Use of an interface contact statistics to rescore protein-protein docked ensembles

Mihaly Mezei

Department of Pharmacological Sciences,
Icahn School of Medicine at Mount Sinai,
New York, NY 10029, USA.

E-Mail: Mihaly.Mezei@mssm.edu
Tel.: +1-212-659-5475

Keywords: Protein-protein interface, circular variance, residue contact propensity

Short title: Rescoring with interface statistics

Abstract

The recently developed statistical measure for the type of residue-residue contact at protein complex interfaces, based on a parameter-free definition of contact, has been used to define a contact score that is correlated with the likelihood of correctness of a proposed complex structure. Comparing the proposed contact scores on the native structure and on a set of model structures the proposed measure was shown to generally favor the native structure but in itself was not able to reliably score the native structure to be the best. Adjusting the scores of redocking experiments with the contact score showed that the adjusted score was able to move up the ranking of the native-like structure among the proposed complexes when the native-like was not ranked the best by the respective program. Tests on docking of unbound proteins compared the contact scores of the complexes with the contact score of the crystal structure again showing the tendency of the contact score to favor native-like conformations. The possibility of using the contact score to improve the determination of biological dimers in a crystal structure was also explored.

1. **Introduction and background.**

Proteins form an enormously varied ensemble of macromolecules, performing a wide variety of biological functions. In most cases these functions are executed by complexes of proteins. Therefore, molecular level knowledge of such associations is a prerequisite of understanding the mechanism of their actions. It turns out, however, that even when the structures of individual proteins are known the structure of the complex(es) formed is difficult to predict, as witnessed by the CAPRI (Critical Assessment of PRediction of Interactions) competition[1], held two to four times each year since 2001. The difficulty of such predictions is also reflected in the fact that most programs/servers return a set of putative complex structures, in most cases along with a score (or scores) assigned to each, instead of a single model.

Recent work, using a parameter-free definition of intermolecular contacts, showed that there is a wide variety in the propensities of contact between different residue types (about two orders of magnitude) in experimentally determined protein-protein interfaces[2]. The present paper examines the possibility of using the comparison of observed contacts with their respective propensities to help in selecting the conformation that is the most native like among an ensemble of putative complexes, typically generated by protein-protein docking

programs/servers. Given that the different docking algorithms use scoring functions that have vastly different source, it is expected that the possibility and extent of improvement depends on the scoring function used.

The contact statistics were developed on a set of 1172 protein complex structures, obtained from the Protein Data Bank (PDB)[3]. The contact propensity for residue (amino acid) pair (*i*,*j*), $PR_{i,j}$ was defined as

$$PR_{i,j} = \frac{N_{i,j}}{\sum_{i,j=1}^{20} N_{i,j}} / [P_i * P_j * (2 - \delta_{i,j})] \quad (1)$$

where $N_{i,j}$ is the number of (*i*,*j*) contacts in the data set and P_i, P_j are the propensities of residue types *i* and *j* to be on the surface of the protein (in the same data set) and $\delta_{i,j}$ is the Kronecker delta. Surface (heavy) atoms are defined as atoms with exposed VdW surface larger than 3% and circular variance[4] (calculated with respect to the rest of the protein atoms) less than 0.8. Atoms *i* and *j* are defined to be in contact if they are mutually proximal[5], *i.e.*, atom *i* of protein 1 is nearest to atom *j* of protein 2 AND atom *j* of protein 2 is nearest to atom *i* of protein 1. The $PR_{i,j}$ values found are in the range [0.21,17.2].

2. **Materials and Methods.**

Based on the fundamental logarithmic relation between free energy and probability two measures are proposed for quantifying the extent a proposed protein-protein

interface adheres to the contact statistics established in the previous study: *S* and its normalized variant, S/N_{ct} as follows.

$$S = \sum_{(i,j)} -kT \ \ln[PR_{i,j}] \quad (2)$$

$$S_N = S/N_{ct} \quad (3)$$

where the summation is over all contacts, *i* and *j* are the residue types of each contact and N_{ct} is the number of contacts in the complex. For *kT* the rounded value of 0.6 was used, corresponding to ambient temperature and units of kcal/mol.

The contact score *S* was used to define a corrected scores S_c as follows:

$$S_c = S_M - w * S \quad (4)$$

where S_M is the score returned by the program generating the ensemble and *w* is the correction weight that should primarily correct for the difference in units used for *S* and S_M. In the present work it was varied to find its optimal value.

3. **Results and Discussion.**

For the first two tests of usefulness of the contact scores *S* and S_N a set of 18 complexes of various sizes were selected from the PDB that were not used for the development of the contact propensities. Two components of these complexes, representing biological dimers according to the PDB annotation, were submitted to

the protein-protein docking servers ClusPro[6-9] and PatchDock[10,11]. Note that such redocking experiments are inadequate to test a given docking algorithm since the components submitted to the servers were in the binding conformation thus the docking is significantly easier than in a 'real life' docking problem where the monomer conformations are obtained without the knowledge of the complex conformation. However, for the present purpose this is an advantage since the better the result of the server, the harder it is to improve on it.

The PDB IDs, the chain IDs used, the number of residues in the two components, as well as the number of contacts in the crystal structure are shown in Table I. The number of putative structures returned by the servers generally varied. Since ClusPro gave at most 30 structures, the number of structures used from PatchDock was also limited to the top-scoring 30 (although in most cases PatchDock generated many more). S and S_N were calculated for all model structures, as well as for the crystal structure (also referred to as native).

The first test aimed at measuring the extent that either of the contact scores in itself is diagnostic of the accuracy of a model. It consisted of counting the number of structures whose scores were better that the score of the x-ray structure. If the calculated scores are good measures of the accuracy of the model than the answer should be zero while for the measure to be a complete failure the answer should

fluctuate around half the number of models considered. Besides the calculated number of models beating the x-ray score, the table also gives the number of residues in the complexes and the RMSD between the model and the crystal structures (calculated by overlaying the first protein of the model to the crystal structure's first protein and calculating the RMSD between the two structures for the second protein) and the number of contacts.

The results, presented in Tables II and III for ClusPro and PatchDock, resp., show that both of our measures are strongly correlated with the accuracy of the model - for several complexes the crystal structure's score is the best. However, they are not the best for all, although there are very few complexes where close to 50% of the models 'beat' the crystal structure. One clear conclusion did emerge: S performed significantly better than S_N, especially for ClusPro. This implies that the number of contacts is also correlated with the accuracy of a model.

Tables II and III also contain the number of models whose contact score is better than the contact score of the best model. Furthermore, the RMSD between the top-scoring model and the x-ray structure (based on the C_α atoms) is also shown to quantify how close to the native the best model is. Use of the contact score is only expected to include the ranking when the best model is native like.

For the study of the ability of contact scores to improve the accuracy of the scores of putative complexes the scores returned by the server, S_M , the modified score S_c was calculated with different correction weights w. S_M was obtained from ClusPro as the "Lowest Energy" value and for PatchDock as the negative of the score given. Note that in this work we did not look for the optimal combination of the scores returned by the servers since the aim of this work was to show that the score accuracy can be increased by our proposed correction. Therefore, our results should not be considered as a test of the servers' accuracy, even though both performed rather well.

Tables IV and V show the rank of the complexes found to be closest to the crystal structure together with the number of structures that have better modified scores S_c calculated with different correction weights w. The reason for eschewing the customary comparison of enrichments (ROC curves) is that once a model differs significantly from the native structure the calculated score and RMSD is not expected to have any relation whatsoever. For the same reason, results for the complexes where the lowest RMSD was large (above 10 Å) will not be included in the discussion.

For complexes where the structure with the lowest RMSD had also the lowest score S_M (*i.e.*, there was no room for improvement) the modified scores S_c were

still scoring the lowest, with the exception of two structures using w=10 or larger. For the few cases where the best structure was close to the crystal structure but did not have the lowest score the modified scores did improve the ranking. For ClusPro a compromise weight of 10 is suggested since for w=10 and 20 some complexes showed improvement while some showed a slight decline. For PatchDock, however, even w=200 yielded improvement and (ignoring the cases where the best model was far from the native) in no case did the ranking become worse. For each other servers/software its value should be established individually. Note, that the rescoring had significantly more effect on the PatchDock runs than on the ClusPro runs.

The third test used 20 and 25 unbound monomers structure pairs (not used for the development of contact statistics) from the DOCKGROUND[12] and ZLAB[13] datasets, resp. and were selected and submitted to ClusPro, PatchDock and Gramm-X[14,15]. Since few of the runs produced native-like complexes (defined as RMSD < 10Å) this test first compared the contact scores S and S_N of the x-ray complex and of the generated models. For PatchDock and Gramm-X the top 30 complexes were considered and for ClusPro all models that were generated by the server ($\leq$30).

Table VI and Table VII give the result of the comparison of the contact scores of models and of the corresponding x-ray structure for the complexes from the DOCKGROUND and ZLAB benchmark set, resp. The tables show for each structure the PDB ids of the complexes, the PDB ids of the unbound monomers, the number of residues in the two components, and the number of contacts in the x-ray structure; for ClusPro the number of models generated was also shown. For the models generated by the three servers used the number of models whose score was better than the x-ray structure's score was given. The better the contact score represents the goodness of a model, the smaller is this number.

For most complexes generated by PatchDock and GRAMM-X few, if any, models showed better contact scores than the x-ray structure. However, for about three fourth of the complexes the majority of ClusPro generated models have better contact scores that the corresponding x-ray structure. This implies that the contact scores are less likely to improve the ClusPro rankig of the native structure (if found among the models generated). The software-dependence of the usefulness of the contact score thus implies that for a docking software not used here these tests should be repeated – they would be needed also for the determining of the optimal weight anyway.

There was one complex (1SBB) where the contact scores compared with the x-ray score uniformly badly over the three softwares used. This lead to the idea of questioning the correctness of the experimental complex. This is not as provocative as it sounds since the biological dimer conformation is selected by the software PISA[16] from the several possible pairings in the unit cell and PISA does not claim an accuracy of 100%. To test this, the full unit cell was generated using Simulaid[17] and an other conformation was chosen as the putative biological dimer. This resulted in a better contact score : -1.51 instead of -2.60 (the usual range of the contact scores is 10). Also, the results, shown in the last row of Table VI with PDB ID 1SBBx, improved for all three programs, albeit not by too much. This suggests that the contact score can also be used to help in determining the biological dimers from a crystal structure, but further studies are required to confirm and quantify this proposition.

The fourth test looked at the complexes where the model set generated by docking the unbound complexes included a native-like model and tested the ability of the contacts score to include their rank. 10 such sets were found among the model sets generated by ClusPro and PatchDock – GRAMM-X was not used in this test since the server does not provide a score. Table VIII provides the comparison of the change in ranking with the contact score corrections. ‘Unfortunately’, most native-

like models were found by ClusPro (where the contact score did not perform well on the score comparison test described above). However, for most complexes the contact score still provided improvement or was neutral; only for one complex was the ranking worsened significantly by the contact score comparison.

In closing it is to be emphasized again that the docking results presented here should not be considered to be a comparison of the servers' performance. All dockings were run using default parameters and no attempt was made to optimize their performance.

Conclusion

It has been shown that incorporation of a correction based on the recently developed interface contact statistics offers a way to improve the ranking of the native-like protein-protein complex model structure among the models generated. The possibility of using the contact statistics to improve the success rate of predicting biological dimers in a crystal structure is also explored. The extent of possible improvement depends on the software used to generate the model ensemble. Furthermore, the optimal scaling of the correction has to be established as well; the present work provides suggestion for ClusPro and PatchDock.

Acknowledgments

This work was supported in part through the computational resources and staff expertise provided by the Department of Scientific Computing at the Icahn School of Medicine at Mount Sinai.

Conflicts of Interest

The author declares no conflict of interest.

Table I: Description of the complexes used for the redocking tests.

PDB ID	chain IDs	# of residues		# of models	# of contacts
4ODS	H L	210	214	30	67
4ONL	A B	140	149	19	25
4POZ	C D	211	215	30	64
2QKO	A B	87	131	24	24
4QVF	A B	141	21	7	29
4UHP	E F	132	94	16	30
4X7S	H L	222	218	25	61
4YII	V A	145	72	20	24
4YON	U A	349	176	30	39
4Z95	H L	212	214	25	63
4GUZ	A D	284	284	24	30
4I4N	A B	281	281	15	38
4OFW	A C	387	387	30	38
4PGG	A B	360	360	30	145
4PVC	A B	342	342	30	38
4R1N	A B	282	282	26	47
4WOY	A B	329	329	30	7
4WUM	A B	389	389	30	86

Table II: Contact score comparison between redocked models generated by ClusPro and the crystal structure and the native-like model

PDB ID	# of models	# of models beating the crystal S score		# of models beating the crystal S_N score		#of models beating the native-like S score	RMSD of the best model
4ODS	30	0	0%	4	13%	4	2.0
4ONL	19	4	21%	4	21%	4	2.8
4POZ	30	1	3%	12	40%	0	4.3
2QKO	24	4	16%	3	13%	15	3.6
4QVF	7	3	42%	4	57%	3	4.7
4UHP	16	5	31%	7	43%	0	4.8
4X7S	25	0	0%	7	28%	0	4.6
4YII	20	1	5%	4	20%	4	6.0
4YON	30	10	33%	10	33%	0	4.0
4Z95	25	0	0%	13	52%	13	5.1
4GUZ	24	11	45%	7	29%	2	3.4
4I4N	15	6	40%	8	53%	0	4.7
4OFW	30	28	93%	27	90%	12	8.6
4PGG	30	0	0%	0	0%	1	6.0
4PVC	30	5	16%	5	16%	6	5.8
4R1N	26	24	92%	22	84%	7	5.6
4WOY	30	21	70%	30	100%	19	49.6
4WUM	30	30	100%	29	96%	9	4.5

Table III: Contact score comparison between redocked models generated by PatchDock and the crystal structure and the native-like model

PDB ID	# of models	# of models beating the crystal S score		# of models beating the crystal S_N score		#of models beating the native-like S score	RMSD of the best model
4ODS	30	0	0%	0	0%	3	2.2
4ONL	30	0	0%	0	0%	23	33.4
4POZ	30	1	3%	0	0%	0	1.7
2QKO	30	0	0%	0	0%	2	3.8
4QVF	30	3	10%	5	16%	0	1.4
4UHP	30	1	3%	1	3%	0	1.8
4X7S	30	2	0%	1	3%	1	2.2
4YII	30	0	0%	1	3%	0	1.0
4YON	30	1	3%	1	3%	0	1.7
4Z95	30	1	3%	2	6%	0	2.1
4GUZ	30	3	10%	0	0%	4	12.7
4I4N	30	0	0%	0	0%	4	19.9
4OFW	30	3	10%	3	10%	1	18.8
4PGG	10	0	0%	0	0%	0	0.5
4PVC	30	0	0%	0	0%	1	10.2
4R1N	30	11	36%	7	23%	4	1.6
4WOY	30	3	10%	29	96%	0	40.9
4WUM	30	14	46%	9	30%	6	2.2

Table IV: Rescoring results for redocked models generated by ClusPro

PDB ID	Best RMSD/Å	Rank of best RMSD	# of models beating the model with the best RMSD using the correction factor below					
			0.0	1.0	2.0	5.0	10.0	20.0
4ODS	2.0	1	0	0	0	0	0	0
4ONL	2.8	1	0	0	0	0	0	0
4POZ	4.3	1	0	0	0	0	0	0
2QKO	3.6	1	0	0	0	0	1	1
4QVF	4.7	1	0	0	0	0	0	0
4UHP	4.8	1	0	0	0	0	0	0
4X7S	4.6	8	7	5	4	3	2	2
4YII	6.0	1	0	0	0	0	0	0
4YON	4.0	1	0	0	0	0	0	0
4Z95	5.1	1	0	0	0	0	1	1
4GUZ	3.4	9	8	7	7	7	7	7
4I4N	4.7	1	0	0	0	0	0	0
4OFW	8.6	16	15	15	15	14	14	15
4PGG	6.0	1	0	0	0	0	0	0
4PVC	5.8	4	3	3	3	3	3	3
4R1N	5.6	1	0	0	0	0	0	0
4WOY	49.6	29	28	28	29	29	28	26
4WUM	4.5	1	0	0	0	0	0	0

Table V: Rescoring results for redocked models generated by PatchDock

PDB ID	Best RMSD/Å	Rank of the best RMSD	# of models beating the model with the best RMSD using the correction factor below						
			0.0	1.0	5.0	10.0	50.0	100.0	200.0
4ODS	2.2	1	0	0	0	0	0	0	0
4ONL	33.4	24	23	23	23	23	24	26	27
4POZ	1.7	1	0	0	0	0	0	0	0
4QKO	3.8	17	16	16	16	16	12	8	4
4QVF	1.4	1	0	0	0	0	0	0	0
4UHP	1.8	15	14	14	13	12	6	2	1
4X7S	2.2	1	0	0	0	0	0	0	0
4YII	1.0	5	4	4	4	4	1	1	0
4YON	1.7	2	1	0	0	0	0	0	0
4Z95	2.1	1	0	0	0	0	0	0	0
4GUZ	12.7	21	20	20	21	21	20	18	12
4I4N	19.9	23	21	21	21	21	19	17	14
4OFW	18.8	24	23	23	22	22	17	10	5
4PGG	0.5	1	4	0	0	0	0	0	0
4PVC	10.2	1	0	0	0	0	0	0	0
4R1N	1.6	1	0	0	0	0	0	0	0
4WOY	40.9	10	9	9	9	9	7	6	4
4WUM	2.2	1	0	0	0	0	0	0	0

				# of models beating the contact score of the X-ray score						
				ClusPro 2.0			PatchDock		Gramm-X	
PDB ID	# of residues		nCT	nM	(S)	(S_N)	(S)	(S_N)	(S)	(S_N)
1A2K (1GY6:3RAN)	442	246	19	27	2	0	0	0	0	0
1A2Y (1VFA:3LZT)	353	224	21	21	20	20	3	3	2	2
1AKJ (1I4F:1CD8)	601	373	36	30	26	26	3	4	15	13
1CHO (1K2I:2OVO)	292	236	12	15	8	5	2	0	2	1
1DE4 (1A6Z:1CX8)	1649	371	31	30	17	13	0	0	0	0
1G20 (1L5H:1FP6)	993	608	65	30	27	26	2	1	2	2
1G4A (1DO2:1HT1)	1510	814	27	30	28	30	1	2	1	3
1GPQ (1XS0:3LZT)	254	126	27	23	21	21	4	4	7	8
1N8O (1GG6:1IFG)	365	228	29	16	4	4	0	0	1	1
1OMW (1YM7:1XHM)	993	608	31	30	5	5	0	0	0	0
1RLB (1F86:1KT3)	405	230	17	30	12	18	2	4	5	6
1SBB (1BEC:3SEB)	475	237	13	30	30	30	10	29	18	30
1U0N (1IJK:1P9A)	717	451	41	30	21	20	1	1	6	5
1UEX (1JWI:1IJB)	446	245	25	21	12	9	0	0	0	0
2ATQ (1IH7:2A1K)	1112	897	8	30	26	30	1	29	4	30
2B4S (1F71:1P14)	584	297	23	23	15	19	2	4	4	4
2D26 (1QLP:1QNJ)	601	372	14	30	29	30	9	26	12	25
2G45 (2G43:1YJ1)	187	117	17	15	0	0	0	0	0	0
2GOO (1REU:1BTE)	195	103	19	21	20	20	13	9	11	10
3SIC (1SUP:2SSI)	383	275	25	30	22	21	6	6	6	4
1SBBx(1BEC:3SEB	475	237	17	30	29	30	8	14	15	20

)										

Table VI: Comparison of the contact scores of models generated from unbound monomer structures in the DOCKGROUND dataset, with the corresponding contact score of the dimer X-ray structure.

Table VII: Comparison of the contact scores of models generated from unbound monomer structures in the ZLAB dataset, with the corresponding contact score of the dimer X-ray structure.

				# of models beating the contact score of the X-ray score						
				ClusPro 2.0			PatchDock		Gramm-X	
PDB ID	# of residues		nCT	nM	(S)	(S_N)	(S)	(S_N)	(S)	(S_N)
1AHW (1FGN:1TFH)	428	200	24	30	25	27	2	7	4	9
1BVK (1BVL:3LZT)	224	129	20	20	20	20	0	0	3	3
1D6R (2TGT:1K9B)	220	58	21	25	11	11	11	12	10	11
1DQJ (1DQQ:3LZT)	424	129	27	29	21	21	0	0	2	2
1E6E (1E1N:1CJE)	457	113	27	29	23	25	5	14	9	12
1E6J (1E6O:1A43)	429	210	16	12	11	11	0	0	2	2
1HIA (2PKA:1BX8)	223	48	28	17	4	7	0	0	0	2
1JPS (1JPT:1TFH)	426	200	29	21	20	13	1	2	3	2
1MAH (1J06:1FSC)	533	61	28	23	4	4	0	0	0	0
1MLC (1MLB:3LZT)	432	129	27	8	11	26	0	0	2	2
1VFB (1VFA:8LYZ)	223	129	22	15	15	30	4	4	5	5
1WEJ (1QBL:1HRC)	437	104	18	18	18	21	0	0	-	-
2FD6 (2FAT:1YWH)	420	248	20	30	4	2	0	0	0	0
2I25 (2I24:3LZT)	114	129	25	20	0	3	0	0	0	0
2MTA (2BBK:2RAC)	498	105	22	17	3	2	0	0	1	1
2UUY (1HJ9:2UUX)	223	52	18	27	4	2	1	0	3	1
2VIS (1GIG:2VIU)	431	267	22	30	0	0	0	0	0	0

2VXT (2VXU:1JOS)	416	156	35	27	8	11	0	0	0	0
2W9E (2W9D:1QM1)	427	99	25	28	4	4	0	0	0	0
3HMX (3HMW:1F45)	726	176	27	30	4	6	1	0	0	0
3MXW (3MXV:3M1N)	426	153	26	30	4	3	0	0	0	0
3RVW (3RVT:3F5V)	429	222	25	29	1	1	9	0	0	0
4DN4 (4DN3:1DOL)	427	61	17	24	4	4	0	0	0	0
4FQI (4FQH:2FK0)	1716	176	165	25	9	9	0	0	0	0
4G6J (4G5Z:4I1B)	430	149	32	27	8	10	0	0	1	1

Table VII: Rescoring results for unbound docking ensembles

PDB ID	Best model		Software	# of models beating the model with the best RMSD using the correction factor below								
	RMSD	Rank		0.0	1.0	2.0	5.0	10.0	20.0	50.0	100.0	200.0
1DE4	5.5	8	ClusPro	7	7	7	7	7	7	8		
1E6E	9.2	4	ClusPro	3	2	2	2	2	1	0		
1E6J	6.0	17	ClusPro	16	17	16	16	14	14	14		
1HIA	7.8	16	ClusPro	15	15	15	15	16	16	15		
1HIA	8.5	3	PatchDock	2	0	0		0		0	0	0
1MAH	7.9	2	ClusPro	1	1	1	1	2	3	8		
1MLC	5.4	19	ClusPro	18	19	19	23	24	23	22		
1N8O	10.0	7	ClusPro	6	6	6	5	6	7	6		
3MXW	2.2	6	ClusPro	5	4	3	1	1	1	1		
3SIC	5.7	2	ClusPro	1	1	0	0	0	0	0		

References

1. Janin J. Welcome to CAPRI: A Critical Assessment of PRedicted Interactions Proteins: Structure, Function, and Genetics 2002;47(3):257.
2. Mezei M. Statistical properties of protein-protein interfaces. Algorithms 2015;8:92-99.
3. Berman H, Henrick K, Nakamura H. Announcing the worldwide Protein Data Bank. Nature structural biology 2003;10(12):980.
4. Mezei M. A new method for mapping macromolecular topography. Journal of molecular graphics & modelling 2003;21(5):463-472.
5. Mezei M, Zhou M-M. Dockres: a computer program that analyzes the output of virtual screening of small molecules. Source Code for Biology and Medicine 2010;5:2.
6. Kozakov D, Beglov D, Bohnuud T, Mottarella S, Xia B, Hall DR, Vajda S. How good is automated protein docking? Proteins: Structure, Function, and Bioinformatics 2013;81:2159-2166.
7. Kozakov D, Brenke R, Comeau SR, Vajda S. PIPER: An FFT-based protein docking program with pairwise potentials. Proteins 2006;65:392-406.
8. S.R.Comeau, Gatchell DW, Vajda S, Camacho CJ. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics 2004;20:45-50.
9. S.R.Comeau, Gatchell DW, Vajda S, Camacho CJ. ClusPro: a fully automated algorithm for protein-protein docking. Nucleic Acids Research 2004;32:W96-W99.
10. Duhovny D, Nussinov R, Wolfson HJ. Efficient Unbound Docking of Rigid Molecules. In: al. Ge, editor. Lecture Notes in Computer Science; 2002; Rome, Italy. Springer Verlag. p 185-200. (Lecture Notes in Computer Science).
11. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. Nucl Acids Res 2005;33:W363-367.
12. T. Kirys, A.M. Ruvinsky, D. Singla, A.V.Tuzikov, P.J. Kundrotas, Vakser IA. Simulated unbound structures for benchmarking of protein docking in the DOCKGROUND resource. BMC Bioinformatics 2015;16:243.
13. Vreven T, Moal I, Vangone A, Pierce B, Kastritis P, Torchala M, Chaleil R, Jimenez-Garcia B, Bates P, Fernandez-Recio J, Bonvin A, Weng Z. Updates to the integrated protein-protein interaction benchmarks: Docking benchmark version 5 and affinity benchmark version 2. J Mol Biol 2015;427:3031-3041.
14. Tovchigrechko A, Vakser IA. Development and testing of an automated approach to protein docking. Proteins 2005;60:296-301.
15. A. Tovchigrechko, Vakser. IA. GRAMM-X public web server for protein-protein docking. Nucleic Acids Res 2006;34:W310-314.
16. E. Krissinel, Henrick K. Inference of macromolecular assemblies from crystalline state. J Mol Biol 2007;372:774--797.
17. Mezei M. Simulaid: a simulation facilitator and analysis program. J Comp Chem 2010;31:2658-2668.