 93	**PRO 82**	→	**PRO 268**
PRO 82	→	LYS 271, PRO 268	PRO 268	→	PRO 82
PRO 208	→	**LEU 233**	SER 93	→	PHE 133
PRO 268	→	**PRO 82**	THR 90	→	TRP 209
THR 90	→	TRP 209	TRP 209	→	THR 90, ALA 85
TRP 209	→	THR 90			
VAL 111	→	ARG 120			

A pair of values representing the number of contacting amino acids on the surface of one and of the other part of the complex may serve as a characteristic of the contact, which has a meaning close to the space of the surface area shielded from water molecules (16). These values may serve as characteristics of the size of the interface areas in the complexes and are shown in the two corresponding columns of Table I.

Determination of the Experimental Frequency of Occurrence

Each of the 210 possible AA contacts (the number of possible two from twenty combinations) was performed using a special program counting the contacts of every type in the entire database of contacts.

Calculation of Theoretical Frequency of Occurrence

Performed using a model of random collision of two identical balls, that had surfaces divided randomly into 20 segments, the sizes of which were proportional to the values of the relative space occupied by various amino acids on the surface of a generalized protein representing the totality of proteins as reported by Conte *et al*. (15).

When such balls collide randomly, the probability that i-AA of one subunit of the complex touches j-AA of the other will be:

$$\mathbf{P_{ij}} = \gamma_i \cdot \gamma_j,$$

where γ_k is the portion of the ball occupied by the AA of the k type, and k = (1…20; all amino acid types).

The mathematical expectation of the occurrence of the contact of i-AA with j-AA after N collisions will be in this case:

$$\mathbf{M_{ij}(N) = N \cdot P_{ij}}$$

Table III

Mathematical expectation of the frequencies of possible contact pairs upon random contact formation. Gray boxes mark the positions corresponding to contact pairs for the frequency of which the mathematical expectation is more than 500. As the matrix is diagonally symmetrical only the positions above the diagonal are marked.

	ALA	ARG	ASN	ASP	CYS	GLN	GLU	GLY	HIS	ILE	LEU	LYS	MET	PHE	PRO	SER	THR	TRP	TYR	VAL
ALA	130	289	201	231	24	195	318	146	62	79	134	383	40	66	166	273	237	43	104	114
ARG	289	***640***	446	***511***	52	432	***704***	325	138	174	296	***848***	88	145	368	***604***	***525***	95	231	253
ASN	201	446	311	356	36	301	491	226	96	121	206	***591***	61	101	256	421	366	66	161	176
ASP	231	511	356	408	42	345	***562***	259	110	139	236	***677***	70	116	294	482	419	76	185	202
CYS	24	52	36	42	4	35	58	27	11	14	24	69	7	12	30	49	43	8	19	21
GLN	195	432	301	345	35	292	476	219	93	118	200	***572***	60	98	248	408	355	64	156	171
GLU	318	***704***	491	***562***	58	476	***775***	357	152	192	326	***933***	97	160	405	***665***	***578***	105	255	278
GLY	146	325	226	259	27	219	357	165	70	88	150	430	45	74	186	306	266	48	117	128
HIS	62	138	96	110	11	93	152	70	30	38	64	183	19	31	79	131	114	21	50	55
ILE	79	174	121	139	14	118	192	88	38	47	81	231	24	40	100	164	143	26	63	69
LEU	134	296	206	236	24	200	326	150	64	81	137	392	41	67	170	279	243	44	107	117
LYS	383	***848***	***591***	677	69	***572***	***933***	430	183	231	392	***1123***	117	193	487	***800***	***696***	126	307	335
MET	40	88	61	70	7	60	97	45	19	24	41	117	12	20	51	83	72	13	32	35
PHE	66	145	101	116	12	98	160	74	31	40	67	193	20	33	84	137	119	22	53	58
PRO	166	368	256	294	30	248	405	186	79	100	170	487	51	84	211	347	302	55	133	145
SER	273	***604***	421	482	49	408	***665***	306	131	164	279	***800***	83	137	347	***570***	496	90	218	239
THR	237	***525***	366	419	43	355	***578***	266	114	143	243	***696***	72	119	302	496	431	78	190	208
TRP	43	95	66	76	8	64	105	48	21	26	44	126	13	22	55	90	78	14	34	38
TYR	104	231	161	185	19	156	255	117	50	63	107	307	32	53	133	218	190	34	84	91
VAL	114	253	176	202	21	171	278	128	55	69	117	335	35	58	145	239	208	38	91	100

The calculated values of mathematical expectation of the occurrence of all possible contact AA pairs are shown in Table III.

The Possible Orienting Effect of the Distant Electrostatic Interaction

The subunits was estimated by calculating the energy of the system of two interacting subunits, the mass centers of which were fixed in vacuum at the distance of ~100A, for different angles between subunits. Also the depth of the energy minimums was determined in appropriate orientations.

The calculation of the summary of electrostatic interaction was performed by the method proposed by E. Kong (ftp://dashes.wustl.edu/pub/papers/kong-thesis.pz.gz), which considerably saves the calculation time.

The calculation of the distribution of partial atom charges of AA in protein subunits was performed using the SYBYL package by various methods (Gasteiger, Gasteiger-Huckel, Huckel).

Results

We performed the analysis of AA composition and the structure of the interface areas in 812 PP complexes. The total number of AA-AA contacts, the number and the type of electrostatic contacts (charged AA- oppositely charged AA), and the number of hydrophobic contacts (non-polar AA-non-polar AA) were established for each interface (Table I).

For 289 complexes, the interface structure was encoded in a matrix form. Such presentation shows a number of contacts of each possible type present in the interface. In Table IV, as an example, the matrix of contacts for one complex is shown. The matrix approach was used for the testing of the hypothesis of L. Mekler (31) (Table V) and the related hypothesis of J. Blalock and colleagues (32), who proposed that the structural code for contact areas based on a hypothetical, specific complementary interaction of two AA are determined by the complementarity of their codons. The testing results are documented in Table VI, which shows a proportion of the complexes studied that have at least one AA contact pair postulated by this hypothesis in their interface areas. It can be seen that for each hypothesis, the required specific contacts occur in less than a half of the complexes studied.

Table IV

The matrix of contact pairs in the interface formed by subunits A and B of phosphofructokinase from B. stearothermophylis (6PFK)

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	0	0	**2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
R	0	0	0	**1**	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0
N	**2**	0	0	**1**	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	**1**	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Q	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0
G	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
I	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
K	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0
M	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
S	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
V	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table V

The matrix representing the AA-hypothesis of Mekler. The complementary contact pairs which should be present in complex interfaces according to the hypothesis are designated by "1".

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	0	**1**	0	0	**1**	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0
R	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	**1**	**1**	0	0	0
N	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	***1***
D	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	**1**
C	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0
Q	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0	**1**	0	0	**1**	0	0	0	0	0	0
G	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	**1**	**1**	0	0	0
H	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	**1**
I	0	0	**1**	**1**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0
L	0	0	0	0	0	**1**	**1**	0	0	0	0	**1**	0	0	0	0	0	0	0	0
K	0	0	0	0	0	0	0	0	0	0	**1**	0	0	**1**	0	0	0	0	0	0
M	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	**1**	0	0	0	0	**1**	0	0	0	0	0	0	0	0
P	0	**1**	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	**1**	0	0
S	**1**	**1**	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	**1**	0	0	0
T	0	**1**	0	0	**1**	0	0	**1**	0	0	0	0	0	0	0	**1**	0	0	0	0
W	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0
Y	0	0	0	0	0	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	**1**
V	0	0	**1**	**1**	0	0	0	0	**1**	0	0	0	0	0	0	0	0	0	**1**	0

Table VI

Testing of contact hypotheses. 1-hypothesis - AA hypothesis of Mekler (32), 2-hypothesis - first hypothesis of Blalock *et al*. (34), 3-hypothesis - second of hypothesis of Blalock *et al*. (34).

	1-hypothesis	2- hypothesis	3-hypothesis
The total number of the complexes studied	289	289	289
The number of complexes, the interface of which has at least one contact postulated by the hypothesis.	94	116	150

Table VII

The frequencies of the contact pairs of all possible types in our experimental sample of 820 complexes. Gray boxes mark the positions corresponding to the contact pairs for the frequency of which the mathematical expectation is more than 500. As the matrix is diagonally symmetrical only the positions above the diagonal are marked.

	ALA	ARG	ASN	ASP	CYS	GLN	GLU	GLY	HIS	ILE	LEU	LYS	MET	PHE	PRO	SER	THR	TRP	TYR	VAL
ALA	316	296	256	170	56	172	158	170	108	184	236	102	204	206	96	184	170	102	272	188
ARG	296	**548**	382	**974**	72	406	**876**	330	200	192	**446**	230	156	298	328	290	342	164	**524**	280
ASN	256	382	328	194	24	178	222	266	102	192	268	292	88	172	160	206	240	96	288	126
ASP	170	**974**	194	244	66	198	132	170	284	84	156	**538**	88	136	124	276	220	144	262	88
CYS	56	72	24	66	156	38	46	76	40	36	32	62	38	46	58	52	30	28	68	36
GLN	172	406	178	198	38	400	196	198	130	190	240	188	156	170	200	206	202	88	258	194
GLU	158	**876**	222	132	46	196	216	198	252	200	196	**634**	122	198	208	270	252	80	386	222
GLY	170	330	266	170	76	198	198	360	108	138	202	166	108	124	214	210	202	100	314	190
HIS	108	200	102	284	40	130	252	108	256	128	216	106	138	76	96	138	184	96	240	180
ILE	184	192	192	84	36	190	200	138	128	268	338	150	142	304	78	138	186	130	198	318
LEU	236	**446**	268	156	32	240	196	202	216	338	796	216	208	338	210	274	272	162	378	398
LYS	102	230	292	**538**	62	188	**634**	166	106	150	216	196	102	154	120	216	194	76	218	136
MET	204	156	88	88	38	156	122	108	138	142	208	102	292	174	140	108	104	98	214	192
PHE	206	298	172	136	46	170	198	124	76	304	338	154	174	376	210	162	236	208	312	232
PRO	96	328	160	124	58	200	208	214	96	78	210	120	140	210	252	176	162	138	324	158
SER	184	290	206	276	52	206	270	210	138	138	274	216	108	162	176	320	258	96	218	158
THR	170	342	240	220	30	202	252	202	184	186	272	194	104	236	162	258	368	98	246	232
TRP	102	164	96	144	28	88	80	100	96	130	162	76	98	208	138	96	98	180	136	76
TYR	272	**524**	288	262	68	258	386	314	240	198	378	218	214	312	324	218	246	136	324	246
VAL	188	280	126	88	36	194	222	190	180	318	398	136	192	232	158	158	232	76	246	328

In the extended sample including 812 complexes, we have determined the frequencies of each of the 210 possible types of contact pairs in the interface areas (Table VII). The observed frequency values were compared to theoretical values, and the ratio between the observed frequency and its mathematical expectation (Table VIII) was determined in the assumption that the proportion of the surface occupied by each AA type was equal to the values presented by Conte *et al*. (15), and that the contacts were formed randomly.

Table VIII

The ratio between actual frequencies of the contact pairs to the mathematical expectation calculated for random contact formation. The positions corresponding to contact pairs the actual frequency of which exceeds their mathematical expectation more than 5 times are marked by light gray boxes and more than 10 times, by dark gray boxes. As the matrix is diagonally symmetrical only the positions above the diagonal are marked.

	ALA	**ARG**	ASN	ASP	CYS	GLN	GLU	GLY	HIS	ILE	LEU	LYS	MET	PHE	PRO	SER	THR	TRP	TYR	VAL
ALA	2.4	1.0	1.3	0.7	2.4	0.9	0.5	1.2	1.7	2.3	1.8	0.3	***5.1***	3.1	0.6	0.7	0.7	2.4	2.6	1.6
ARG	1.0	0.9	0.9	1.9	1.4	0.9	1.2	1.0	1.4	1.1	1.5	0.3	1.8	2.0	0.9	0.5	0.7	1.7	2.3	1.1
ASN	1.3	0.9	1.1	0.5	0.7	0.6	0.5	1.2	1.1	1.6	1.3	0.5	1.4	1.7	0.6	0.5	0.7	1.4	1.8	0.7
ASP	0.7	1.9	0.5	0.6	1.6	0.6	0.2	0.7	2.6	0.6	0.7	0.8	1.3	1.2	0.4	0.6	0.5	1.9	1.4	0.4
CYS	2.4	1.4	0.7	1.6	***36.5***	1.1	0.8	2.9	3.5	2.5	1.3	0.9	***5.3***	3.9	1.9	1.1	0.7	3.6	3.6	1.7
GLN	0.9	0.9	0.6	0.6	1.1	1.4	0.4	0.9	1.4	1.6	1.2	0.3	2.6	1.7	0.8	0.5	0.6	1.4	1.7	1.1
GLU	0.5	1.2	0.5	0.2	0.8	0.4	0.3	0.6	1.7	1.0	0.6	0.7	1.3	1.2	0.5	0.4	0.4	0.8	1.5	0.8
GLY	1.2	1.0	1.2	0.7	2.9	0.9	0.6	2.2	1.5	1.6	1.3	0.4	2.4	1.7	1.1	0.7	0.8	2.1	2.7	1.5
HIS	1.7	1.4	1.1	2.6	3.5	1.4	1.7	1.5	***8.6***	3.4	3.4	0.6	***7.2***	2.4	1.2	1.1	1.6	4.7	4.8	3.3
ILE	2.3	1.1	1.6	0.6	2.5	1.6	1.0	1.6	3.4	5.7	4.2	0.7	***5.9***	***7.7***	0.8	0.8	1.3	***5.0***	3.1	4.6
LEU	1.8	1.5	1.3	0.7	1.3	1.2	0.6	1.3	3.4	4.2	***5.8***	0.6	***5.1***	***5.0***	1.2	1.0	1.1	3.7	3.5	3.4
LYS	0.3	0.3	0.5	0.8	0.9	0.3	0.7	0.4	0.6	0.7	0.6	0.2	0.9	0.8	0.2	0.3	0.3	0.6	0.7	0.4
MET	***5.1***	1.8	1.4	1.3	***5.3***	2.6	1.3	2.4	7.2	***5.9***	***5.1***	0.9	***24.1***	***8.7***	2.8	1.3	1.4	7.5	***6.7***	***5.5***
PHE	3.1	2.0	1.7	1.2	3.9	1.7	1.2	1.7	2.4	***7.7***	***5.0***	0.8	***8.7***	***11.4***	2.5	1.2	2.0	***9.6***	***5.9***	4.0
PRO	0.6	0.9	0.6	0.4	1.9	0.8	0.5	1.1	1.2	0.8	1.2	0.2	2.8	2.5	1.2	0.5	0.5	2.5	2.4	1.1
SER	0.7	0.5	0.5	0.6	1.1	0.5	0.4	0.7	1.1	0.8	1.0	0.3	1.3	1.2	0.5	0.6	0.5	1.1	1.0	0.7
THR	0.7	0.7	0.7	0.5	0.7	0.6	0.4	0.8	1.6	1.3	1.1	0.3	1.4	2.0	0.5	0.5	0.9	1.3	1.3	1.1
TRP	2.4	1.7	1.4	1.9	3.6	1.4	0.8	2.1	4.7	***5.0***	3.7	0.6	***7.5***	***9.6***	2.5	1.1	1.3	***12.7***	3.9	2.0
TYR	2.6	2.3	1.8	1.4	3.6	1.7	1.5	2.7	4.8	3.1	3.5	0.7	***6.7***	***5.9***	2.4	1.0	1.3	3.9	3.9	2.7
VAL	1.6	1.1	0.7	0.4	1.7	1.1	0.8	1.5	3.3	4.6	3.4	0.4	***5.5***	4.0	1.1	0.7	1.1	2.0	2.7	3.3

Estimation of the effect the electric fields of partial charges of the atoms of protein molecules have on the mutual orientation of subunits coming closer to each other during Brown collision was performed for ribonuclease homodimers from Streptomyces aureofaciens (1rge PDB). It appeared that in the system of two electrically interacting subunits with their mass centers fixed at the distance of 100A (Fig. 1), changing the angles of the subunits relative to each other through all possible values caused the energy of the system to come through 9 deep minimums of about the same depth. This minimal value (~$1.2 \cdot 10^{-2}$ e^2/Å or ~$2.9 \cdot 10^{-20}$ J) was approximately by an order of magnitude greater than the kT energy at room temperature ($7.04 \cdot 10^{-2}$ e^2/Å or $4.14 \cdot 10^{-21}$ J) (Table IX). Next we analyzed, in each minimum, the "sharpness of tuning," which is the value inverse to the angle of deviation from the orientation with minimal energy that leads to the increase of the energy equal to $^1/_2$ kT. It appeared to be very sensitive to the turns around Y, Z and Y', Z' axes (~0.25 degree^{-1}) and much less so to the turns around X and X' axes (~0.03 degree^{-1}). The calculated angles of orientation of the distant subunits in minimum energy states appeared to be relatively close to the orientation in which they are fixed in the complex [0,0,0,0,0,0], but not exactly equal to it (deviation 15-25 degrees).

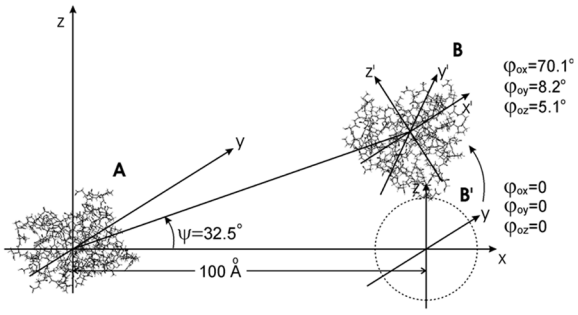

Figure 1: The mutual orientation of ribonuclease subunits A and B, the centers of which fixed at 100 Å distance, at the minimum of electrostatic interaction energy position.

A - subunit A in a fixed position (Brookheiven Protein Bank coordinates).

B' - the position of subunit B displaced from A-B complex without rotation for 100 Å along the x axis, which connected the centers of subunits A and B.

B - subunit B removed from position B' by rotation around A and around its own center with the distance between the centers of A and B fixed at 100 Å to the position which makes minimal electrostatic interaction energy of system.

φ_{ox} (φ_{oy}, φ_{oz}) - rotation angles relative to the B' coordinate axis.

Discussion

As it follows from Table VI, the results of the testing do not prove the Mekler (31) hypothesis and Blalock *et al*. (32) hypothesis, because neither can explain the existence of even half of the actually observed com-

Table IX

The minimums for electrostatic energy of the system of the two distant ribonuclease subunits (100A between the mass centers), in vacuum. N - the number of potential minimum; E – energy; Acos x(y,z), cosinuses of the turning angles of subunit A around the x(y,z) - axis relative to the position in the complex (for complex (1,1,1)); Bcos x(y,z), cosinuses of the turning angles of subunit B around the x'(y',z') - axis relative to the position in the complex (for complex (1,1,1))
E - energy (e^2 /Å) (ΠPN T=300°K kT = 8.04 E-2 e^2 /Å)

Gasteiger (method for parcial charges calculation)

N		1	2	3	4	5	6	7	8	9
E		-1.1B2E-02	-1.165E-02	-1.157E-02	-1.162E-02	-1.167E-02	-1.163E-02	-1.167E-02	-1.161E-02	-1.161E-02
	x	-0.08	-0.08	0.36	-0.48	0.79	-0.82	0.92	0.95	0.97
Acos	y	0.97	0.97	0.88	0.87	0.53	0.57	0.29	0.31	0.14
	z	-0.23	-0.23	-0.30	-0.13	0.29	0.02	0.25	0.11	-0.22
	x	0.30	-0.05	-0.29	0.45	-0.50	0.65	-0.66	0.75	0.72
Bcos	y	-0.24	-0.11	0.00	-0.25	0.02	-0.28	0.08	-0.32	0.14
	Z	-0.92	-0.99	-0.96	-0.86	-0.87	-0.71	-0.75	-0.58	-0.68

Gasteiger-Huckel

N		1	2	3	4	5	6	7	8	9
E		-1.252E-02	-1.251E-02	-1.249E-02	-1.254E-02	-1.255E-02	-1.250E-02	-1.253E-02	-1.247E-02	-1.242E-02
	x	-0.08	-0.08	0.16	-0.46	0.79	-0.82	0.92	-0.95	0.96
Acos	y	0.97	0.97	0.95	0.67	0.53	0.57	0.29	0.31	0.26
	Z	-0.23	-0.23	-0.27	-0.13	-0.29	0.02	-0.25	0.11	0.08
	x	0.15	-0.11	-0.28	0.45	0.50	0.65	-0.86	0.75	-0.79
Bcos	y	-0.19	-0.11	-0.07	-0.25	0.02	-0.28	0.08	-0.32	0.31
	Z	-0.97	-0.99	-0.96	-0.86	-0.87	-0.71	-0.75	-0.58	-0.52

Huckel

N		1	2	3	4	5	6	7	8
E		-1,185E-02	-1.186E-02	-1,183E-02	-1.188E-02	-1.182E-02	-1.187E.-02	-1.181E-02	-1.183E-02
	x	-0.08	-0.08	- 0.48	0.79	-0.82	0.92	-0.95	0. 97
Acos	y	0.97	0.97	0.87	0.53	0.57	0.29	0.31	0.14
	z	-0.23	-0.23	-0.13	-0.29	0.02	-0.25	0.11	-0.22
	x	0.30	-0.11	0.45	-0.50	0.65	-0..00	0.75	-0.72
Bcos	y	-0.24	-0.11	-0.25	0.02	0.28	0.08	-0.32	0.14
	z	-0.92	-0.99	-0.86	-0.87	-0.71	-0.75	-0.58	-0.88

plexes. Thus, these elegant theories, unfortunately, do not provide a key to understanding the molecular mechanisms of PP recognition.

The analysis of the frequency of contact AA pairs in the interfaces of the complexes in our sample of 812 objects (Table VII) showed that the more frequent contacts occur between polar and charged AA. Apparently, it argues for the mechanism based on electrostatic interaction of AA residues located on the surfaces of interacting subunits. However, as it can be seen from Table VIII, only in the case of 13 contact pairs is the observed frequency much higher (more than four times) than the theoretical one calculated for random contact formation: C:C; H:H; L:L; F:F; W:W; M:C; M:H; M:I; M:L; M:M; M:F; M:W; M:Y. It is interesting that all these pairs with the exception of two, which contain positively charged histidine, are formed by non polar (hydrophobic) AA. However, one cannot believe that these more frequent contacts are the ones defining the recognition, because they are observed only in 24% of the complexes studied.

The observed high frequency of the contacts formed by oppositely charged AA (electrostatic contacts) shows an important role of electric interactions of the surface AA residues in the complex formation mechanism. At the same time in 56% of the complexes studied, no electrostatic contacts were observed. The assumption that the subunits recognition may require either hydrophobic or electrostatic contacts also does not account for the situation, because 46% of complexes have neither one, but the complex formation rates and the stability are still high.

The results obtained, in our opinion, show that the mechanism of complex formation cannot be regarded as purely an interaction between the surface AA of the subunits. To understand this process, one has to take into account the interactions between AA residues located in the inner parts of protein molecules, below their surface. There are sufficient grounds to think (27) that such interactions may be realized through electric fields generated by the totalities of the partial atom charges of all the AA of each subunit forming electrically neutral protein multipole. So the precise orientation likely may be achieved because of the electric interaction of the multipoles of two protein molecules.

Deep energy minimums (~10kT at T=300°K) existed at certain angles of mutual orientation of ribonuclease subunits (1rge PDB) (fixed at the distance of 100 Å, globule diameter ~50 Å), were observed in our preliminary calculations. It argues for the possibility of the strong, long distance, orienting effect produced by an electrostatic interaction of the protein multipoles and supports the view on the important role of electrostatic interaction in the preorientation of protein molecules before Brown collision.

Our calculations were performed for the charges in a vacuum. However, the real molecules are immersed in water, a dielectric medium, which could be believed to screen and dramatically decrease electrostatic intensity. We believe, though, that it does not compromise our results. The very notion that the layer of "water" between subunits should decrease their electrostatic interaction is based on the formal application of the laws, which are true for electric fields in a continuous medium. It is clear that in the calculation of an electric field produced by the protein molecule at distances comparable with their size (and we are interested in this particular situation), one cannot consider the "medium" around the protein molecule as continuous and the electric field as homogeneous. Therefore, it is very doubtful that the interaction in this case may be described using the dielectric constant, i.e. the law laid down for homogeneous electric field in continuous media. But this problem, of course, requires a more detailed study.

Acknowledgments

Authors express their gratitude to senior staff fellow of the Institute of Biomedical Chemistry of the Russian Academy of Medical Sciences V. S. Skvortsov who provided an expert help in calculations of partial charges of the atoms of a ribonuclease molecule.

References and Footnotes

1. Chotia, C., Janin, J. *Nature 256*, 705-708 (1975).
2. Rashin, A. A., Honig, B. *J. Mol. Biol. 173*, 515-521 (1984).
3. Thornton, J. M., Singh, J., Campbell, S., Blundell, T. L. *Biochem. Soc. Trans. 16*, 927-930 (1988).
4. Young, L., Jernigan, R. L., Covell, D. G. *Protein. Sci. 3*, 717-29 (1994).
5. Janin, J. *Biochimie. 77*, 497-505 (1995).
6. Jones, S., Thornton, J. *Proc. Natl. Acad. Sci. USA 93*, 13-20 (1996).
7. Janin, J. *Prog. Biophys. Mol. Biol. 64*, 145-166 (1996).
8. Norel, R., Petrey, D., Wolfson, H. J., Nussinov, R. *Proteins 36*, 307-17 (1999).
9. Tsai, J., Lin, S. L., Wolfson, H., Nussinov, R. *J. Mol. Biol. 260*, 604-620 (1996).
10. Ausiello, G., Cesareni, G., Helmer-Citterich, M. *Proteins 28*, 556-67 (1997).
11. Jones, S., Thornton, J. *J. Mol. Biol. 272*, 121-132 (1997).
12. Janin, J., Miller, S., Chotia, C. *J. Mol. Biol. 204*, 155-164 (1988).
13. Lawrence, M. C., Colman, P. M. *J. Mol. Biol. 234*, 946-950 (1993).
14. Janin, J. *Struc. Fold. Des. 7*, 277-279 (1999).
15. Conte, L. L., Chothia, C., Janin, J. *J. Mol. Biol. 285*, 2177-2198 (1999).
16. Palma, P. N., Krippahl, L., Wampler, J. E., Moura, J. J. *Proteins 39*, 372-384 (2000).
17. Xu, D., Tsai, C. J., Nussinov, R. *Protein Eng. 10*, 999-1012 (1997).
18. Larsen, N. A., Olson, A. J., Goodsell, D. S. *Structure 6*, 412-427 (1998).
19. Buckle, A. M., Schreiber, G., Fersht, A. R. *Biochemistry 33*, 8878-89 (1994).
20. Zhou, H. X. *Biophys. J. 64*, 1711-1716 (1993).
21. Sheinerman, F. B., Norel, R., Honig, B. *Curr. Opin. Struct. Biol. 10*, 153-159 (2000).

22. Norel, R., Sheinerman, F., Petrey, D., Honig, B. *Protein Sci. 10*, 2147-2161 (2001).
23. Northrup, S. H., Erickson, H. P. *Proc. Natl. Acad. Sci. USA 89*, 3338-3342 (1992).
24. Bogan, A. A. *J. Mol. Biol. 280*, 1-9 (1998).
25. Froloff, N., Windeman, A., Honig, B. *Protein Sci. 6*, 1293-1301 (1997).
26. Schapira, M., Totrov, M., Abagyan, R. *J. Mol. Recogn. 12*, 177-190 (1999).
27. Reddy, V. S., Giesing, H. A., Morton, R., Kumar, A., Post, C. B., Brooks, C. L. III, Johnson, J. E. *Biophys. J. 74*, 546-558 (1998).
28. Xu, D., Lin, S. L., Nussinov, R. *J. Mol. Biol. 265*, 68-84 (1997).
29. Xiao, L., Honig, B. *J. Mol. Biol. 289*, 1435-1444 (1999).
30. Li, A. J., Nussinov, R. *Proteins 32*, 111-127 (1998).
31. Mekler, L. B. *Biophisika* (in Russian) *14*, 581-584 (1970).
32. Bost, K. L., Smith, E. M., Blalock, J. E. *Proc. Natl. Acad. Sci. USA 82*, 1372-1375 (1985).
33. Drozdov-Tikhomirov, L. N., Linde, D. M., Poroikov, V. V., Alexandrov, A. A., Skurida, G. I. *J. Biomol. Struct. Dyn. 19*, 279-284 (2001).

Date Received: June 13, 2003

Communicated by the Editor Valery Ivanov