bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Prediction of residue-residue contacts in CASP12 targets from its predicted tertiary structures

Piyush Agrawal[1], Sandeep Singh[1], Gandharva Nagpal[1], Deepti Sethi[1], Gajendra P.S. Raghava*[,1,2]

1. CSIR-Institute of Microbial Technology, Chandigarh 160036, India.

2. Indraprastha Institute of Information Technology, New Delhi 110020, India.

Emails of Authors:

PA: piyush_11@imtech.res.in

SS: sandybioteck@gmail.com

GN: gnagpal@imtech.res.in

DS: sethideepti120@gmail.com

*** Corresponding author**

Professor, Centre for Computational Biology, Indraprastha Institute of Information Technology,

Okhla Industrial Estate, Phase III, New Delhi 110020,

India. Tel.: +91 011 26907444

E-mail address: raghava@iiitd.ac.in

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

One of the challenges in the field of structural proteomics is to predict residue-residue contacts in a protein. It is an integral part of CASP competitions due to its importance in the field of structural biology. This manuscript describes RRCPred 2.0 a method participated in CASP12 and predicted residue-residue contact in targets with high precision. In this approach, firstly 150 predicted protein structures were obtained from CASP12 Stage 2 tarball and ranked using clustering-based quality assessment software. Secondly, residue-residue contacts were assigned in top 10 protein structures based on distance between residues. Finally, residue-residue contacts were predicted in target protein based on consensus/average in top 10 predicted structures. This simple approach performs better than most of CASP12 methods in the categories of TBM and TBM/FM. It ranked 1st in following categories; i) TBM domain on list size L/5, ii) TBM/FM domain on list size L/5 and iii) TBM/FM domain on Top 10. These observations indicate that predicted tertiary structure of a protein can be used for predicting residue-residue contacts in protein with high accuracy.

Keywords: Residue-residue contact; CASP; Quality Assessment; Protein Structures; Clustering Approach.

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Numbers of methods have been developed in the past for predicting three-dimension **(**3D) structure of a protein from its primary structure or amino acid sequence. Broadly, these approaches can be classified in three classes; i) template or homology based approach, ii) threading based approach and iii) *ab initio* method[1]. In recent years, *ab initio* method is gaining importance in the field of structure prediction since traditional method of template based structure prediction is limited by the number of structural templates in the Protein Data Bank (RCSB-PDB)[2]. However, recent studies showed residue-residue contact map guided structure prediction, as an important intermediate step in *ab initio* structure prediction of a protein[3,4]. Apart from structure prediction, residue-residue contact predictions has also shown promising application in the field of drug designing[5], model ranking, selection and its evaluation[6,7]. Although the importance of residue-residue contact prediction in modeling tertiary structure of a protein was introduced decades ago[8,9], the thorough implementation of the idea came into practice when several groups across the world have shown how residues in contact can be predicted with improved and reasonable accuracy[10,11]. One of the main reasons of development in this field is introduction of a separate section of residue-residue contact prediction in CASP (Critical Assessment of Techniques for Protein Structure Prediction). This section was introduced in the CASP2[12] but it was fully standardized during CASP6-CASP9[13-16]. Definition of residue-residue contact has not been changed since previous CASP rounds. Two residues are said to be in contact when distance between their Cβ atoms (Cα in case of glycine) is shorter than 8.0 Å. These contacts are also classified in three different categories: short, medium and long range contacts based on separation along the sequence. Short range contacts are those contacts where residues are separated by 6-11 residues, Medium range contacts are those where residues are separated by 12-23 residues whereas Long range contacts are those where residues are separated by minimum 24 residues or more than that[17]. Among all the contacts, 'long range contact' prediction is of much interest and more valuable for structure prediction[18]. These contacts are

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

evaluated using precision i.e. number of contacts correctly predicted out of all the predicted contacts[19]

Current existing methods on residue-residue contact prediction can be classified into two categories. First is sequence based method, where no information other than sequence of a protein is present. Contacts in the protein are predicted utilizing information which can be derived directly from its sequence. Number of methods have been developed utilizing various machine-learning techniques such as support vector machines[20,21], hidden markov models[22], neural networks[23,24], deep network and boosting techniques[25], evolutionary information[26,27], covariation signal information[28], etc. Second class is template/structure-based methods where structure is used for extracting contact information. In this method, firstly structure of protein is modeled using best available template and then contacts are predicted from the modeled structure[21].

As reported in literature, prediction of contacts depends very much on the quality of the template structure[21]. Therefore, keeping this in mind, in the present study, we have developed a simple approach to predict residue-residue contact from high quality tertiary structure predicted by various methods participated in structure prediction category of CASP12. We ranked all predicted structure and used top 10 structures for predicting residue-residue contact using consensus or average based approach.

Results and Discussion

A large number of groups participated in CASP12 residue-residue contact prediction category, where they submitted their predicted models to CASP12 site (http://predictioncenter.org/casp12/doc/CASP12_Abstracts.pdf). In this paper, we analyzed and discussed performance of models manually submitted by our group “raghavagps”. Our prediction models were derived from Stage 2 tarballs released by CASP12 sites. These tarballs comprises models submitted by online servers to CASP12 site (See Materials & Methods). In this CASP12, we have analyzed the results present at the site (http://predictioncenter.org/casp12/rrc_results.cgi).

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Here, our result was evaluated on 66 targets out of total 94 targets for 'long' and 'long+medium' range contacts by the CASP team. Out of our 66 targets, 38 targets were present in Free Modeling (FM) category, 12 in Template Based Modeling (TBM) category and 16 in Template Based Modeling/Free Modeling (TBM/FM) category.

Prediction of 'long' range Contacts

Firstly, we analyzed the prediction performance of all groups in predicting 'long' range contacts on targets of TBM domain category (Table 1). The performance of models were extracted from CASP12 sites using probability >0 on list size L/5 (Supplementary Fig. S1). As shown in Table 1, method RRCPred 2.0 (raghavagps) got average precision value of 85.028%, which is better than other methods in TBM category. This method RRCPred 2.0 (raghavagps) was followed by "naive" and "FLOUDAS_SERVER", with 2nd and 3rd position respectively. Similarly, performance was extracted from CASP site for Top10 contacts for all groups on targets of TBM category. On Top 10 contacts method "naive" got average precision value of 86.486%, which is better than other methods (Table 1). As shown in Table 1, method "naïve" was followed by "RaptorX-Contact" and RRCPred 2.0 (raghavagps) with 2nd and 3rd rank respectively (Supplementary Fig. S2).

Table 1: The performance or precision of different methods for 'long' range contacts on CASP12 TBM targets.

Methods/groups	**L/5***		**Top10**	
	Precision (%)	**Rank**	**Precision (%)**	**Rank**
RRCPred 2.0 (raghavagps)	85.028	1	83.333	3
Naïve	80.247	2	86.486	1
FLOUDAS_SERVER	77.321	3	77.941	6
Deepfold-Contact	76.801	4	82.432	4

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RaptorX-Contact	76.745	5	83.784	2
Distill	75.051	6	77.568	7
iFold_1	72.105	7	77.500	8
Yang-Server	72.088	8	79.444	5
MetaPSICOV	71.582	9	76.571	9
PconsC31	69.672	10	75.405	10

* where L is the sequence length of the protein.

In addition to TBM targets, we also extracted the performance of different methods on targets falling in TBM/FM category using probability value >0. The precision of predicting 'long' range contact of different methods for L/5 is shown in Table 2. In this category (TBM/FM) for L/5, method RRCPred 2.0 (raghavagps) achieved precision 72.244% with recall 16.466%, which is better than other methods. This method is followed by “iFold_1” and “MetaPSICOV” with precision 58.117% and 53.755% respectively (Supplementary Fig. S3). Similar trend was observed for Top10 contacts where method RRCPred 2.0 (raghavagps) achieved precision 73.125%, followed by “iFold_1” and “PconsC31” which got precision 62.353% and 61.579% respectively (Supplementary Fig. S4). It is important to note that performance of all methods/groups decreased substantially for TBM/FM targets in comparison to TBM targets.

Most challenging targets in CASP competition are targets of FM category. We extracted the performance of all methods from CASP site for FM targets. As shown in Table 3, method “RaptorX-Contact” with average precision of 47.095% in L/5 list size outperformed other methods. The performance of all methods was very poor for FM category targets. The method “RaptorX-Contact” was followed by “iFold_1” and “MetaPSICOV” who scored 2nd and 3rd rank respectively. Unfortunately, our method RRCPred 2.0 (raghavagps) performed very poor on FM targets and achieved precision value of 32.883% and ranked 17th (Supplementary Fig. S5). We also analysed the performance of different methods for Top10 contacts on FM targets and observed similar trends

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Supplementary Fig. S6). This means method RRCPred 2.0 (raghavagps) is not suitable for FM category targets.

Table 2. The performance or precision of different methods for 'long' range contacts on CASP12 TBM/FM targets.

Methods	**L/5***		**Top10**	
	Precision (%)	**Rank**	**Precision (%)**	**Rank**
RRCPred 2.0 (raghavagps)	72.244	1	73.125	1
iFold_1	58.117	2	62.353	2
MetaPSICOV	53.755	3	59.474	4
PconsC31	52.342	4	61.579	3
Naive	51.988	5	54.211	7
Deepfold-Contact	51.026	6	53.684	8
myprotein-me	47.892	7	55.000	6
RaptorX-Contact	46.579	8	47.368	12
RBO-Epsilon	46.286	9	57.368	5
Pcons-net	45.721	10	50.526	10

* where L is the sequence length of the protein.

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 3. The performance or precision of different methods for 'long' range contacts on CASP12 FM targets.

Methods	**L/5***		**Top10**	
	Precision (%)	**Rank**	**Precision (%)**	**Rank**
RaptorX-Contact	47.095	1	54.211	1
iFold_1	46.864	2	50.270	2
Deepfold-Contact	45.549	3	48.889	4
MetaPSICOV	44.007	4	47.368	6
MULTICOM-CLUSTER	41.691	5	48.947	3
MULTICOM-CUNSTRUCT	41.637	6	48.421	5
Naive	41.388	7	45.556	8
Pcons-net	41.013	8	46.579	7
IGBteam	38.784	9	43.421	9
Zhang_Contact	37.801	10	40.833	13
RRCPred 2.0 (raghavagps)	32.883	17	37.632	18

* where L is the sequence length of the protein.

When we selected TBM and TBM/FM options collectively, we observed that our method ranked 1st again in both the list size L/5 as well and Top10 with precision of 77.723% and 77.500% respectively (Supplementary Fig. S7 and S8 respectively). In this section, method “naïve” secured 2nd position. When domain TBM/FM and FM selected collectively, our method performed little low in comparison to other methods and ranked 6th with average precision of 44.546% in L/5 list size (Supplementary Fig. S9) and 8th with average precision of 48.148% in Top10 list size (Supplementary Fig. S10). Method “iFold_1” secured 1st position in this section with average precision of >50.00% in both the list size category. In the case of TBM and FM domain, our method

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

didn't performed so well and showed average precision of 45.398% and 48.600% in L/5 and Top10 list size respectively (Supplementary Fig. S11 and S12 respectively). Method "RaptorX-Conatct" outperformed other methods in this section. When all the three domains were selected at a given time and performance was evaluated, our methods showed descent performance with average precision of 51.906% in L/5 list size (Supplementary Fig. S13) and 54.545% in Top10 list size (Supplementary Fig. S14). In this section method "Deepfold-Contact" ranked 1st in L/5 list size category with average precision of 59.249% whereas method "RaptorX-Conatct" ranked 1st in L/5 list size category with average precision of 64.468%.

Prediction of 'long+medium' range contacts

We extracted the performance of methods on 'long+medium' range contacts from CASP12 site for TBM targets using probability >0 on list size L/5. As shown in Table 4, our method RRCPred 2.0 (raghavagps) achieved average precision of 89.101 and stood 1st in the list (Supplementary Fig. S15), it was followed by "naive" and "RaptorX-Contact" at 2nd and 3rd rank respectively. In case of Top10 list size "naive" got 1st rank with average precision 91.892% and was followed by RRCPred 2.0 (raghavagps) and "RaptorX-Contact" with average precision of 89.167% and 88.378% at 2nd and 3rd rank respectively (Supplementary Fig. S16).

Table 4. The performance or precision of different methods for 'long+medium' range contacts on CASP12 TBM targets.

Methods	**L/5***		**Top10**	
	Precision	**Rank**	**Precision**	**Rank**

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

	(%)		**(%)**	
RRCPred 2.0 (raghavagps)	89.101	1	89.167	2
Naïve	85.750	2	91.892	1
RaptorX-Contact	84.229	3	88.378	3
iFold_1	84.087	4	88.333	4
Deepfold-Contact	82.462	5	86.486	6
Yang-Server	82.362	6	87.222	5
FLOUDAS_SERVER	80.448	7	80.294	11
MetaPSICOV	79.437	8	84.000	7
Distill	79.149	9	81.351	9
Pcons-net	74.998	10	83.243	8

*** where L is the sequence length of the protein.**

In addition to TBM targets, we also analyzed the performance of different methods on TBM/FM targets for list size L/5 and Top10 and probability >0 (Table 5). In this category (TBM/FM), RRCPred 2.0 (raghavagps) obtained average precision of 78.358% for L/5 list size which is better than other methods (Supplementary Fig. S17). "MetaPSICOV" was present at 2nd position and "RaptorX-Contact" at 3rd position for L/5 list size. Similar trend was observed for Top10 list size where RRCPred 2.0 (raghavagps) achieved precision of 79.375% (Supplementary Fig. S18) and performed better than other methods.

In case of FM domain targets as shown in Table 6, our method didn't performed well in both list size L/5 (Supplementary Fig. S19) and Top10 list size (Supplementary Fig. S20) categories. Here, method "RaptorX-Contact" outperformed other methods with average precision of 55.831% in L/5 list size and 63.158% in Top10 list size.

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 5. The performance or precision of different methods for 'long+medium' range contacts on CASP12 TBM/FM targets.

Methods	**L/5***		**Top10**	
	Precision (%)	**Rank**	**Precision (%)**	**Rank**
RRCPred 2.0 (raghavagps)	78.378	1	79.375	1
MetaPSICOV	69.109	2	74.737	2
RaptorX-Contact	65.995	3	72.632	5
iFold_1	65.896	4	73.529	3
Deepfold-Contact	65.461	5	72.632	4
Naive	64.436	6	71.053	7
RBO-Epsilon	62.587	7	71.579	6
PconsC31	61.171	8	70.526	8
MULTICOM-CONSTRUCT	61.046	9	68.421	10
Pcons-net	60.756	10	69.474	9

* *where L is the sequence length of the protein.*

Table 6. The performance or precision of different methods for 'long+medium' range contacts on CASP12 FM targets.

Methods	**L/5***		**Top10**	
	Precision (%)	**Rank**	**Precision (%)**	**Rank**

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RaptorX-Contact	55.831	1	63.158	1
iFold_1	52.292	2	57.027	5
MetaPSICOV	51.307	3	58.684	3
MULTICOM-CUNSTRUCT	50.449	4	58.947	2
Pcons-net	49.536	5	55.000	6
Deepfold-Contact	49.023	6	54.444	7
RBO-Epsilon	48.601	7	57.632	4
Yang-Server	47.715	8	52.973	9
FALCON_COLORS	47.253	9	52.895	10
PconsC31	45.728	10	50.256	12
RRCPred 2.0 (raghavagps)	40.269	17	46.842	16

* where L is the sequence length of the protein.

In case of TBM and TBM/FM domain selection, our method showed highest average precision of 82.974% when list size L/5 was selected (Supplementary Fig. S21), followed by method "naive" and "RaptorX-Contact" at 2nd and 3rd ppsition respectively. In list size Top10, method "naive" ranked 1st with average precision of 84.821% and our method RRCPred 2.0 (raghavagps) ranked second with average precision value of 83.571% (Supplementary Fig. S22). When the domain selection TBM and TBM/FM was changed to TBM/FM and FM, our method perform little poor in comparison to other methods and ranked 8th in L/5 list size (Supplementary Fig. S23) whereas 11th in Top10 list size (Supplementary Fig. S24). Method "RaptorX-Contact" outperformed other methods in both list size. In the case of TBM and FM domain, our method showed average precision value of 51.989% and 57.000% in L/5 (Supplementary Fig. S25) and Top10 list size (Supplementary Fig. S26). "RaptorX-Contact" was found to be at the top with average precision of 69.840% in L/5 and 75.600% in Top10 list size. When all the three domains were selected, our method showed reasonable performance of average precision value of 58.386% (Supplementary Fig. S27) in L/5 list size and 62.424% in Top10 list size (Supplementary Fig. S28).

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Target wise analysis

In addition to average performance of methods, we analyzed the performance of our method target wise. The performance for each target using different options were extracted from CASP12 site using probability >0. As shown in Table 7, for a large number of targets our method got 100% precision. In case of 'long' range contacts for L/5, "raghavagps" achieved more than 70% precision for 28 targets that includes 4 targets with 100% precision (Supplementary Table S1). This performance improved for Top10 contacts, where our method predicts 10 targets with 100% precision (Supplementary Table S2). This shows utility of our method in real life. As expected, performance on 'long+medium' range was better than 'long' range contacts. In case of 'long+medium' range contacts for L/5 our method achieved more than 70% for 31 targets with 5 targets showing precision of 100% (Supplementary Table S3). In case of Top10 for 'long+medium' range contacts, our method predicts 38 targets with more than 70% precision (Supplementary Table S4). For more detailed information, see (Supplementary Table S5-S8) where performances of all methods have been compiled for L/5 'long' range contacts (Supplementary Table S5), Top10 'long' range contacs (Supplementary Table S6), L/5 'long+medium' range contacts (Supplementary Table S7) and Top10 'long+medium' range contacts (Supplementary Table S8) respectively.

Table 7: Summary of target-wise performance of method RRCPred 2.0 (raghavagps) on long range and long+medium range contact.

Precision (%)	**Number of CASP12 targets**			
	'long+medium' range contacts		**'long' range contacts**	
Range (percentage)	**L/5**	**Top10**	**L/5**	**Top10**
≥ 100	5	10	4	10
90 to 100	14	25	12	21

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

80 to 100	26	33	22	27
70 to 100	31	38	28	29

Evaluation and ranking by CASP12 organizers

Once competition is over, CASP organizers evaluate the performance of methods participated in competition and provides the results in the form of presentation; CASP12 results, analyzed by the organizers is also available from its website (http://predictioncenter.org/casp12/doc/presentations/CASP12_RR-Gaeta.pdf). We found that the organizers have analyzed the results on FM category targets where the common parameters were contacts which was of 'medium' and 'long' range combined and SumZ (>-2) whereas varying parameters were list size (L/2, L/5 and Full list), probability (>0 and >0.5). Weights taken into consideration were Precision, F1, F1prob., and AUC1. Method "RaptorX-Contact" was found to be present at top position in most of the evaluations and hence becoming one of the promising methods for FM targets contact prediction. We found that our method RRCPred 2.0 (raghavagps) also performed well and was present among the top 10 methods in most of the evaluations. The result suggests that our method performed comparably on FM targets when compared to other top methods and can be improved further. Unfortunately, the performance of TBM and TBM/FM targets were not evaluated and provided by the organizers in the presentation. This is the reason our method was not considered in successful predictors.

Conclusion

In last two decades, number of methods has been developed for predicting residue-residue contact. Despite tremendous efforts made by scientific community, the performance of these methods is far from satisfactory[29]. In particular, prediction of 'long' range contact is a tedious job. Due to its importance, it is highly desirable to develop highly accurate methods for predicting residue-residue contacts in a protein from amino acids sequence. This is the reason that residue-residue contact prediction is part of CASP competition. One of the major goals of CASP is to assess the quality of

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tertiary structure predicted by different methods and groups. Overall aim is to improve the performance of tertiary structure prediction methods. This raises a question whether one can use the predicted tertiary structures for predicting residue-residue contact in a protein. In order to address this question, our group 'raghavagps' used 150 predicted tertiary structures released by CASP12 for each target from Stage 2 tarballs. We ranked these structures using quality assessment software QASproCL then top 10 structures were used to assign/predict residue-residue contact in a target (see Materials & Methods). These predicted residue-residue contacts were submitted to CASP12 in manual submission category. In order to avoid any biasness, we extracted performance of models for all groups from CASP12 web site. This paper presents performance of method RRCPred 2.0 (raghavagps) and its comparison with performance of other methods. As shown in Results section in comparison to other models, this method performs better than other methods or in top 3 methods for both categories TBM and TBM/FM. This means predicted tertiary structure can be used to predict residue-residue contacts with high precision, particularly for easy targets (TBM or TBM/FM). The performance of RRCPred 2.0 (raghavagps) was poor for FM targets; still it is in list of top methods. We used QASproCL for ranking predicted structures, which was not the best method for ranking predicted structure as evaluated by CASP12. Thus it is possible to improve the performance of this method if we used best quality assessment method for ranking predicted structure.

Methodology

Dataset and Algorithm

We extracted Stage 2 tarball for each target from CASP12 web site, where each tarball consists of best 150 server models (predicted structures). These models/structures were predicted by servers which participated in TS category in CASP12. These tarballs were given as input in the quality assessment (QA) category in CASP12 competitions and were open to everyone for downloading from their web site. QASproCL, an in-house developed clustering based quality assessment software, was used for computing quality of each model for a given CASP12 target

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[**http://predictioncenter.org/casp12/doc/CASP12_Abstracts.pdf**]. These models were ranked based on their QASproCL model quality score (MQ) where the model getting the highest quality score was ranked first.. In order to compute residue-residue contact in a protein target, we used top/best 10 models. Following equation 1 was used for computing residue-residue contact score between two residues in a protein.

$$RRC_{i,j}^{m} = \frac{1}{10}\sum_{m=1}^{10} MQ_m \left[1 - \frac{[D_{i,j}-4]}{8}\right] \left[if\, D_{i,j} \leq 4 then D_{i,j} = 4\right] \quad (1)$$

Where m is the mth model, i and j are two residues forming a pair.

where $RRC_{i,j}$ is residue-residue contact score, MQ_m is model quality of mth model as predicted by QASproCL and D_i is distance between two residues in Angstrom (Å).

Schematic representation of complete algorithm applied in method "raghavagps" is presented in the form of flowchart (Figure 1).

Insert Figure 1 here

Parameters for Evaluation

The performance of prediction models was measured using different parameters (e.g. F1 score, precision, recall), which has been used in earlier CASP competition's for evaluating group performance[17]. These parameters can be computed using following equation 2-4.

$$F1score = \frac{2*precision*recall}{precision+recall} \quad (2)$$

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

$$Recall = \frac{TP}{(TP+FN)} \quad (4)$$

Where *TP* is true positives or correctly predicted contacts, *FP* is false positive or non-contacts being predicted as contacts, *TN* is true negative or correctly predicted non-contacts and FN is false negative or contacts being predicted as non-contacts.

Criteria for Evaluation

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In the present CASP12 competition, 'long' and 'long+medium' range contacts were analyzed since 'long' range contacts are hard to predict. In order to evaluate performance of all contact predictors, number of contacts was kept same by trimming the list to N contacts with highest probabilities. The List Size parameter N determines contact lists length and takes different values i.e. L/2, L/5, Top 10 and FL (Full Length). Here L represents the length of protein sequence; Top 10 means first 10 contacts submitted for a given sequence whereas FL represents all predicted contacts, which is different for different predictions. Different types of scores for each target as well as for each group were calculated for evaluating performance.

References

1 Dorn, M., MB, E. S., Buriol, L. S. & Lamb, L. C. Three-dimensional protein structure prediction: Methods and computational strategies. *Computational biology and chemistry* **53PB**, 251-276, doi:10.1016/j.compbiolchem.2014.10.001 (2014).

2 Adhikari, B. & Cheng, J. Protein Residue Contacts and Prediction Methods. *Methods in molecular biology* **1415**, 463-476, doi:10.1007/978-1-4939-3572-7_24 (2016).

3 Marks, D. S. *et al.* Protein 3D structure computed from evolutionary sequence variation. *PloS one* **6**, e28766, doi:10.1371/journal.pone.0028766 (2011).

4 Adhikari, B., Bhattacharya, D., Cao, R. & Cheng, J. CONFOLD: Residue-residue contact-guided ab initio protein folding. *Proteins* **83**, 1436-1449, doi:10.1002/prot.24829 (2015).

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

5 Kliger, Y. *et al.* Peptides modulating conformational changes in secreted chaperones: from in silico design to preclinical proof of concept. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13797-13801, doi:10.1073/pnas.0906514106 (2009).

6 Miller, C. S. & Eisenberg, D. Using inferred residue contacts to distinguish between correct and incorrect protein models. *Bioinformatics* **24**, 1575-1582, doi:10.1093/bioinformatics/btn248 (2008).

7 Wang, Z., Eickholt, J. & Cheng, J. APOLLO: a quality assessment service for single and multiple protein models. *Bioinformatics* **27**, 1715-1716, doi:10.1093/bioinformatics/btr268 (2011).

8 Vendruscolo, M. & Domany, E. Protein folding using contact maps. *Vitamins and hormones* **58**, 171-212 (2000).

9 Mirny, L. & Domany, E. Protein fold recognition and dynamics in the space of contact maps. *Proteins* **26**, 391-410, doi:10.1002/(SICI)1097-0134(199612)26:4<391::AID-PROT3>3.0.CO;2-F (1996).

10 Rohl, C. A., Strauss, C. E., Misura, K. M. & Baker, D. Protein structure prediction using Rosetta. *Methods in enzymology* **383**, 66-93, doi:10.1016/S0076-6879(04)83004-0 (2004).

11 Jones, D. T. Predicting novel protein folds by using FRAGFOLD. *Proteins* **Suppl 5**, 127-132 (2001).

12 Lesk, A. M. CASP2: report on ab initio predictions. *Proteins* **Suppl 1**, 151-166 (1997).

13 Grana, O. *et al.* CASP6 assessment of contact prediction. *Proteins* **61 Suppl 7**, 214-224, doi:10.1002/prot.20739 (2005).

14 Izarzugaza, J. M., Grana, O., Tress, M. L., Valencia, A. & Clarke, N. D. Assessment of intramolecular contact predictions for CASP7. *Proteins* **69 Suppl 8**, 152-158, doi:10.1002/prot.21637 (2007).

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

15 Ezkurdia, I., Grana, O., Izarzugaza, J. M. & Tress, M. L. Assessment of domain boundary predictions and the prediction of intramolecular contacts in CASP8. *Proteins* **77 Suppl 9**, 196-209, doi:10.1002/prot.22554 (2009).

16 Monastyrskyy, B., Fidelis, K., Tramontano, A. & Kryshtafovych, A. Evaluation of residue-residue contact predictions in CASP9. *Proteins* **79 Suppl 10**, 119-125, doi:10.1002/prot.23160 (2011).

17 Monastyrskyy, B., D'Andrea, D., Fidelis, K., Tramontano, A. & Kryshtafovych, A. Evaluation of residue-residue contact prediction in CASP10. *Proteins* **82 Suppl 2**, 138-153, doi:10.1002/prot.24340 (2014).

18 Monastyrskyy, B., D'Andrea, D., Fidelis, K., Tramontano, A. & Kryshtafovych, A. New encouraging developments in contact prediction: Assessment of the CASP11 results. *Proteins* **84 Suppl 1**, 131-144, doi:10.1002/prot.24943 (2016).

19 Sathyapriya, R., Duarte, J. M., Stehr, H., Filippis, I. & Lappe, M. Defining an essence of structure determining residue contacts in proteins. *PLoS computational biology* **5**, e1000584, doi:10.1371/journal.pcbi.1000584 (2009).

20 Cheng, J. & Baldi, P. Improved residue contact prediction using support vector machines and a large feature set. *BMC bioinformatics* **8**, 113, doi:10.1186/1471-2105-8-113 (2007).

21 Wu, S. & Zhang, Y. A comprehensive assessment of sequence-based and template-based methods for protein contact prediction. *Bioinformatics* **24**, 924-931, doi:10.1093/bioinformatics/btn069 (2008).

22 Bjorkholm, P. *et al.* Using multi-data hidden Markov models trained on local neighborhoods of protein structure to predict residue-residue contacts. *Bioinformatics* **25**, 1264-1270, doi:10.1093/bioinformatics/btp149 (2009).

23 Fariselli, P., Olmea, O., Valencia, A. & Casadio, R. Prediction of contact maps with neural networks and correlated mutations. *Protein engineering* **14**, 835-843 (2001).

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

24 Hamilton, N., Burrage, K., Ragan, M. A. & Huber, T. Protein contact prediction using patterns of correlation. *Proteins* **56**, 679-684, doi:10.1002/prot.20160 (2004).

25 Eickholt, J. & Cheng, J. Predicting protein residue-residue contacts using deep networks and boosting. *Bioinformatics* **28**, 3066-3072, doi:10.1093/bioinformatics/bts598 (2012).

26 Wang, Z. & Xu, J. Predicting protein contact map using evolutionary and physical constraints by integer programming. *Bioinformatics* **29**, i266-273, doi:10.1093/bioinformatics/btt211 (2013).

27 Schneider, M. & Brock, O. Combining physicochemical and evolutionary information for protein contact prediction. *PloS one* **9**, e108438, doi:10.1371/journal.pone.0108438 (2014).

28 Jones, D. T., Singh, T., Kosciolek, T. & Tetchner, S. MetaPSICOV: combining coevolution methods for accurate prediction of contacts and long range hydrogen bonding in proteins. *Bioinformatics* **31**, 999-1006, doi:10.1093/bioinformatics/btu791 (2015).

29 Kundrotas, P. J. & Alexov, E. G. Predicting residue contacts using pragmatic correlated mutations method: reducing the false positives. *BMC bioinformatics* **7**, 503, doi:10.1186/1471-2105-7-503 (2006).

Acknowledgement

Authors are thankful to Dr. Chinmayee Choudhury for her kind help during QASproCL algorithm development and funding agencies Department of Science and Technology (DST-INSPIRE), Council of Scientific and Industrial Research (CSIR), J.C.Bose National Fellowship from Department of Science and Technology (DST), for fellowships and financial support.

AUTHORS' CONTRIBUTION:

S.S., G.N., P.A. and D.S. developed the algorithm QASproCL. P.A. and S.S extracted the data and predicted the contacts. P.A. submitted the results manually to the CASP site. P.A. extracted the results from CASP12 site and anlayzed the whole data with the help of G.P.S.R. P.A. and G.P.S.R. wrote the manuscript. G.P.S.R. conceived the idea and coordinated the project.

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

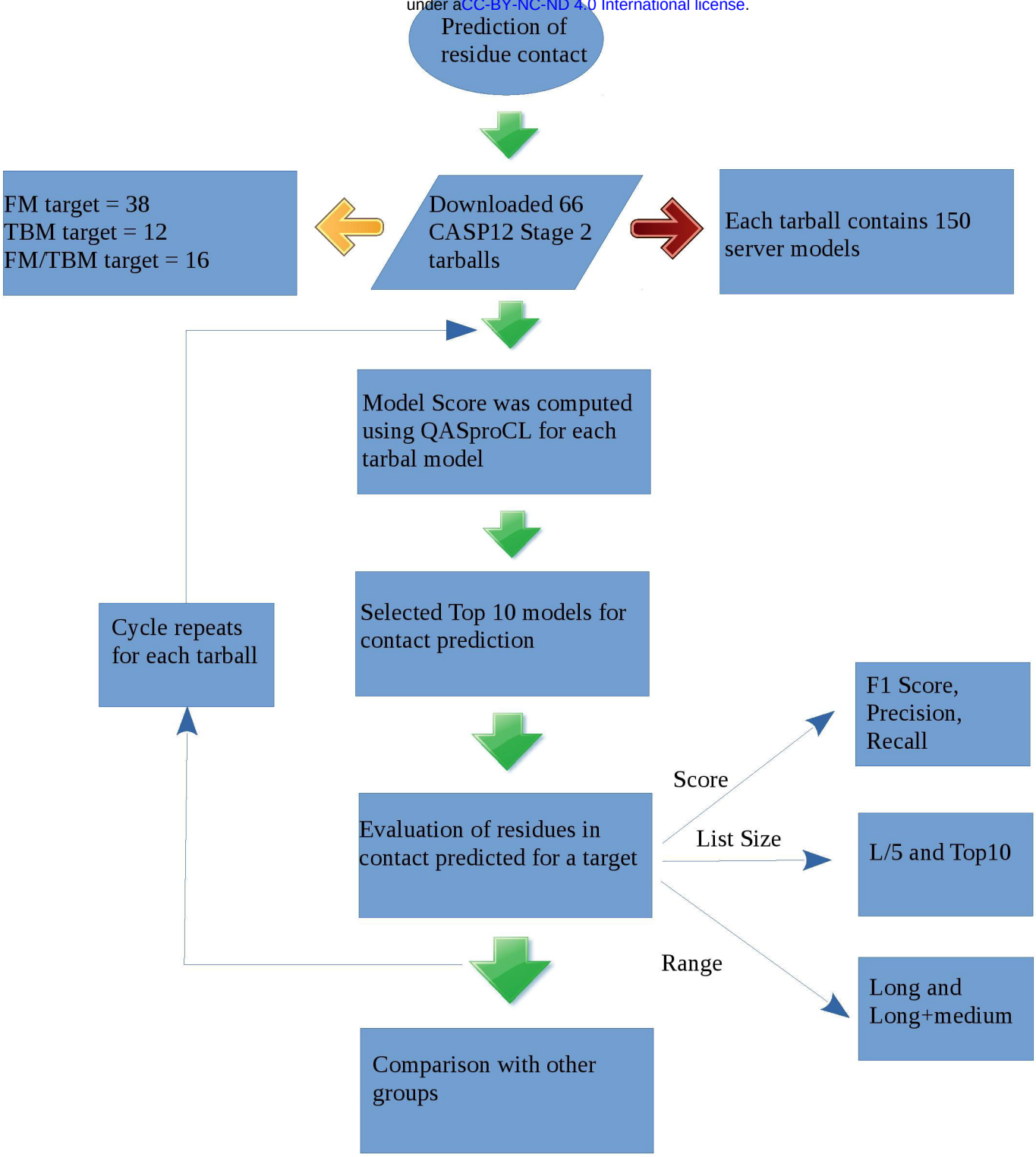
COMPETING INTERESTS:

The authors' declare that they have no conflict of interest.

Figure Legend

Figure 1. Schematic represntation of the methodolgy followed for predicting residue-residue contact in the method RRCPred 2.0 (raghavagps).

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Prediction of residue-residue contacts in CASP12 targets from its predicted tertiary structures

Piyush Agrawal[1], Sandeep Singh[1], Gandharva Nagpal[1], Deepti Sethi[1], Gajendra P.S. Raghava*[,1,2]

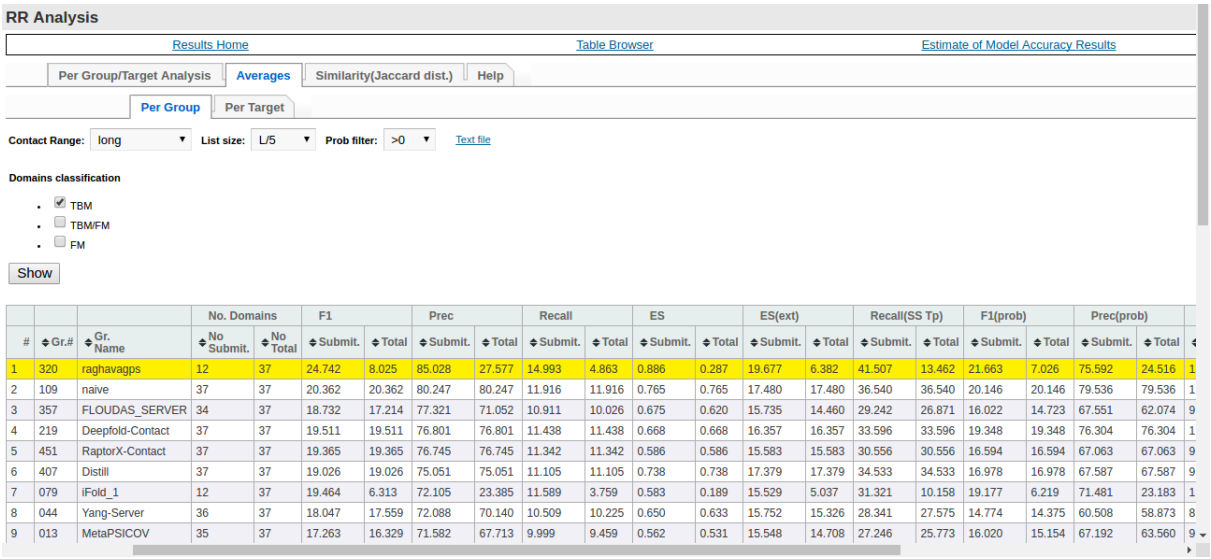
1. CSIR-Institute of Microbial Technology, Chandigarh 160036, India.

2. Indraprastha Institute of Information Technology, New Delhi 110020, India.

Supplementary Information

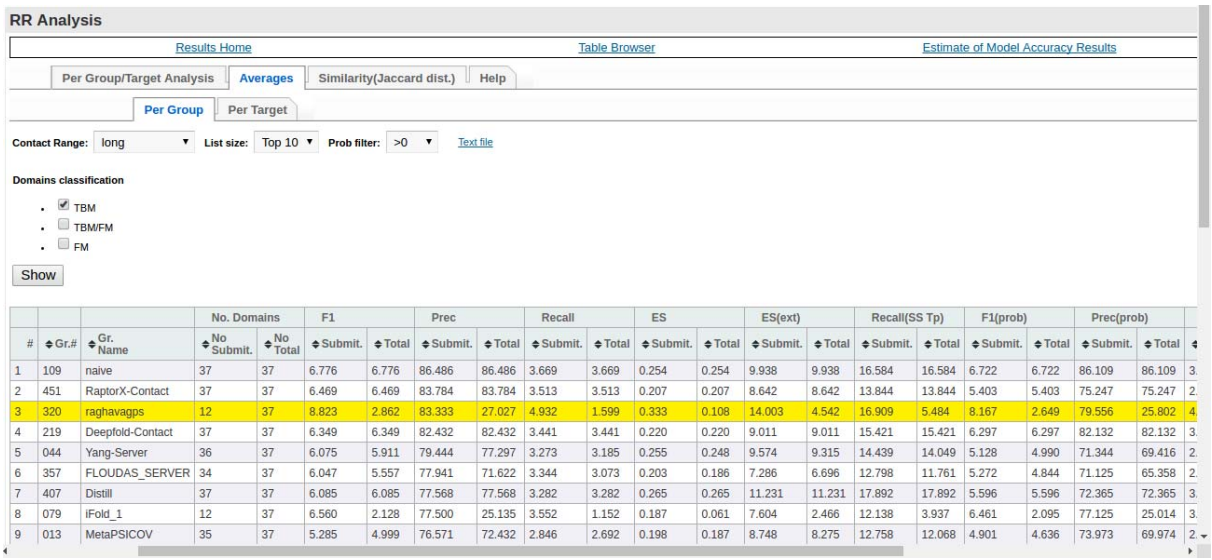
bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S1. Screenshot of the performance or precision of different methods for L/5 list size and 'long' range contacts on CASP12 TBM targets.



			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	320	raghavagps	12	37	24.742	8.025	85.028	27.577	14.993	4.863	0.886	0.287	19.677	6.382	41.507	13.462	21.663	7.026	75.592	24.516
2	109	naive	37	37	20.362	20.362	80.247	80.247	11.916	11.916	0.765	0.765	17.480	17.480	36.540	36.540	20.146	20.146	79.536	79.536
3	357	FLOUDAS_SERVER	34	37	18.732	17.214	77.321	71.052	10.911	10.026	0.675	0.620	15.735	14.460	29.242	26.871	16.022	14.723	67.551	62.074
4	219	Deepfold-Contact	37	37	19.511	19.511	76.801	76.801	11.438	11.438	0.668	0.668	16.357	16.357	33.596	33.596	19.348	19.348	76.304	76.304
5	451	RaptorX-Contact	37	37	19.365	19.365	76.745	76.745	11.342	11.342	0.586	0.586	15.583	15.583	30.556	30.556	16.594	16.594	67.063	67.063
6	407	Distill	37	37	19.026	19.026	75.051	75.051	11.105	11.105	0.738	0.738	17.379	17.379	34.533	34.533	16.978	16.978	67.587	67.587
7	079	iFold_1	12	37	19.464	6.313	72.105	23.385	11.589	3.759	0.583	0.189	15.529	5.037	31.321	10.158	19.177	6.219	71.481	23.183
8	044	Yang-Server	36	37	18.047	17.559	72.088	70.140	10.509	10.225	0.650	0.633	15.752	15.326	28.341	27.575	14.774	14.375	60.508	58.873
9	013	MetaPSICOV	35	37	17.263	16.329	71.582	67.713	9.999	9.459	0.562	0.531	15.548	14.708	27.246	25.773	16.020	15.154	67.192	63.560

Figure S2. Screenshot of the performance or precision of different methods for Top10 list size and 'long' range contacts on CASP12 TBM targets.



			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	109	naive	37	37	6.776	6.776	86.486	86.486	3.669	3.669	0.254	0.254	9.938	9.938	16.584	16.584	6.722	6.722	86.109	86.109
2	451	RaptorX-Contact	37	37	6.469	6.469	83.784	83.784	3.513	3.513	0.207	0.207	8.642	8.642	13.844	13.844	5.403	5.403	75.247	75.247
3	320	raghavagps	12	37	8.823	2.862	83.333	27.027	4.932	1.599	0.333	0.108	14.003	4.542	16.909	5.484	8.167	2.649	79.556	25.802
4	219	Deepfold-Contact	37	37	6.349	6.349	82.432	82.432	3.441	3.441	0.220	0.220	9.011	9.011	15.421	15.421	6.297	6.297	82.132	82.132
5	044	Yang-Server	36	37	6.075	5.911	79.444	77.297	3.273	3.185	0.255	0.248	9.574	9.315	14.439	14.049	5.128	4.990	71.344	69.416
6	357	FLOUDAS_SERVER	34	37	6.047	5.557	77.941	71.622	3.344	3.073	0.203	0.186	7.286	6.696	12.798	11.761	5.272	4.844	71.125	65.358
7	407	Distill	37	37	6.085	6.085	77.568	77.568	3.282	3.282	0.265	0.265	11.231	11.231	17.892	17.892	5.596	5.596	72.365	72.365
8	079	iFold_1	12	37	6.560	2.128	77.500	25.135	3.552	1.152	0.187	0.061	7.604	2.466	12.138	3.937	6.461	2.095	77.125	25.014
9	013	MetaPSICOV	35	37	5.285	4.999	76.571	72.432	2.846	2.692	0.198	0.187	8.748	8.275	12.758	12.068	4.901	4.636	73.973	69.974

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S3. Screenshot of the performance or precision of different methods for L/5 list size and 'long' range contacts on CASP12 TBM/FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long List size: L/5 Prob filter: >0 Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	
1	320	raghavagps	16	19	25.365	21.360	72.244	60.837	16.466	13.866	0.446	0.376	8.434	7.102	42.068	35.425	19.836	16.704	57.987	48.832	12.
2	079	iFold_1	17	19	19.684	17.612	58.117	51.999	12.620	11.292	0.343	0.306	8.210	7.346	33.366	29.854	19.135	17.121	56.761	50.786	12.
3	013	MetaPSICOV	19	19	19.419	19.419	53.755	53.755	12.594	12.594	0.285	0.285	7.327	7.327	32.855	32.855	13.008	13.008	38.314	38.314	8.3
4	097	PconsC31	19	19	17.461	17.461	52.342	52.342	11.126	11.126	0.301	0.301	6.906	6.906	28.074	28.074	7.558	7.558	24.811	24.811	4.7
5	109	naive	19	19	18.218	18.218	51.988	51.988	11.752	11.752	0.314	0.314	7.459	7.459	30.742	30.742	17.691	17.691	50.745	50.745	11.
6	219	Deepfold-Contact	19	19	17.342	17.342	51.026	51.026	11.109	11.109	0.311	0.311	7.143	7.143	29.744	29.744	16.872	16.872	49.926	49.926	10.
7	251	myprotein-me	16	19	14.453	12.171	47.892	40.330	9.076	7.643	0.192	0.162	5.143	4.331	21.725	18.295	3.194	2.690	11.866	9.992	1.9
8	451	RaptorX-Contact	19	19	15.903	15.903	46.579	46.579	10.126	10.126	0.281	0.281	7.657	7.657	21.185	21.185	10.369	10.369	32.693	32.693	6.4
9	020	RBO-Epsilon	19	19	17.703	17.703	46.286	46.286	11.587	11.587	0.261	0.261	6.534	6.534	23.908	23.908	8.840	8.840	22.939	22.939	5.8

Figure S4. Screenshot of the performance or precision of different methods for Top10 list size and 'long' range contacts on CASP12 TBM/FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long List size: Top 10 Prob filter: >0 Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	
1	320	raghavagps	16	19	15.252	12.844	73.125	61.579	9.393	7.910	0.231	0.195	6.355	5.352	29.898	25.177	12.007	10.111	61.877	52.107	7.
2	079	iFold_1	17	19	11.619	10.396	62.353	55.789	7.118	6.368	0.153	0.137	5.717	5.115	23.157	20.719	11.360	10.164	61.461	54.991	6.
3	097	PconsC31	19	19	11.733	11.733	61.579	61.579	7.078	7.078	0.161	0.161	5.379	5.379	21.498	21.498	4.671	4.671	31.594	31.594	2.
4	013	MetaPSICOV	19	19	13.332	13.332	59.474	59.474	8.207	8.207	0.166	0.166	5.424	5.424	26.392	26.392	8.747	8.747	45.209	45.209	5.
5	020	RBO-Epsilon	19	19	13.472	13.472	57.368	57.368	8.335	8.335	0.162	0.162	4.463	4.463	19.102	19.102	7.365	7.365	34.225	34.225	4.
6	251	myprotein-me	16	19	8.797	7.408	55.000	46.316	5.276	4.443	0.082	0.069	3.522	2.966	16.602	13.981	1.744	1.469	15.906	13.394	1.
7	109	naive	19	19	11.332	11.332	54.211	54.211	6.929	6.929	0.145	0.145	5.463	5.463	20.162	20.162	11.043	11.043	53.306	53.306	6.
8	219	Deepfold-Contact	19	19	10.591	10.591	53.684	53.684	6.421	6.421	0.143	0.143	5.276	5.276	19.816	19.816	10.319	10.319	52.875	52.875	6.
9	024	PconsC2	19	19	9.479	9.479	50.526	50.526	5.759	5.759	0.131	0.131	4.698	4.698	17.349	17.349	1.864	1.864	14.652	14.652	1.

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S5. Screenshot of the performance or precision of different methods for L/5 list size and 'long' range contacts on CASP12 FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results | RR Assessm

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long | List size: L/5 | Prob filter: >0 | Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)		Prec-FDR(prob)		
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	
1	451	RaptorX-Contact	38	38	12.800	12.800	47.095	47.095	7.499	7.499	0.210	0.210	6.512	6.512	22.752	22.752	9.374	9.374	35.552	35.552	5.455	5.455	7.610	7.610	47.
2	079	iFold_1	37	38	12.639	12.307	46.864	45.630	7.402	7.208	0.207	0.202	6.081	5.921	23.995	23.363	12.381	12.056	46.034	44.823	7.247	7.057	-4.016	-3.910	46.
3	219	Deepfold-Contact	36	38	12.423	11.769	45.549	43.152	7.298	6.914	0.219	0.207	6.410	6.073	21.172	20.058	12.178	11.537	44.860	42.499	7.146	6.770	-6.691	-6.339	45.
4	013	MetaPSICOV	38	38	11.463	11.463	44.007	44.007	6.658	6.658	0.221	0.221	6.888	6.888	21.198	21.198	7.546	7.546	30.250	30.250	4.344	4.344	12.176	12.176	44.
5	287	MULTICOM-CLUSTER	38	38	11.020	11.020	41.691	41.691	6.419	6.419	0.195	0.195	5.679	5.679	24.409	24.409	9.194	9.194	35.569	35.569	5.330	5.330	-1.244	-1.244	42.
6	236	MULTICOM-CONSTRUCT	38	38	10.978	10.978	41.637	41.637	6.392	6.392	0.194	0.194	5.651	5.651	24.058	24.058	7.148	7.148	28.379	28.379	4.123	4.123	7.899	7.899	42.
7	109	naive	36	38	11.088	10.504	41.388	39.210	6.483	6.142	0.212	0.201	6.169	5.844	21.832	20.683	10.839	10.269	40.534	38.401	6.336	6.002	-14.978	-14.189	41.
8	432	Pcons-net	38	38	10.981	10.981	41.013	41.013	6.417	6.417	0.176	0.176	5.091	5.091	23.642	23.642	6.140	6.140	23.518	23.518	3.572	3.572	0.582	0.582	41.
9	310	IGBteam	38	38	10.455	10.455	38.784	38.784	6.127	6.127	0.134	0.134	4.437	4.437	15.868	15.868	7.112	7.112	27.186	27.186	4.139	4.139	-5.294	-5.294	39.
10	373	Zhang_Contact	36	38	10.083	9.552	37.801	35.811	5.907	5.596	0.189	0.179	5.908	5.597	22.840	21.638	6.229	5.902	24.553	23.261	3.608	3.418	-4.709	-4.461	38.
11	097	PconsC31	38	38	9.911	9.911	37.772	37.772	5.772	5.772	0.160	0.160	4.453	4.453	20.203	20.203	5.549	5.549	22.224	22.224	3.198	3.198	-0.460	-0.460	38.
12	348	FALCON_COLORS	38	38	9.775	9.775	35.935	35.935	5.735	5.735	0.183	0.183	5.967	5.967	23.253	23.253	5.071	5.071	19.508	19.508	2.944	2.944	-3.202	-3.202	36.
13	020	RBO-Epsilon	38	38	9.636	9.636	35.186	35.186	5.648	5.648	0.201	0.201	6.152	6.152	24.941	24.941	4.633	4.633	17.504	17.504	2.696	2.696	-1.313	-1.313	37.
14	431	Shen-Group	38	38	9.518	9.518	34.215	34.215	5.610	5.610	0.227	0.227	6.998	6.998	23.057	23.057	6.713	6.713	24.122	24.122	3.958	3.958	-22.143	-22.143	34.
15	044	Yang-Server	37	38	9.106	8.867	33.849	32.958	5.326	5.186	0.155	0.150	5.112	4.978	21.582	21.014	4.608	4.487	17.671	17.206	2.680	2.609	-9.605	-9.352	35.
16	389	AkbAR	36	38	8.959	8.488	33.336	31.582	5.256	4.979	0.233	0.220	6.581	6.234	26.068	24.696	3.581	3.392	14.269	13.518	2.066	1.957	2.376	2.251	35.
17	320	raghavagps	38	38	9.476	9.476	32.883	32.883	5.626	5.626	0.216	0.216	5.716	5.716	13.055	13.055	6.517	6.517	22.599	22.599	3.874	3.874	-17.382	-17.382	34.
18	024	PconsC2	36	38	8.430	7.987	32.842	31.113	4.883	4.626	0.161	0.152	4.711	4.463	17.807	16.870	2.188	2.072	8.896	8.427	1.254	1.188	5.986	5.671	37.

Figure S6. Screenshot of the performance or precision of different methods for Top10 list size and 'long' range contacts on CASP12 FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results | RR Assessm

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long | List size: Top 10 | Prob filter: >0 | Text file

Domains classification

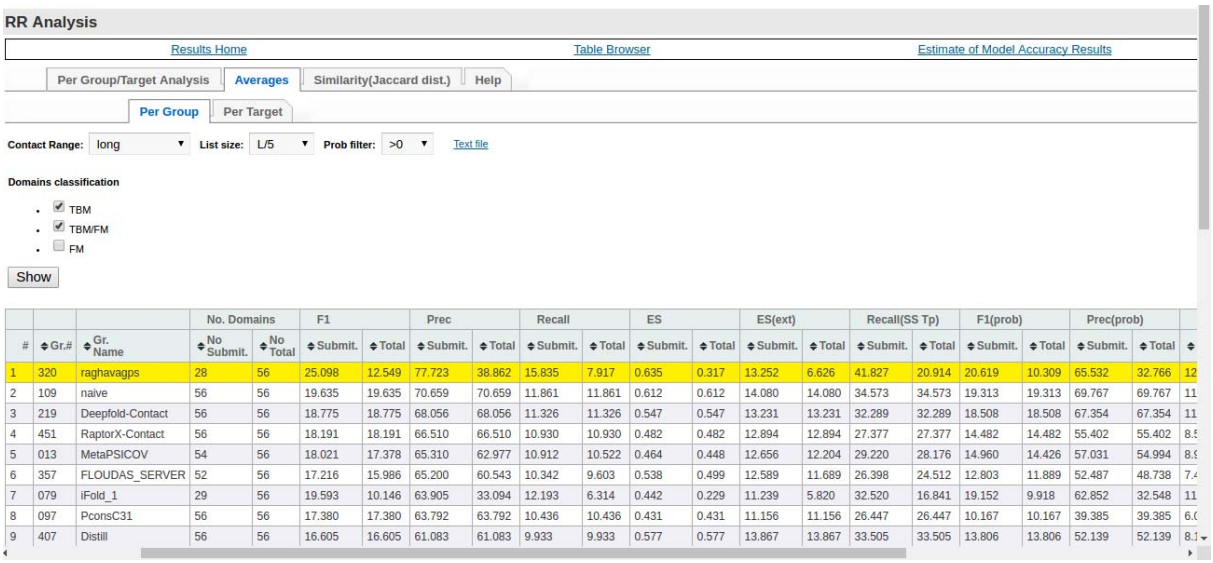
- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)		Prec-FDR(prob)		
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	
1	451	RaptorX-Contact	38	38	6.706	6.706	54.211	54.211	3.620	3.620	0.109	0.109	4.103	4.103	14.061	14.061	5.089	5.089	42.688	42.688	2.737	2.737	17.318	17.318	54.
2	079	iFold_1	37	38	6.356	6.189	50.270	48.947	3.452	3.361	0.082	0.080	3.358	3.269	13.394	13.042	6.279	6.114	49.786	48.476	3.410	3.320	2.044	1.990	50.
3	287	MULTICOM-CLUSTER	38	38	5.523	5.523	48.947	48.947	2.966	2.966	0.088	0.088	2.757	2.757	14.571	14.571	4.707	4.707	43.456	43.456	2.520	2.520	10.472	10.472	49.
4	219	Deepfold-Contact	36	38	6.348	6.013	48.889	46.316	3.452	3.270	0.084	0.080	3.461	3.279	12.402	11.749	6.270	5.940	48.436	45.887	3.408	3.229	-1.092	-1.035	48.
5	236	MULTICOM-CONSTRUCT	38	38	5.516	5.516	48.421	48.421	2.963	2.963	0.091	0.091	3.003	3.003	14.770	14.770	3.884	3.884	36.783	36.783	2.075	2.075	16.871	16.871	49.
6	013	MetaPSICOV	38	38	5.599	5.599	47.368	47.368	3.012	3.012	0.090	0.090	3.501	3.501	11.644	11.644	4.161	4.161	37.361	37.361	2.230	2.230	17.976	17.976	48.
7	432	Pcons-net	38	38	5.401	5.401	46.579	46.579	2.909	2.909	0.083	0.083	2.954	2.954	13.340	13.340	3.219	3.219	29.550	29.550	1.728	1.728	7.648	7.648	46.
8	109	naive	36	38	5.623	5.327	45.556	43.158	3.040	2.880	0.096	0.091	3.951	3.743	15.860	15.025	5.530	5.239	44.951	42.585	2.989	2.832	-7.690	-7.285	45.
9	097	PconsC31	38	38	4.899	4.899	43.421	43.421	2.636	2.636	0.077	0.077	2.755	2.755	11.869	11.869	2.890	2.890	28.164	28.164	1.542	1.542	5.518	5.518	43.
10	310	IGBteam	38	38	5.549	5.549	43.421	43.421	3.015	3.015	0.076	0.076	2.941	2.941	10.802	10.802	4.010	4.010	33.167	33.167	2.170	2.170	-0.131	-0.131	44.
11	020	RBO-Epsilon	38	38	5.209	5.209	41.316	41.316	2.818	2.818	0.095	0.095	3.177	3.177	15.009	15.009	2.898	2.898	25.052	25.052	1.561	1.561	3.637	3.637	42.
12	348	FALCON_COLORS	38	38	5.021	5.021	41.316	41.316	2.714	2.714	0.087	0.087	3.327	3.327	12.607	12.607	2.955	2.955	26.032	26.032	1.589	1.589	1.791	1.791	41.
13	373	Zhang_Contact	36	38	4.607	4.364	40.833	38.684	2.479	2.349	0.089	0.084	3.301	3.127	14.711	13.937	3.262	3.090	30.141	28.554	1.749	1.657	-0.773	-0.732	40.
14	044	Yang-Server	37	38	4.804	4.677	39.730	38.684	2.592	2.523	0.084	0.082	3.056	2.976	14.523	14.141	2.729	2.657	23.515	22.896	1.471	1.432	-6.299	-6.134	40.
15	431	Shen-Group	38	38	4.881	4.881	39.474	39.474	2.644	2.644	0.106	0.106	4.453	4.453	13.872	13.872	3.688	3.688	30.201	30.201	1.996	1.996	-15.713	-15.713	39.
16	389	AkbAR	36	38	4.692	4.445	39.167	37.105	2.538	2.405	0.113	0.107	4.348	4.119	16.186	15.334	2.214	2.097	20.439	19.363	1.186	1.123	6.084	5.764	40.
17	024	PconsC2	36	38	4.430	4.197	38.611	36.579	2.382	2.257	0.090	0.086	3.607	3.418	13.198	12.503	1.356	1.285	13.345	12.642	0.722	0.684	9.728	9.216	41.
18	320	raghavagps	38	38	4.861	4.861	37.632	37.632	2.647	2.647	0.112	0.112	3.433	3.433	8.157	8.157	3.662	3.662	28.609	28.609	1.996	1.996	-12.943	-12.943	38.

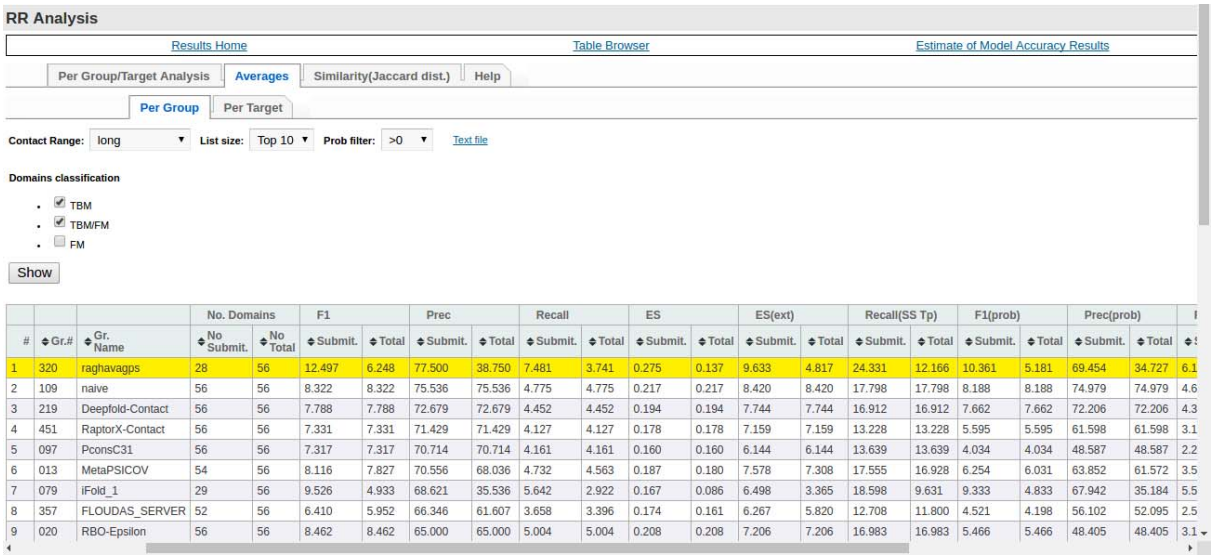
bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S7. Screenshot of the performance or precision of different methods for L/5 list size and 'long' range contacts on CASP12 TBM and TBM/FM targets.



			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	
1	320	raghavagps	28	56	25.098	12.549	77.723	38.862	15.835	7.917	0.635	0.317	13.252	6.626	41.827	20.914	20.619	10.309	65.532	32.766	12
2	109	naive	56	56	19.635	19.635	70.659	70.659	11.861	11.861	0.612	0.612	14.080	14.080	34.573	34.573	19.313	19.313	69.767	69.767	11
3	219	Deepfold-Contact	56	56	18.775	18.775	68.056	68.056	11.326	11.326	0.547	0.547	13.231	13.231	32.289	32.289	18.508	18.508	67.354	67.354	11
4	451	RaptorX-Contact	56	56	18.191	18.191	66.510	66.510	10.930	10.930	0.482	0.482	12.894	12.894	27.377	27.377	14.482	14.482	55.402	55.402	8.5
5	013	MetaPSICOV	54	56	18.021	17.378	65.310	62.977	10.912	10.522	0.464	0.448	12.656	12.204	29.220	28.176	14.960	14.426	57.031	54.994	8.9
6	357	FLOUDAS_SERVER	52	56	17.216	15.986	65.200	60.543	10.342	9.603	0.538	0.499	12.589	11.689	26.398	24.512	12.803	11.889	52.487	48.738	7.4
7	079	iFold_1	29	56	19.593	10.146	63.905	33.094	12.193	6.314	0.442	0.229	11.239	5.820	32.520	16.841	19.152	9.918	62.852	32.548	11
8	097	PconsC31	56	56	17.380	17.380	63.792	63.792	10.436	10.436	0.431	0.431	11.156	11.156	26.447	26.447	10.167	10.167	39.385	39.385	6.0
9	407	Distill	56	56	16.605	16.605	61.083	61.083	9.933	9.933	0.577	0.577	13.867	13.867	33.505	33.505	13.806	13.806	52.139	52.139	8.1

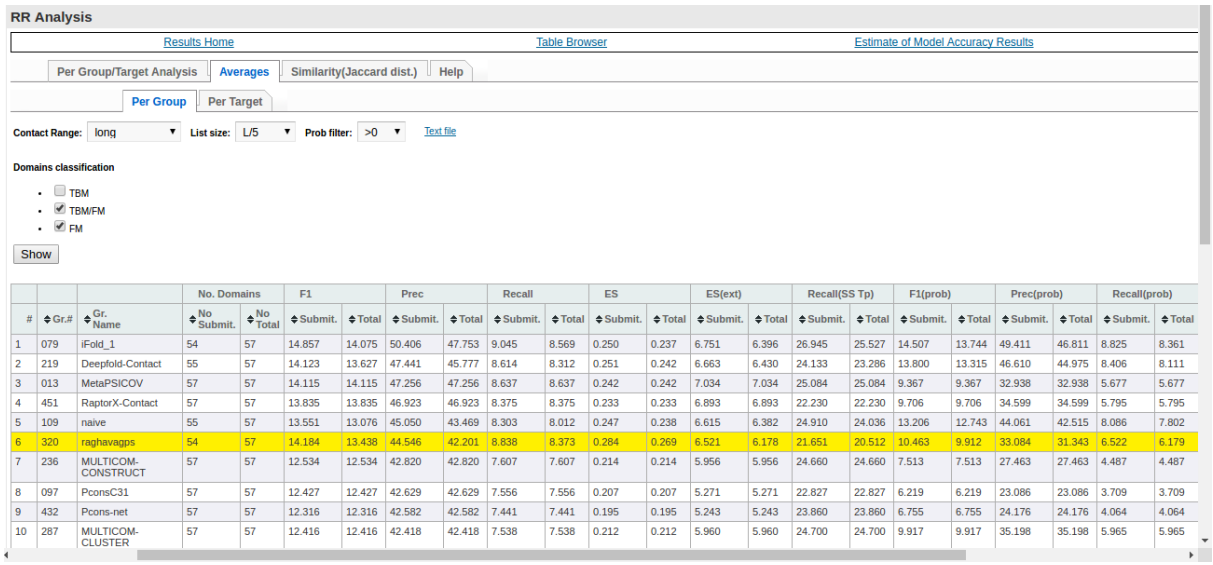
Figure S8. Screenshot of the performance or precision of different methods for Top10 list size and 'long' range contacts on CASP12 TBM and TBM/FM targets.



			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	
1	320	raghavagps	28	56	12.497	6.248	77.500	38.750	7.481	3.741	0.275	0.137	9.633	4.817	24.331	12.166	10.361	5.181	69.454	34.727	6.1
2	109	naive	56	56	8.322	8.322	75.536	75.536	4.775	4.775	0.217	0.217	8.420	8.420	17.798	17.798	8.188	8.188	74.979	74.979	4.6
3	219	Deepfold-Contact	56	56	7.788	7.788	72.679	72.679	4.452	4.452	0.194	0.194	7.744	7.744	16.912	16.912	7.662	7.662	72.206	72.206	4.3
4	451	RaptorX-Contact	56	56	7.331	7.331	71.429	71.429	4.127	4.127	0.178	0.178	7.159	7.159	13.228	13.228	5.595	5.595	61.598	61.598	3.1
5	097	PconsC31	56	56	7.317	7.317	70.714	70.714	4.161	4.161	0.160	0.160	6.144	6.144	13.639	13.639	4.034	4.034	48.587	48.587	2.2
6	013	MetaPSICOV	54	56	8.116	7.827	70.556	68.036	4.732	4.563	0.187	0.180	7.578	7.308	17.555	16.928	6.254	6.031	63.852	61.572	3.5
7	079	iFold_1	29	56	9.526	4.933	68.621	35.536	5.642	2.922	0.167	0.086	6.498	3.365	18.598	9.631	9.333	4.833	67.942	35.184	5.5
8	357	FLOUDAS_SERVER	52	56	6.410	5.952	66.346	61.607	3.658	3.396	0.174	0.161	6.267	5.820	12.708	11.800	4.521	4.198	56.102	52.095	2.5
9	020	RBO-Epsilon	56	56	8.462	8.462	65.000	65.000	5.004	5.004	0.208	0.208	7.206	7.206	16.983	16.983	5.466	5.466	48.405	48.405	3.1

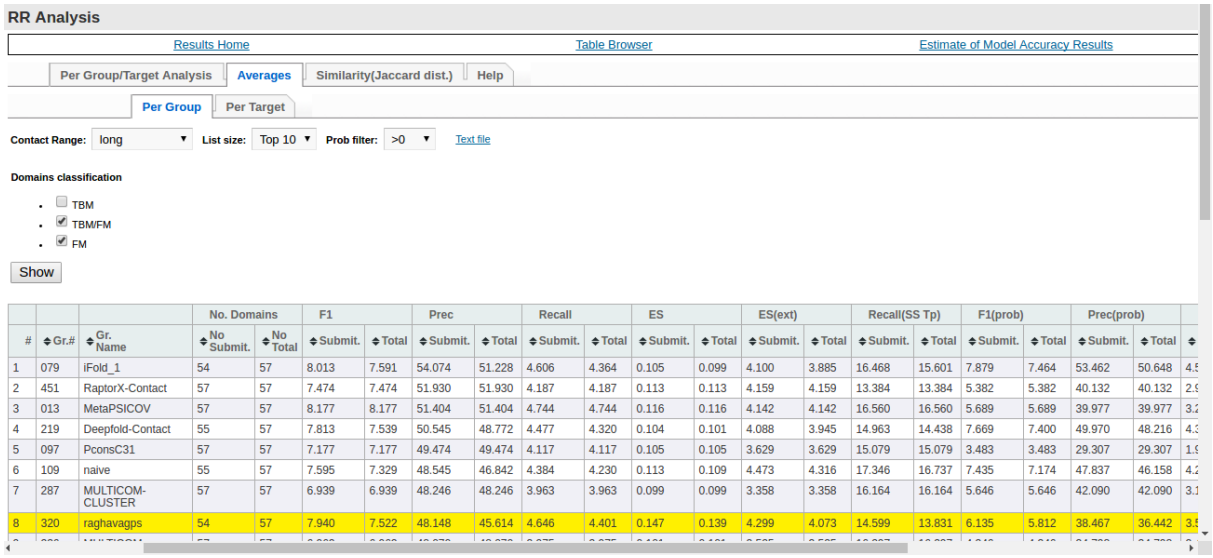
bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S9. Screenshot of the performance or precision of different methods for L/5 list size and 'long' range contacts on CASP12 TBM/FM and FM targets.



			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	079	iFold_1	54	57	14.857	14.075	50.406	47.753	9.045	8.569	0.250	0.237	6.751	6.396	26.945	25.527	14.507	13.744	49.411	46.811	8.825	8.361
2	219	Deepfold-Contact	55	57	14.123	13.627	47.441	45.777	8.614	8.312	0.251	0.242	6.663	6.430	24.133	23.286	13.800	13.315	46.610	44.975	8.406	8.111
3	013	MetaPSICOV	57	57	14.115	14.115	47.256	47.256	8.637	8.637	0.242	0.242	7.034	7.034	25.084	25.084	9.367	9.367	32.938	32.938	5.677	5.677
4	451	RaptorX-Contact	57	57	13.835	13.835	46.923	46.923	8.375	8.375	0.233	0.233	6.893	6.893	22.230	22.230	9.706	9.706	34.599	34.599	5.795	5.795
5	109	naive	55	57	13.551	13.076	45.050	43.469	8.303	8.012	0.247	0.238	6.615	6.382	24.910	24.036	13.206	12.743	44.061	42.515	8.086	7.802
6	320	raghavagps	54	57	14.184	13.438	44.546	42.201	8.838	8.373	0.284	0.269	6.521	6.178	21.651	20.512	10.463	9.912	33.084	31.343	6.522	6.179
7	236	MULTICOM-CONSTRUCT	57	57	12.534	12.534	42.820	42.820	7.607	7.607	0.214	0.214	5.956	5.956	24.660	24.660	7.513	7.513	27.463	27.463	4.487	4.487
8	097	PconsC31	57	57	12.427	12.427	42.629	42.629	7.556	7.556	0.207	0.207	5.271	5.271	22.827	22.827	6.219	6.219	23.086	23.086	3.709	3.709
9	432	Pcons-net	57	57	12.316	12.316	42.582	42.582	7.441	7.441	0.195	0.195	5.243	5.243	23.860	23.860	6.755	6.755	24.176	24.176	4.064	4.064
10	287	MULTICOM-CLUSTER	57	57	12.416	12.416	42.418	42.418	7.538	7.538	0.212	0.212	5.960	5.960	24.700	24.700	9.917	9.917	35.198	35.198	5.965	5.965

Figure S10. Screenshot of the performance or precision of different methods for Top10 list size and 'long' range contacts on CASP12 TBM/FM and FM targets.



			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	079	iFold_1	54	57	8.013	7.591	54.074	51.228	4.606	4.364	0.105	0.099	4.100	3.885	16.468	15.601	7.879	7.464	53.462	50.648
2	451	RaptorX-Contact	57	57	7.474	7.474	51.930	51.930	4.187	4.187	0.113	0.113	4.159	4.159	13.384	13.384	5.382	5.382	40.132	40.132
3	013	MetaPSICOV	57	57	8.177	8.177	51.404	51.404	4.744	4.744	0.116	0.116	4.142	4.142	16.560	16.560	5.689	5.689	39.977	39.977
4	219	Deepfold-Contact	55	57	7.813	7.539	50.545	48.772	4.477	4.320	0.104	0.101	4.088	3.945	14.963	14.438	7.669	7.400	49.970	48.216
5	097	PconsC31	57	57	7.177	7.177	49.474	49.474	4.117	4.117	0.105	0.105	3.629	3.629	15.079	15.079	3.483	3.483	29.307	29.307
6	109	naive	55	57	7.595	7.329	48.545	46.842	4.384	4.230	0.113	0.109	4.473	4.316	17.346	16.737	7.435	7.174	47.837	46.158
7	287	MULTICOM-CLUSTER	57	57	6.939	6.939	48.246	48.246	3.963	3.963	0.099	0.099	3.358	3.358	16.164	16.164	5.646	5.646	42.090	42.090
8	320	raghavagps	54	57	7.940	7.522	48.148	45.614	4.646	4.401	0.147	0.139	4.299	4.073	14.599	13.831	6.135	5.812	38.467	36.442

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S11. Screenshot of the performance or precision of different methods for L/5 list size and 'long' range contacts on CASP12 TBM and FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results | RR Assessm

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long | List size: L/5 | Prob filter: >0 | Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)		Prec-FDR(prob)		F
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	S
1	451	RaptorX-Contact	75	75	16.039	16.039	61.722	61.722	9.395	9.395	0.395	0.395	10.987	10.987	26.602	26.602	12.936	12.936	51.097	51.097	7.535	7.535	29.847	29.847	62.
2	219	Deepfold-Contact	73	75	16.015	15.588	61.389	59.752	9.396	9.146	0.446	0.434	11.452	11.147	27.469	26.737	15.812	15.390	60.797	59.176	9.270	9.023	23.985	23.346	61.
3	109	naive	73	75	15.789	15.368	61.084	59.455	9.237	8.991	0.492	0.479	11.902	11.584	29.287	28.506	15.556	15.141	60.302	58.694	9.098	8.855	23.153	22.536	61.
4	013	MetaPSICOV	73	75	14.243	13.864	57.228	55.702	8.260	8.040	0.384	0.374	11.040	10.746	24.098	23.455	11.609	11.299	47.962	46.683	6.698	6.519	29.657	28.867	57.
5	357	FLOUDAS_SERVER	63	75	13.634	11.453	54.687	45.937	7.950	6.678	0.438	0.368	10.411	8.745	19.914	16.728	10.171	8.544	42.498	35.699	5.869	4.930	34.032	28.587	55.
6	097	PconsC31	75	75	13.575	13.575	53.509	53.509	7.897	7.897	0.326	0.326	8.836	8.836	22.871	22.871	8.488	8.488	34.382	34.382	4.910	4.910	15.414	15.414	53.
7	079	iFold_1	49	75	14.311	9.350	53.045	34.656	8.428	5.506	0.299	0.196	8.395	5.484	25.789	16.849	14.046	9.176	52.266	34.147	8.264	5.399	7.957	5.199	53.
8	432	Pcons-net	75	75	13.424	13.424	52.880	52.880	7.809	7.809	0.320	0.320	9.063	9.063	23.962	23.962	8.496	8.496	34.172	34.172	4.924	4.924	13.848	13.848	53.
9	044	Yang-Server	73	75	13.515	13.155	52.707	51.301	7.882	7.672	0.399	0.388	10.359	10.083	24.915	24.251	9.622	9.365	38.796	37.762	5.569	5.420	17.642	17.172	53.
10	287	MULTICOM-CLUSTER	75	75	12.953	12.953	51.727	51.727	7.510	7.510	0.359	0.359	9.791	9.791	24.363	24.363	11.342	11.342	46.114	46.114	6.551	6.551	14.345	14.345	52.
11	236	MULTICOM-CONSTRUCT	75	75	12.909	12.909	51.604	51.604	7.483	7.483	0.358	0.358	9.778	9.778	24.185	24.185	9.795	9.795	40.555	40.555	5.636	5.636	21.185	21.185	52.
12	310	IGBteam	75	75	11.756	11.756	47.819	47.819	6.785	6.785	0.297	0.297	8.585	8.585	18.390	18.390	9.348	9.348	38.980	38.980	5.367	5.367	7.406	7.406	48.
13	407	Distill	75	75	12.199	12.199	47.290	47.290	7.123	7.123	0.426	0.426	10.541	10.541	23.560	23.560	10.007	10.007	39.545	39.545	5.825	5.825	12.228	12.228	47.
14	020	RBO-Epsilon	75	75	12.649	12.649	47.210	47.210	7.486	7.486	0.421	0.421	11.218	11.218	28.761	28.761	7.635	7.635	28.896	28.896	4.524	4.524	9.358	9.358	49.
15	373	Zhang_Contact	73	75	11.891	11.574	46.849	45.599	6.948	6.763	0.333	0.324	9.489	9.236	22.886	22.276	8.808	8.573	35.755	34.802	5.115	4.979	5.689	5.537	47.
16	348	FALCON_COLORS	75	75	11.814	11.814	45.976	45.976	6.897	6.897	0.331	0.331	9.918	9.918	24.067	24.067	7.882	7.882	31.884	31.884	4.564	4.564	9.945	9.945	46.
17	320	raghavagps	50	75	13.140	8.760	45.398	30.265	7.874	5.250	0.377	0.251	9.067	6.044	19.883	13.255	10.152	6.768	35.317	23.545	6.083	4.056	1.811	1.208	46.
18	024	PconsC2	65	75	10.727	9.297	44.918	38.929	6.151	5.331	0.374	0.325	9.445	8.186	22.186	19.228	4.264	3.695	18.817	16.308	2.419	2.097	12.820	11.111	48.
19	251	myprotein-me	53	75	10.313	7.288	41.699	29.467	6.010	4.247	0.248	0.175	7.057	4.987	16.805	11.875	3.358	2.373	14.673	10.369	1.922	1.358	2.914	2.059	43.

Figure S12. Screenshot of the performance or precision of different methods for Top10 list size and 'long' range contacts on CASP12 TBM and FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results | RR Assessm

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long | List size: Top 10 | Prob filter: >0 | Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)		Prec-FDR(prob)		
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	
1	451	RaptorX-Contact	75	75	6.589	6.589	68.800	68.800	3.567	3.567	0.157	0.157	6.343	6.343	13.954	13.954	5.244	5.244	58.750	58.750	2.825	2.825	40.888	40.888	68.
2	109	naive	73	75	6.208	6.042	66.301	64.533	3.359	3.269	0.176	0.172	6.986	6.799	16.227	15.794	6.134	5.971	65.812	64.057	3.318	3.229	33.127	32.244	66.
3	219	Deepfold-Contact	73	75	6.348	6.179	65.890	64.133	3.446	3.355	0.153	0.149	6.274	6.106	13.932	13.561	6.284	6.116	65.515	63.768	3.410	3.319	32.390	31.527	65.
4	013	MetaPSICOV	73	75	5.449	5.303	61.370	59.733	2.932	2.854	0.142	0.138	6.017	5.856	12.178	11.853	4.515	4.395	54.915	53.450	2.421	2.357	37.138	36.148	61.
5	044	Yang-Server	73	75	5.431	5.286	59.315	57.733	2.928	2.850	0.168	0.164	6.271	6.103	14.482	14.095	3.912	3.808	47.102	45.846	2.100	2.044	25.820	25.131	59.
6	097	PconsC31	75	75	4.974	4.974	59.200	59.200	2.650	2.650	0.118	0.118	4.621	4.621	10.751	10.751	3.293	3.293	42.544	42.544	1.743	1.743	24.613	24.613	59.
7	432	Pcons-net	75	75	5.081	5.081	59.200	59.200	2.711	2.711	0.117	0.117	4.434	4.434	11.839	11.839	3.437	3.437	42.910	42.910	1.827	1.827	24.710	24.710	59.
8	287	MULTICOM-CLUSTER	75	75	5.050	5.050	57.200	57.200	2.698	2.698	0.137	0.137	5.254	5.254	12.847	12.847	4.454	4.454	52.838	52.838	2.374	2.374	23.722	23.722	57.
9	079	iFold_1	49	75	6.406	4.185	56.939	37.200	3.476	2.271	0.108	0.070	4.398	2.873	13.087	8.550	6.323	4.131	56.482	36.901	3.431	2.241	15.146	9.895	56.
10	236	MULTICOM-CONSTRUCT	75	75	5.034	5.034	56.800	56.800	2.690	2.690	0.138	0.138	5.376	5.376	12.948	12.948	3.886	3.886	48.123	48.123	2.066	2.066	29.189	29.189	57.
11	357	FLOUDAS_SERVER	66	75	4.911	4.322	55.758	49.067	2.690	2.367	0.139	0.123	4.952	4.358	9.562	8.414	3.286	2.892	43.375	38.170	1.780	1.567	35.227	31.000	55.
12	020	RBO-Epsilon	75	75	5.544	5.544	54.933	54.933	3.053	3.053	0.162	0.162	5.859	5.859	15.446	15.446	3.683	3.683	40.165	40.165	2.036	2.036	20.031	20.031	55.
13	310	IGBteam	75	75	4.806	4.806	53.867	53.867	2.578	2.578	0.119	0.119	5.058	5.058	10.480	10.480	3.797	3.797	46.061	46.061	2.027	2.027	15.889	15.889	54.
14	348	FALCON_COLORS	75	75	4.773	4.773	52.800	52.800	2.575	2.575	0.131	0.131	5.602	5.602	11.861	11.861	3.264	3.264	40.735	40.735	1.751	1.751	18.881	18.881	52.
15	024	PconsC2	65	75	4.547	3.941	51.538	44.667	2.434	2.109	0.149	0.129	6.349	5.503	12.721	11.025	1.941	1.682	26.941	23.349	1.028	0.891	20.417	17.695	52.
16	373	Zhang_Contact	73	75	4.429	4.311	50.411	49.067	2.395	2.332	0.125	0.121	5.070	4.935	12.505	12.172	3.452	3.359	42.165	41.041	1.859	1.809	11.377	11.073	50.
17	407	Distill	75	75	4.273	4.273	49.733	49.733	2.302	2.302	0.159	0.159	6.745	6.745	13.698	13.698	3.541	3.541	43.599	43.599	1.904	1.904	13.639	13.639	49.
18	320	raghavagps	50	75	5.812	3.875	48.600	32.400	3.195	2.130	0.165	0.110	5.970	3.980	10.258	6.838	4.743	3.162	40.837	27.224	2.609	1.739	5.530	3.687	49.
19	251	myprotein-me	53	75	3.912	2.764	47.736	33.733	2.095	1.480	0.091	0.065	4.063	2.871	8.583	6.066	1.364	0.964	19.550	13.815	0.720	0.509	7.859	5.554	48.

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S13. Screenshot of the performance or precision of different methods for L/5 list size and 'long' range contacts on CASP12 all category targets (TBM, FM and TBM/FM).

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long | List size: L/5 | Prob filter: >0 | Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	219	Deepfold-Contact	92	94	16.289	15.943	59.249	57.989	9.750	9.542	0.418	0.409	10.562	10.337	27.939	27.345	16.031	15.690	58.552	57.306	9.585	9.381
2	109	naive	92	94	16.290	15.944	59.205	57.946	9.756	9.549	0.455	0.446	10.984	10.751	29.587	28.958	15.997	15.657	58.328	57.087	9.573	9.370
3	451	RaptorX-Contact	94	94	16.012	16.012	58.661	58.661	9.543	9.543	0.372	0.372	10.314	10.314	25.507	25.507	12.417	12.417	47.377	47.377	7.321	7.321
4	013	MetaPSICOV	92	94	15.312	14.986	56.511	55.308	9.155	8.960	0.364	0.356	10.273	10.055	25.907	25.355	11.898	11.645	45.969	44.991	7.037	6.887
5	079	iFold_1	66	94	15.695	11.020	54.352	38.162	9.508	6.676	0.310	0.218	8.347	5.861	27.740	19.477	15.356	10.782	53.424	37.510	9.293	6.525
6	097	PconsC31	94	94	14.361	14.361	53.273	53.273	8.550	8.550	0.321	0.321	8.446	8.446	23.923	23.923	8.300	8.300	32.447	32.447	4.874	4.874
7	357	FLOUDAS_SERVER	81	94	13.793	11.886	51.935	44.753	8.243	7.103	0.403	0.347	9.574	8.250	20.161	17.373	9.405	8.104	38.395	33.085	5.479	4.721
8	320	raghavagps	66	94	16.104	11.307	51.906	36.445	9.957	6.991	0.394	0.276	8.913	6.258	25.261	17.737	12.499	8.776	40.813	28.656	7.714	5.416
9	432	Pcons-net	94	94	13.740	13.740	51.433	51.433	8.149	8.149	0.302	0.302	8.352	8.352	24.029	24.029	8.393	8.393	32.418	32.418	4.949	4.949
10	236	MULTICOM-CONSTRUCT	94	94	13.462	13.462	50.307	50.307	7.999	7.999	0.337	0.337	9.129	9.129	24.524	24.524	9.481	9.481	37.538	37.538	5.551	5.551

Figure S14. Screenshot of the performance or precision of different methods for Top10 list size and 'long' range contacts on CASP12 all category targets (TBM, FM and TBM/FM).

predictioncenter.org/casp12/rrc_avrg_results.cgi

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long | List size: Top 10 | Prob filter: >0 | Text file

Domains classification

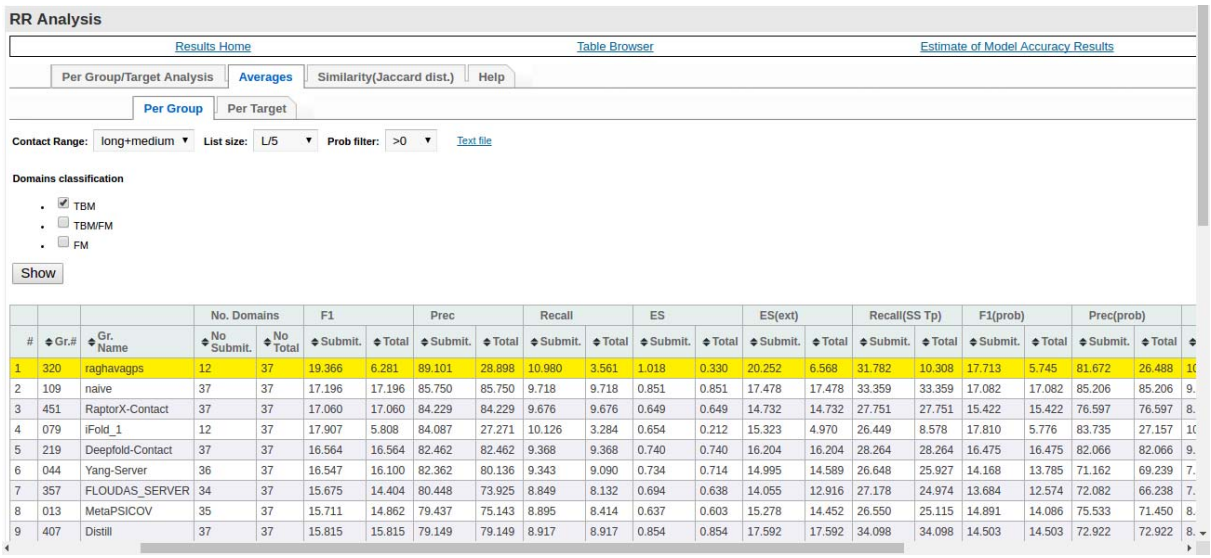
- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	451	RaptorX-Contact	94	94	7.079	7.079	64.468	64.468	3.922	3.922	0.150	0.150	5.924	5.924	13.565	13.565	5.390	5.390	53.954	53.954	2.956	2.956
2	109	naive	92	94	7.266	7.111	63.804	62.447	4.096	4.009	0.170	0.166	6.671	6.529	17.039	16.677	7.148	6.996	63.229	61.884	4.026	3.940
3	219	Deepfold-Contact	92	94	7.224	7.071	63.370	62.021	4.061	3.974	0.151	0.148	6.068	5.939	15.147	14.825	7.117	6.966	62.905	61.566	3.995	3.910
4	013	MetaPSICOV	92	94	7.077	6.926	60.978	59.681	4.022	3.936	0.147	0.144	5.894	5.769	15.113	14.792	5.389	5.275	52.910	51.785	3.011	2.947
5	097	PconsC31	94	94	6.340	6.340	59.681	59.681	3.545	3.545	0.126	0.126	4.774	4.774	12.923	12.923	3.572	3.572	40.331	40.331	1.946	1.946
6	079	iFold_1	66	94	7.749	5.441	58.333	40.957	4.414	3.099	0.119	0.084	4.737	3.326	15.680	11.010	7.621	5.351	57.764	40.558	4.337	3.045
7	432	Pcons-net	94	94	5.922	5.922	57.447	57.447	3.276	3.276	0.113	0.113	4.238	4.238	12.631	12.631	3.743	3.743	40.432	40.432	2.057	2.057
8	020	RBO-Epsilon	94	94	7.147	7.147	55.426	55.426	4.120	4.120	0.162	0.162	5.577	5.577	16.185	16.185	4.428	4.428	38.965	38.965	2.531	2.531
9	287	MULTICOM-CLUSTER	94	94	6.005	6.005	55.106	55.106	3.357	3.357	0.134	0.134	5.114	5.114	14.161	14.161	5.075	5.075	50.113	50.113	2.814	2.814
10	236	MULTICOM-CONSTRUCT	94	94	6.008	6.008	54.894	54.894	3.359	3.359	0.135	0.135	5.212	5.212	14.242	14.242	4.166	4.166	44.627	44.627	2.282	2.282
11	044	Yang-Server	92	94	5.692	5.571	54.565	53.404	3.157	3.089	0.150	0.147	5.577	5.458	13.252	12.970	3.614	3.537	41.164	40.289	1.960	1.918
12	320	raghavagps	66	94	8.100	5.687	54.545	38.298	4.698	3.299	0.181	0.127	6.063	4.257	15.019	10.545	6.504	4.567	45.937	32.254	3.753	2.635
13	357	FLOUDAS_SERVER	84	94	5.379	4.807	53.333	47.660	3.024	2.703	0.135	0.121	4.821	4.308	10.199	9.114	3.247	2.902	40.022	35.764	1.789	1.598

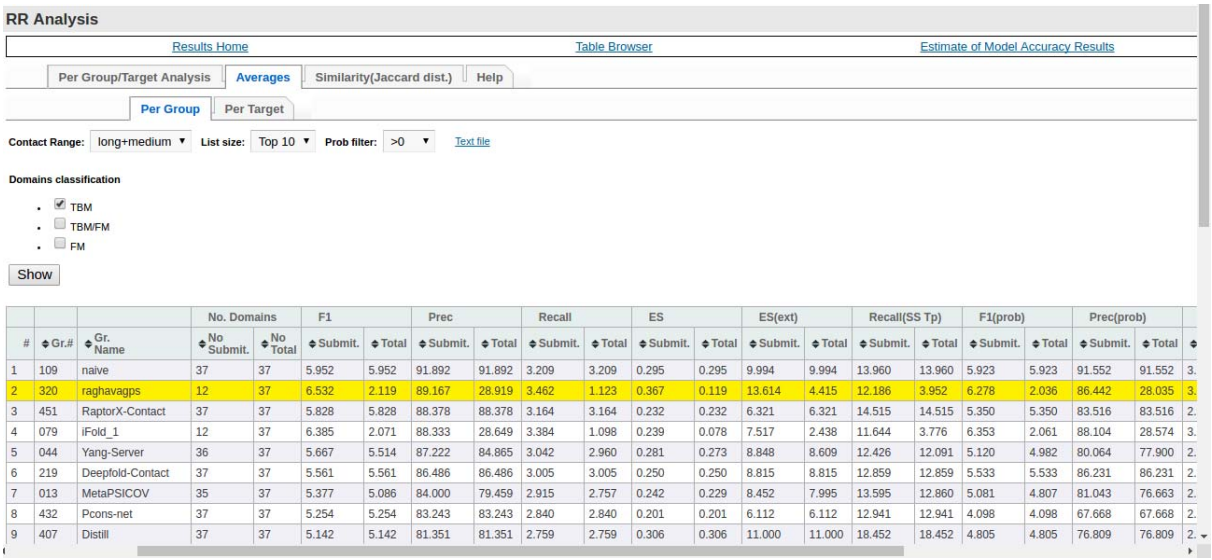
bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S15. Screenshot of the performance or precision of different methods for L/5 list size and 'long+medium' range contacts on CASP12 TBM targets.



			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	320	raghavagps	12	37	19.366	6.281	89.101	28.898	10.980	3.561	1.018	0.330	20.252	6.568	31.782	10.308	17.713	5.745	81.672	26.488
2	109	naive	37	37	17.196	17.196	85.750	85.750	9.718	9.718	0.851	0.851	17.478	17.478	33.359	33.359	17.082	17.082	85.206	85.206
3	451	RaptorX-Contact	37	37	17.060	17.060	84.229	84.229	9.676	9.676	0.649	0.649	14.732	14.732	27.751	27.751	15.422	15.422	76.597	76.597
4	079	iFold_1	12	37	17.907	5.808	84.087	27.271	10.126	3.284	0.654	0.212	15.323	4.970	26.449	8.578	17.810	5.776	83.735	27.157
5	219	Deepfold-Contact	37	37	16.564	16.564	82.462	82.462	9.368	9.368	0.740	0.740	16.204	16.204	28.264	28.264	16.475	16.475	82.066	82.066
6	044	Yang-Server	36	37	16.547	16.100	82.362	80.136	9.343	9.090	0.734	0.714	14.995	14.589	26.648	25.927	14.168	13.785	71.162	69.239
7	357	FLOUDAS_SERVER	34	37	15.675	14.404	80.448	73.925	8.849	8.132	0.694	0.638	14.055	12.916	27.178	24.974	13.684	12.574	72.082	66.238
8	013	MetaPSICOV	35	37	15.711	14.862	79.437	75.143	8.895	8.414	0.637	0.603	15.278	14.452	26.550	25.115	14.891	14.086	75.533	71.450
9	407	Distill	37	37	15.815	15.815	79.149	79.149	8.917	8.917	0.854	0.854	17.592	17.592	34.098	34.098	14.503	14.503	72.922	72.922

Figure S16. Screenshot of the performance or precision of different methods for Top10 list size and 'long+medium' range contacts on CASP12 TBM targets.



			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	109	naive	37	37	5.952	5.952	91.892	91.892	3.209	3.209	0.295	0.295	9.994	9.994	13.960	13.960	5.923	5.923	91.552	91.552
2	320	raghavagps	12	37	6.532	2.119	89.167	28.919	3.462	1.123	0.367	0.119	13.614	4.415	12.186	3.952	6.278	2.036	86.442	28.035
3	451	RaptorX-Contact	37	37	5.828	5.828	88.378	88.378	3.164	3.164	0.232	0.232	6.321	6.321	14.515	14.515	5.350	5.350	83.516	83.516
4	079	iFold_1	12	37	6.385	2.071	88.333	28.649	3.384	1.098	0.239	0.078	7.517	2.438	11.644	3.776	6.353	2.061	88.104	28.574
5	044	Yang-Server	36	37	5.667	5.514	87.222	84.865	3.042	2.960	0.281	0.273	8.848	8.609	12.426	12.091	5.120	4.982	80.064	77.900
6	219	Deepfold-Contact	37	37	5.561	5.561	86.486	86.486	3.005	3.005	0.250	0.250	8.815	8.815	12.859	12.859	5.533	5.533	86.231	86.231
7	013	MetaPSICOV	35	37	5.377	5.086	84.000	79.459	2.915	2.757	0.242	0.229	8.452	7.995	13.595	12.860	5.081	4.807	81.043	76.663
8	432	Pcons-net	37	37	5.254	5.254	83.243	83.243	2.840	2.840	0.201	0.201	6.112	6.112	12.941	12.941	4.098	4.098	67.668	67.668
9	407	Distill	37	37	5.142	5.142	81.351	81.351	2.759	2.759	0.306	0.306	11.000	11.000	18.452	18.452	4.805	4.805	76.809	76.809

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S17. Screenshot of the performance or precision of different methods for L/5 list size and 'long+medium' range contacts on CASP12 TBM/FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long+medium List size: L/5 Prob filter: >0 Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		R
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	S
1	320	raghavagps	16	19	20.866	17.572	78.378	66.003	12.741	10.729	0.542	0.457	9.211	7.756	39.441	33.214	16.987	14.305	64.627	54.423	10.
2	013	MetaPSICOV	19	19	17.891	17.891	69.109	69.109	10.634	10.634	0.377	0.377	7.493	7.493	27.311	27.311	14.254	14.254	55.702	55.702	8.4
3	451	RaptorX-Contact	19	19	16.726	16.726	65.995	65.995	9.881	9.881	0.309	0.309	6.455	6.455	23.321	23.321	13.103	13.103	53.455	53.455	7.6
4	079	iFold_1	17	19	16.222	14.515	65.896	58.960	9.656	8.640	0.369	0.331	7.291	6.524	28.616	25.604	16.015	14.329	65.092	58.241	9.5
5	219	Deepfold-Contact	19	19	16.776	16.776	65.461	65.461	10.006	10.006	0.380	0.380	7.069	7.069	25.460	25.460	16.314	16.314	63.798	63.798	9.7
6	109	naive	19	19	16.629	16.629	64.436	64.436	9.929	9.929	0.386	0.386	7.590	7.590	26.644	26.644	16.122	16.122	62.610	62.610	9.6
7	020	RBO-Epsilon	19	19	16.818	16.818	62.587	62.587	10.121	10.121	0.321	0.321	6.386	6.386	24.699	24.699	11.915	11.915	45.017	45.017	7.1
8	097	PconsC31	19	19	15.067	15.067	61.171	61.171	8.917	8.917	0.337	0.337	6.344	6.344	22.309	22.309	9.328	9.328	39.296	39.296	5.4
9	236	MULTICOM-	19	19	15.524	15.524	61.046	61.046	9.280	9.280	0.314	0.314	6.281	6.281	22.318	22.318	11.425	11.425	46.794	46.794	6.7

Figure S18. Screenshot of the performance or precision of different methods for Top10 list size and 'long+medium' range contacts on CASP12 TBM/FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long+medium List size: Top 10 Prob filter: >0 Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	
1	320	raghavagps	16	19	11.943	10.057	79.375	66.842	6.809	5.734	0.268	0.225	6.523	5.493	25.506	21.478	9.955	8.383	68.793	57.931	5.
2	013	MetaPSICOV	19	19	11.196	11.196	74.737	74.737	6.265	6.265	0.189	0.189	4.784	4.784	19.004	19.004	9.222	9.222	63.298	63.298	5.
3	079	iFold_1	17	19	9.286	8.308	73.529	65.789	5.187	4.641	0.189	0.169	5.086	4.551	17.154	15.348	9.201	8.232	72.975	65.294	5.
4	219	Deepfold-Contact	19	19	10.477	10.477	72.632	72.632	5.891	5.891	0.194	0.194	4.792	4.792	16.870	16.870	10.112	10.112	70.983	70.983	5.
5	451	RaptorX-Contact	19	19	10.770	10.770	72.632	72.632	6.013	6.013	0.166	0.166	3.390	3.390	15.020	15.020	8.545	8.545	61.116	61.116	4.
6	020	RBO-Epsilon	19	19	11.315	11.315	71.579	71.579	6.390	6.390	0.179	0.179	3.274	3.274	18.440	18.440	8.656	8.656	57.078	57.078	4.
7	109	naive	19	19	10.419	10.419	71.053	71.053	5.860	5.860	0.203	0.203	5.032	5.032	18.706	18.706	10.025	10.025	69.264	69.264	5.
8	097	PconsC31	19	19	9.120	9.120	70.526	70.526	5.049	5.049	0.186	0.186	4.395	4.395	17.079	17.079	6.002	6.002	48.832	48.832	3.
9	432	Pcons-net	19	19	10.374	10.374	69.474	69.474	5.863	5.863	0.181	0.181	4.071	4.071	17.534	17.534	6.940	6.940	49.990	49.990	3.

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S19. Screenshot of the performance or precision of different methods for L/5 list size and 'long+medium' range contacts on CASP12 FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results | RR Assess

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long+medium | List size: L/5 | Prob filter: >0 | Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)		Prec-FDR(prob)		
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	
1	451	RaptorX-Contact	38	38	12.386	12.386	55.831	55.831	7.029	7.029	0.252	0.252	5.604	5.604	22.282	22.282	9.551	9.551	44.436	44.436	5.391	5.391	18.254	18.254	56
2	079	iFold_1	37	38	11.212	10.917	52.292	50.916	6.320	6.153	0.236	0.230	5.667	5.518	18.875	18.378	11.070	10.779	51.744	50.382	6.238	6.074	6.321	6.154	52
3	013	MetaPSICOV	38	38	10.919	10.919	51.307	51.307	6.155	6.155	0.254	0.254	6.361	6.361	20.123	20.123	7.857	7.857	38.265	38.265	4.403	4.403	16.788	16.788	52
4	236	MULTICOM-CONSTRUCT	38	38	10.847	10.847	50.449	50.449	6.121	6.121	0.244	0.244	5.139	5.139	20.528	20.528	7.661	7.661	36.717	36.717	4.303	4.303	14.280	14.280	51
5	432	Pcons-net	38	38	10.810	10.810	49.536	49.536	6.113	6.113	0.209	0.209	4.602	4.602	17.174	17.174	6.773	6.773	31.993	31.993	3.810	3.810	7.447	7.447	50
6	219	Deepfold-Contact	36	38	10.559	10.003	49.023	46.442	5.958	5.644	0.250	0.237	6.120	5.798	18.710	17.725	10.431	9.882	48.539	45.984	5.884	5.574	-0.471	-0.446	49
7	020	RBO-Epsilon	38	38	10.736	10.736	48.601	48.601	6.081	6.081	0.230	0.230	4.779	4.779	20.528	20.528	7.095	7.095	32.812	32.812	4.007	4.007	8.757	8.757	50
8	044	Yang-Server	37	38	10.462	10.186	47.715	46.460	5.918	5.763	0.210	0.205	4.833	4.706	18.174	17.696	6.389	6.221	30.139	29.346	3.599	3.504	0.788	0.767	49
9	348	FALCON_COLORS	38	38	10.387	10.387	47.253	47.253	5.880	5.880	0.236	0.236	5.723	5.723	18.441	18.441	6.479	6.479	30.759	30.759	3.645	3.645	6.442	6.442	48
10	097	PconsC31	38	38	9.734	9.734	45.728	45.728	5.483	5.483	0.191	0.191	3.759	3.759	14.637	14.637	6.127	6.127	29.723	29.723	3.435	3.435	5.367	5.367	46
11	310	IGBteam	38	38	9.428	9.428	45.596	45.596	5.288	5.288	0.178	0.178	3.838	3.838	16.433	16.433	7.284	7.284	35.639	35.639	4.079	4.079	-1.447	-1.447	46
12	109	naive	36	38	9.517	9.016	44.343	42.009	5.367	5.085	0.251	0.237	6.130	5.808	19.795	18.753	9.371	8.878	43.732	41.431	5.284	5.006	-9.637	-9.130	44
13	389	AkbAR	36	38	9.396	8.901	43.718	41.417	5.300	5.021	0.270	0.256	5.968	5.654	21.897	20.745	5.039	4.773	24.717	23.416	2.821	2.673	7.744	7.336	46
14	287	MULTICOM-CLUSTER	38	38	9.114	9.114	42.661	42.661	5.138	5.138	0.205	0.205	5.385	5.385	17.395	17.395	7.680	7.680	36.733	36.733	4.316	4.316	-0.341	-0.341	43
15	431	Shen-Group	38	38	8.984	8.984	40.938	40.938	5.085	5.085	0.255	0.255	6.659	6.659	20.402	20.402	7.243	7.243	33.073	33.073	4.099	4.099	-13.457	-13.457	41
16	373	Zhang_Contact	36	38	8.561	8.111	40.385	38.259	4.824	4.570	0.206	0.195	5.851	5.543	17.584	16.658	5.485	5.196	26.679	25.274	3.077	2.915	-2.064	-1.955	40
17	320	raghavagps	38	38	9.005	9.005	40.269	40.269	5.110	5.110	0.270	0.270	5.352	5.352	14.582	14.582	6.606	6.606	29.598	29.598	3.746	3.746	-11.414	-11.414	40
18	024	PconsC2	36	38	7.755	7.347	36.919	34.976	4.361	4.132	0.211	0.200	4.951	4.690	17.073	16.174	2.188	2.072	10.931	10.356	1.220	1.156	7.617	7.216	40
19	345	MULTICOM-NOVEL	38	38	8.235	8.235	36.630	36.630	4.680	4.680	0.191	0.191	3.755	3.755	18.887	18.887	7.054	7.054	31.687	31.687	4.003	4.003	-21.907	-21.907	36

Figure S20. Screenshot of the performance or precision of different methods for Top10 list size and 'long+medium' range contacts on CASP12 FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results | RR Assessn

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long+medium | List size: Top 10 | Prob filter: >0 | Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)		Prec-FDR(prob)		
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	
1	451	RaptorX-Contact	38	38	6.150	6.150	63.158	63.158	3.264	3.264	0.125	0.125	2.860	2.860	11.804	11.804	4.870	4.870	52.436	52.436	2.578	2.578	28.787	28.787	63
2	236	MULTICOM-CONSTRUCT	38	38	5.305	5.305	58.947	58.947	2.806	2.806	0.113	0.113	2.769	2.769	12.752	12.752	3.996	3.996	46.889	46.889	2.107	2.107	26.267	26.267	59
3	013	MetaPSICOV	38	38	5.362	5.362	58.684	58.684	2.839	2.839	0.109	0.109	2.828	2.828	11.039	11.039	4.214	4.214	47.765	47.765	2.226	2.226	26.380	26.380	59
4	020	RBO-Epsilon	38	38	5.396	5.396	57.632	57.632	2.863	2.863	0.114	0.114	2.264	2.264	13.601	13.601	4.113	4.113	46.421	46.421	2.177	2.177	23.543	23.543	58
5	079	iFold_1	37	38	5.561	5.414	57.027	55.526	2.950	2.873	0.111	0.108	3.302	3.215	10.854	10.568	5.516	5.371	56.664	55.173	2.927	2.850	15.128	14.730	57
6	432	Pcons-net	38	38	5.006	5.006	55.000	55.000	2.649	2.649	0.099	0.099	2.572	2.572	10.537	10.537	3.381	3.381	38.989	38.989	1.784	1.784	15.516	15.516	55
7	219	Deepfold-Contact	36	38	5.422	5.137	54.444	51.579	2.879	2.727	0.108	0.102	3.198	3.029	10.134	9.601	5.375	5.092	54.069	51.224	2.854	2.703	9.571	9.067	54
8	389	AkbAR	36	38	5.043	4.778	53.333	50.526	2.673	2.532	0.139	0.132	3.371	3.194	15.781	14.951	3.083	2.921	34.592	32.772	1.627	1.541	13.945	13.211	53
9	044	Yang-Server	37	38	5.005	4.873	52.973	51.579	2.654	2.584	0.102	0.100	2.501	2.435	10.151	9.884	3.454	3.363	38.156	37.152	1.828	1.780	7.370	7.176	53
10	348	FALCON_COLORS	38	38	4.903	4.903	52.895	52.895	2.599	2.599	0.110	0.110	2.833	2.833	11.106	11.106	3.448	3.448	39.186	39.186	1.822	1.822	14.517	14.517	53
11	310	IGBteam	38	38	4.679	4.679	51.053	51.053	2.478	2.478	0.104	0.104	2.813	2.813	13.526	13.526	3.785	3.785	42.139	42.139	2.002	2.002	4.801	4.801	51
12	097	PconsC31	38	38	4.546	4.546	50.526	50.526	2.401	2.401	0.093	0.093	2.269	2.269	8.862	8.862	3.054	3.054	36.151	36.151	1.607	1.607	11.823	11.823	50
13	431	Shen-Group	38	38	4.605	4.605	50.263	50.263	2.440	2.440	0.122	0.122	3.571	3.571	14.203	14.203	3.905	3.905	43.247	43.247	2.067	2.067	1.202	1.202	50
14	287	MULTICOM-CLUSTER	38	38	4.435	4.435	49.474	49.474	2.345	2.345	0.099	0.099	2.749	2.749	10.412	10.412	3.820	3.820	44.371	44.371	2.015	2.015	11.039	11.039	49
15	109	naive	36	38	4.667	4.421	48.889	46.316	2.473	2.343	0.112	0.106	3.575	3.386	12.253	11.608	4.610	4.368	48.426	45.877	2.443	2.315	-1.393	-1.320	48
16	320	raghavagps	38	38	4.254	4.254	46.842	46.842	2.252	2.252	0.129	0.129	3.454	3.454	7.721	7.721	3.460	3.460	38.674	38.674	1.832	1.832	-1.857	-1.857	47
17	345	MULTICOM-NOVEL	38	38	4.405	4.405	45.263	45.263	2.340	2.340	0.094	0.094	1.876	1.876	11.527	11.527	3.854	3.854	40.208	40.208	2.045	2.045	-7.636	-7.636	45
18	281	PLCT	38	38	4.040	4.040	44.211	44.211	2.139	2.139	0.080	0.080	1.510	1.510	7.624	7.624	2.002	2.002	22.885	22.885	1.057	1.057	1.690	1.690	45
19	373	Zhang_Contact	36	38	3.815	3.614	43.056	40.789	2.017	1.911	0.100	0.095	3.328	3.152	10.656	10.096	2.753	2.608	32.000	30.316	1.454	1.377	2.128	2.016	42

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S21. Screenshot of the performance or precision of different methods for L/5 list size and 'long+medium' range contacts on CASP12 TBM and TBM/FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long+medium List size: L/5 Prob filter: >0 Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	320	raghavagps	28	56	20.223	10.112	82.974	41.487	11.986	5.993	0.746	0.373	13.942	6.971	36.159	18.079	17.298	8.649	71.932	35.966	10.220	5.110
2	109	naive	56	56	17.004	17.004	78.518	78.518	9.790	9.790	0.693	0.693	14.123	14.123	31.081	31.081	16.756	16.756	77.539	77.539	9.646	9.646
3	451	RaptorX-Contact	56	56	16.947	16.947	78.042	78.042	9.745	9.745	0.534	0.534	11.923	11.923	26.248	26.248	14.635	14.635	68.745	68.745	8.378	8.378
4	219	Deepfold-Contact	56	56	16.636	16.636	76.694	76.694	9.585	9.585	0.618	0.618	13.105	13.105	27.312	27.312	16.420	16.420	75.868	75.868	9.458	9.458
5	013	MetaPSICOV	54	56	16.478	15.889	75.803	73.096	9.507	9.167	0.546	0.526	12.539	12.091	26.818	25.860	14.667	14.143	68.555	66.107	8.443	8.141
6	044	Yang-Server	55	56	15.803	15.521	74.161	72.836	9.050	8.888	0.588	0.577	11.589	11.382	24.830	24.387	12.160	11.943	58.510	57.465	6.922	6.799
7	079	iFold_1	29	56	16.920	8.762	73.423	38.023	9.851	5.101	0.487	0.252	10.615	5.497	27.719	14.355	16.758	8.678	72.807	37.703	9.756	5.052
8	097	PconsC31	56	56	15.144	15.144	70.187	70.187	8.711	8.711	0.470	0.470	10.243	10.243	22.561	22.561	10.362	10.362	49.239	49.239	5.932	5.932
9	432	Pcons-net	56	56	15.344	15.344	70.166	70.166	8.864	8.864	0.465	0.465	9.797	9.797	24.326	24.326	10.723	10.723	50.356	50.356	6.162	6.162
10	357	FLOUDAS_SERVER	52	56	14.510	13.474	69.539	64.572	8.342	7.746	0.568	0.528	11.559	10.734	24.201	22.472	11.292	10.485	57.778	53.651	6.361	5.906
11	236	MULTICOM-	56	56	14.510	14.510	67.031	67.031	8.370	8.370	0.507	0.507	11.349	11.349	23.083	23.083	11.825	11.825	56.355	56.355	6.777	6.777

Figure S22. Screenshot of the performance or precision of different methods for Top10 list size and 'long+medium' range contacts on CASP12 TBM and TBM/FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long+medium List size: Top 10 Prob filter: >0 Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	109	naive	56	56	7.468	7.468	84.821	84.821	4.108	4.108	0.264	0.264	8.310	8.310	15.570	15.570	7.315	7.315	83.990	83.990	4.021	4.021
2	320	raghavagps	28	56	9.624	4.812	83.571	41.786	5.375	2.687	0.310	0.155	9.562	4.781	19.797	9.899	8.379	4.190	76.357	38.178	4.664	2.332
3	451	RaptorX-Contact	56	56	7.505	7.505	83.036	83.036	4.131	4.131	0.210	0.210	5.327	5.327	14.686	14.686	6.434	6.434	75.916	75.916	3.526	3.526
4	219	Deepfold-Contact	56	56	7.229	7.229	81.786	81.786	3.984	3.984	0.231	0.231	7.450	7.450	14.220	14.220	7.086	7.086	81.058	81.058	3.903	3.903
5	013	MetaPSICOV	54	56	7.425	7.159	80.741	77.857	4.094	3.947	0.224	0.216	7.161	6.906	15.498	14.945	6.538	6.305	74.800	72.128	3.599	3.471
6	079	iFold_1	29	56	8.086	4.187	79.655	41.250	4.441	2.300	0.210	0.109	6.092	3.155	14.874	7.703	8.022	4.154	79.236	41.033	4.406	2.281
7	044	Yang-Server	55	56	6.709	6.589	78.909	77.500	3.667	3.601	0.236	0.232	6.856	6.734	13.846	13.599	5.384	5.288	67.717	66.508	2.926	2.874
8	432	Pcons-net	56	56	6.991	6.991	78.571	78.571	3.866	3.866	0.194	0.194	5.419	5.419	14.499	14.499	5.062	5.062	61.670	61.670	2.781	2.781
9	097	PconsC31	56	56	6.430	6.430	77.500	77.500	3.515	3.515	0.190	0.190	5.322	5.322	13.619	13.619	4.619	4.619	59.858	59.858	2.514	2.514
10	020	RBO-Epsilon	56	56	7.385	7.385	74.821	74.821	4.116	4.116	0.208	0.208	4.791	4.791	14.486	14.486	6.089	6.089	64.517	64.517	3.389	3.389
11	236	MULTICOM-	56	56	6.389	6.389	72.500	72.500	3.518	3.518	0.194	0.194	5.874	5.874	11.996	11.996	5.333	5.333	64.119	64.119	2.926	2.926

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S23. Screenshot of the performance or precision of different methods for L/5 list size and 'long+medium' range contacts on CASP12 TBM/FM and FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long+medium List size: L/5 Prob filter: >0 Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	451	RaptorX-Contact	57	57	13.833	13.833	59.219	59.219	7.980	7.980	0.271	0.271	5.887	5.887	22.628	22.628	10.735	10.735	47.442	47.442	6.152	6.152
2	013	MetaPSICOV	57	57	13.243	13.243	57.241	57.241	7.648	7.648	0.295	0.295	6.738	6.738	22.519	22.519	9.989	9.989	44.077	44.077	5.757	5.757
3	079	iFold_1	54	57	12.789	12.116	56.575	53.597	7.370	6.982	0.278	0.263	6.179	5.854	21.941	20.786	12.627	11.962	55.946	53.001	7.276	6.893
4	219	Deepfold-Contact	55	57	12.707	12.261	54.701	52.782	7.357	7.098	0.295	0.284	6.448	6.222	21.042	20.303	12.463	12.026	53.810	51.922	7.213	6.960
5	236	MULTICOM-CONSTRUCT	57	57	12.406	12.406	53.982	53.982	7.174	7.174	0.267	0.267	5.520	5.520	21.125	21.125	8.916	8.916	40.076	40.076	5.126	5.126
6	432	Pcons-net	57	57	12.387	12.387	53.276	53.276	7.174	7.174	0.246	0.246	4.967	4.967	19.920	19.920	7.860	7.860	35.024	35.024	4.525	4.525
7	020	RBO-Epsilon	57	57	12.763	12.763	53.263	53.263	7.427	7.427	0.260	0.260	5.315	5.315	21.918	21.918	8.701	8.701	36.881	36.881	5.051	5.051
8	320	raghavagps	54	57	12.519	11.861	51.561	48.847	7.371	6.983	0.351	0.332	6.495	6.153	21.948	20.793	9.682	9.172	39.977	37.873	5.704	5.404
9	044	Yang-Server	56	57	11.796	11.589	51.416	50.514	6.793	6.673	0.244	0.240	4.937	4.850	19.264	18.926	7.056	6.933	31.631	31.076	4.041	3.971
10	109	naive	55	57	11.974	11.554	51.284	49.485	6.943	6.699	0.297	0.287	6.634	6.402	22.161	21.383	11.703	11.293	50.254	48.490	6.785	6.547

Figure S24. Screenshot of the performance or precision of different methods for Top10 list size and 'long+medium' range contacts on CASP12 TBM/FM and FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long+medium List size: Top 10 Prob filter: >0 Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	451	RaptorX-Contact	57	57	7.690	7.690	66.316	66.316	4.180	4.180	0.139	0.139	3.037	3.037	12.876	12.876	6.095	6.095	55.329	55.329	3.298	3.298
2	013	MetaPSICOV	57	57	7.306	7.306	64.035	64.035	3.981	3.981	0.136	0.136	3.480	3.480	13.694	13.694	5.884	5.884	52.942	52.942	3.201	3.201
3	020	RBO-Epsilon	57	57	7.369	7.369	62.281	62.281	4.039	4.039	0.136	0.136	2.600	2.600	15.214	15.214	5.628	5.628	49.973	49.973	3.072	3.072
4	079	iFold_1	54	57	6.734	6.379	62.222	58.947	3.654	3.462	0.136	0.129	3.863	3.660	12.837	12.161	6.676	6.325	61.799	58.547	3.623	3.433
5	236	MULTICOM-CONSTRUCT	57	57	6.694	6.694	62.105	62.105	3.646	3.646	0.128	0.128	3.114	3.114	13.778	13.778	5.168	5.168	50.224	50.224	2.804	2.804
6	219	Deepfold-Contact	55	57	7.168	6.917	60.727	58.596	3.919	3.782	0.137	0.133	3.748	3.617	12.461	12.024	7.012	6.766	59.912	57.810	3.831	3.696
7	432	Pcons-net	57	57	6.795	6.795	59.825	59.825	3.720	3.720	0.126	0.126	3.071	3.071	12.869	12.869	4.567	4.567	42.656	42.656	2.490	2.490
8	097	PconsC31	57	57	6.071	6.071	57.193	57.193	3.284	3.284	0.124	0.124	2.978	2.978	11.601	11.601	4.036	4.036	40.378	40.378	2.176	2.176
9	348	FALCON_COLORS	57	57	6.560	6.560	57.193	57.193	3.578	3.578	0.133	0.133	3.247	3.247	13.331	13.331	4.652	4.652	42.227	42.227	2.525	2.525
10	109	naive	55	57	6.654	6.421	56.545	54.561	3.643	3.515	0.144	0.139	4.078	3.935	14.483	13.974	6.481	6.253	55.624	53.673	3.545	3.421
11	320	raghavagps	54	57	6.532	6.188	56.481	53.509	3.602	3.413	0.170	0.161	4.363	4.134	12.991	12.307	5.384	5.101	47.598	45.093	2.968	2.812
12	044	Yang-Server	56	57	6.253	6.143	56.429	55.439	3.399	3.339	0.119	0.117	2.698	2.651	12.317	12.101	4.278	4.203	40.248	39.542	2.315	2.275

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S25. Screenshot of the performance or precision of different methods for L/5 list size and 'long+medium' range contacts on CASP12 TBM and FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results | RR Assessn

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long+medium | List size: L/5 | Prob filter: >0 | Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)		Prec-FDR(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	451	RaptorX-Contact	75	75	14.692	14.692	69.840	69.840	8.335	8.335	0.448	0.448	10.107	10.107	24.980	24.980	12.447	12.447	60.302	60.302	7.043	7.043	41.241	41.241
2	219	Deepfold-Contact	73	75	13.602	13.240	65.972	64.212	7.687	7.482	0.498	0.485	11.231	10.932	23.552	22.924	13.494	13.134	65.532	63.785	7.624	7.421	32.710	31.838
3	109	naive	73	75	13.409	13.051	65.330	63.588	7.573	7.371	0.555	0.540	11.882	11.565	26.670	25.958	13.280	12.926	64.753	63.026	7.499	7.299	31.341	30.505
4	044	Yang-Server	73	75	13.463	13.104	64.801	63.073	7.607	7.404	0.469	0.456	9.844	9.582	22.353	21.757	10.225	9.952	50.369	49.026	5.763	5.610	29.343	28.561
5	013	MetaPSICOV	73	75	13.216	12.864	64.794	63.066	7.469	7.269	0.438	0.426	10.636	10.353	23.205	22.586	11.229	10.930	56.133	54.636	6.334	6.165	37.752	36.745
6	432	Pcons-net	75	75	12.997	12.997	62.097	62.097	7.360	7.360	0.373	0.373	8.203	8.203	20.428	20.428	8.896	8.896	43.401	43.401	5.023	5.023	23.512	23.512
7	236	MULTICOM-CONSTRUCT	75	75	12.397	12.397	60.145	60.145	7.000	7.000	0.423	0.423	9.487	9.487	21.982	21.982	9.816	9.816	48.828	48.828	5.525	5.525	28.828	28.828
8	097	PconsC31	75	75	12.422	12.422	60.079	60.079	7.023	7.023	0.362	0.362	7.946	7.946	18.610	18.610	8.478	8.478	41.870	41.870	4.781	4.781	22.006	22.006
9	079	iFold_1	49	75	12.851	8.396	60.079	39.251	7.252	4.738	0.338	0.221	8.032	5.248	20.730	13.543	12.721	8.311	59.578	38.924	7.176	4.689	21.464	14.023
10	020	RBO-Epsilon	75	75	12.411	12.411	58.190	58.190	7.066	7.066	0.422	0.422	9.041	9.041	22.746	22.746	9.261	9.261	43.571	43.571	5.285	5.285	21.073	21.073
11	348	FALCON_COLORS	75	75	11.895	11.895	56.557	56.557	6.737	6.737	0.393	0.393	9.527	9.527	21.055	21.055	8.823	8.823	42.994	42.994	4.985	4.985	20.740	20.740
12	357	FLOUDAS_SERVER	65	75	11.339	9.827	56.260	48.759	6.407	5.553	0.451	0.391	9.340	8.095	18.398	15.945	8.514	7.379	44.007	38.139	4.763	4.128	36.339	31.494
13	310	IGBteam	75	75	11.077	11.077	55.373	55.373	6.240	6.240	0.367	0.367	8.453	8.453	18.280	18.280	9.420	9.420	47.351	47.351	5.309	5.309	14.609	14.609
14	407	Distill	75	75	10.786	10.786	53.113	53.113	6.079	6.079	0.504	0.504	10.571	10.571	22.925	22.925	9.113	9.113	45.342	45.342	5.126	5.126	16.196	16.196
15	287	MULTICOM-CLUSTER	75	75	10.546	10.546	53.080	53.080	5.903	5.903	0.370	0.370	9.595	9.595	19.118	19.118	9.349	9.349	47.770	47.770	5.220	5.220	16.055	16.055
16	320	raghavagps	50	75	11.492	7.661	51.989	34.659	6.519	4.346	0.449	0.300	8.928	5.952	18.710	12.473	9.272	6.181	42.096	28.064	5.257	3.505	8.520	5.680
17	389	AkbAR	48	75	10.981	7.028	51.193	32.764	6.201	3.969	0.389	0.249	8.237	5.272	23.340	14.938	6.758	4.325	32.855	21.027	3.796	2.429	16.154	10.338
18	373	Zhang_Contact	73	75	9.865	9.602	49.242	47.928	5.539	5.391	0.353	0.344	9.566	9.311	17.953	17.474	7.347	7.151	37.487	36.487	4.115	4.005	8.132	7.915

Figure S26. Screenshot of the performance or precision of different methods for Top10 list size and 'long+medium' range contacts on CASP12 TBM and FM targets.

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results | RR Assessn

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long+medium | List size: Top 10 | Prob filter: >0 | Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)		Prec-FDR(prob)	
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total
1	451	RaptorX-Contact	75	75	5.991	5.991	75.600	75.600	3.215	3.215	0.178	0.178	4.568	4.568	13.141	13.141	5.107	5.107	67.769	67.769	2.738	2.738	51.409	51.409
2	013	MetaPSICOV	73	75	5.369	5.226	70.822	68.933	2.875	2.798	0.173	0.168	5.524	5.377	12.265	11.938	4.630	4.507	63.720	62.021	2.480	2.414	46.643	45.399
3	219	Deepfold-Contact	73	75	5.492	5.346	70.685	68.800	2.942	2.864	0.180	0.175	6.045	5.884	11.515	11.208	5.455	5.310	70.370	68.494	2.923	2.845	41.704	40.592
4	109	naive	73	75	5.318	5.177	70.685	68.800	2.846	2.770	0.205	0.199	6.828	6.646	13.118	12.768	5.276	5.135	70.284	68.410	2.824	2.748	41.659	40.548
5	044	Yang-Server	73	75	5.331	5.189	69.863	68.000	2.845	2.769	0.190	0.185	5.631	5.481	11.273	10.972	4.276	4.162	58.823	57.255	2.282	2.221	37.648	36.644
6	432	Pcons-net	75	75	5.128	5.128	68.933	68.933	2.743	2.743	0.150	0.150	4.318	4.318	11.723	11.723	3.735	3.735	53.137	53.137	1.993	1.993	35.798	35.798
7	020	RBO-Epsilon	75	75	5.381	5.381	66.933	66.933	2.905	2.905	0.168	0.168	3.895	3.895	13.036	13.036	4.438	4.438	57.233	57.233	2.401	2.401	36.407	36.407
8	236	MULTICOM-CONSTRUCT	75	75	5.058	5.058	66.667	66.667	2.699	2.699	0.163	0.163	4.826	4.826	11.407	11.407	4.104	4.104	57.220	57.220	2.188	2.188	38.595	38.595
9	097	PconsC31	75	75	4.794	4.794	65.600	65.600	2.562	2.562	0.141	0.141	4.009	4.009	10.332	10.332	3.475	3.475	50.640	50.640	1.852	1.852	32.109	32.109
10	079	iFold_1	49	75	5.763	3.765	64.694	42.267	3.057	1.997	0.143	0.093	4.334	2.832	11.047	7.217	5.721	3.738	64.364	42.051	3.035	1.983	30.184	19.720
11	348	FALCON_COLORS	75	75	4.647	4.647	61.600	61.600	2.483	2.483	0.146	0.146	4.512	4.512	10.326	10.326	3.684	3.684	51.697	51.697	1.968	1.968	29.685	29.685
12	310	IGBteam	75	75	4.455	4.455	60.533	60.533	2.386	2.386	0.151	0.151	4.718	4.718	11.564	11.564	3.859	3.859	53.910	53.910	2.070	2.070	22.247	22.247
13	389	AkbAR	48	75	5.191	3.322	59.583	38.133	2.752	1.761	0.179	0.115	4.616	2.955	14.451	9.249	3.420	2.189	42.992	27.515	1.804	1.154	23.658	15.141
14	287	MULTICOM-CLUSTER	75	75	4.062	4.062	58.933	58.933	2.135	2.135	0.144	0.144	4.976	4.976	9.131	9.131	3.617	3.617	54.626	54.626	1.898	1.898	25.864	25.864
15	407	Distill	75	75	3.894	3.894	57.200	57.200	2.075	2.075	0.190	0.190	6.662	6.662	12.880	12.880	3.308	3.308	50.529	50.529	1.760	1.760	20.749	20.749
16	320	raghavagps	50	75	4.801	3.201	57.000	38.000	2.542	1.695	0.186	0.124	5.892	3.928	8.793	5.862	4.136	2.758	50.138	33.426	2.190	1.460	16.827	11.218
17	357	FLOUDAS_SERVER	67	75	3.914	3.497	55.821	49.867	2.107	1.882	0.149	0.133	4.599	4.109	7.118	6.359	2.686	2.399	44.708	39.939	1.426	1.274	36.858	32.926
18	431	Shen-Group	50	75	4.693	3.129	55.000	36.667	2.487	1.658	0.154	0.103	4.778	3.185	13.130	8.754	3.967	2.644	47.434	31.623	2.100	1.400	9.437	6.291

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure S27. Screenshot of the performance or precision of different methods for L/5 list size and 'long+medium' range contacts on CASP12 all category targets (TBM, FM and TBM/FM).

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results | RR Assessme

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long+medium | List size: L/5 | Prob filter: >0 | Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)		Prec-FDR(prob)		Pr
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Su
1	451	RaptorX-Contact	94	94	15.103	15.103	69.063	69.063	8.647	8.647	0.420	0.420	9.369	9.369	24.645	24.645	12.580	12.580	58.918	58.918	7.170	7.170	39.033	39.033	69.6
2	219	Deepfold-Contact	92	94	14.258	13.954	65.866	64.465	8.166	7.992	0.474	0.464	10.372	10.151	23.946	23.437	14.076	13.777	65.174	63.788	8.060	7.888	33.064	32.360	66.4
3	013	MetaPSICOV	92	94	14.182	13.880	65.685	64.287	8.122	7.949	0.425	0.416	9.987	9.775	24.053	23.541	11.854	11.602	56.044	54.852	6.774	6.630	37.156	36.366	66.5
4	109	naive	92	94	14.074	13.775	65.145	63.759	8.059	7.888	0.520	0.509	10.996	10.762	26.664	26.097	13.867	13.572	64.311	62.942	7.939	7.770	31.539	30.868	65.7
5	044	Yang-Server	92	94	13.655	13.364	63.525	62.173	7.790	7.625	0.436	0.427	8.872	8.684	22.153	21.682	9.839	9.630	47.100	46.097	5.586	5.467	25.833	25.284	64.5
6	432	Pcons-net	94	94	13.511	13.511	61.826	61.826	7.752	7.752	0.362	0.362	7.697	7.697	21.435	21.435	9.126	9.126	42.933	42.933	5.211	5.211	22.276	22.276	62.6
7	079	iFold_1	66	94	13.720	9.633	61.577	43.235	7.871	5.527	0.346	0.243	7.841	5.506	22.761	15.981	13.569	9.527	60.998	42.829	7.784	5.465	24.188	16.983	61.6
8	236	MULTICOM-CONSTRUCT	94	94	13.029	13.029	60.327	60.327	7.461	7.461	0.401	0.401	8.839	8.839	22.050	22.050	10.141	10.141	48.416	48.416	5.777	5.777	27.521	27.521	61.4
9	097	PconsC31	94	94	12.957	12.957	60.299	60.299	7.406	7.406	0.357	0.357	7.622	7.622	19.358	19.358	8.650	8.650	41.350	41.350	4.923	4.923	21.192	21.192	61.0
10	020	RBO-Epsilon	94	94	13.302	13.302	59.079	59.079	7.684	7.684	0.402	0.402	8.504	8.504	23.141	23.141	9.797	9.797	43.863	43.863	5.659	5.659	21.324	21.324	60.9
11	320	raghavagps	66	94	13.764	9.664	58.386	40.995	8.027	5.636	0.472	0.331	8.996	6.316	23.736	16.666	11.142	7.823	47.558	33.392	6.493	4.559	17.713	12.437	58.8
12	348	FALCON_COLORS	94	94	12.466	12.466	56.688	56.688	7.146	7.146	0.374	0.374	8.908	8.908	21.215	21.215	8.980	8.980	42.025	42.025	5.122	5.122	19.533	19.533	57.6
13	310	IGBteam	94	94	11.353	11.353	54.685	54.685	6.442	6.442	0.350	0.350	8.078	8.078	18.696	18.696	9.661	9.661	46.740	46.740	5.486	5.486	12.989	12.989	55.2
14	357	FLOUDAS_SERVER	83	94	11.550	10.198	54.671	48.273	6.619	5.844	0.425	0.375	8.799	7.769	18.437	16.279	8.137	7.184	41.134	36.320	4.583	4.047	30.958	27.335	55.1
15	287	MULTICOM-CLUSTER	94	94	11.054	11.054	52.436	52.436	6.297	6.297	0.357	0.357	8.941	8.941	19.001	19.001	9.548	9.548	46.329	46.329	5.414	5.414	13.777	13.777	52.9

Figure S28. Screenshot of the performance or precision of different methods for Top10 list size and 'long+medium' range contacts on CASP12 all category targets (TBM, FM and TBM/FM).

RR Analysis

Results Home | Table Browser | Estimate of Model Accuracy Results | RR Assessme

Per Group/Target Analysis | Averages | Similarity(Jaccard dist.) | Help

Per Group | Per Target

Contact Range: long+medium | List size: Top 10 | Prob filter: >0 | Text file

Domains classification

- TBM
- TBM/FM
- FM

Show

			No. Domains		F1		Prec		Recall		ES		ES(ext)		Recall(SS Tp)		F1(prob)		Prec(prob)		Recall(prob)		Prec-FDR(prob)		Pr
#	Gr.#	Gr. Name	No Submit.	No Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Submit.	Total	Su
1	451	RaptorX-Contact	94	94	6.957	6.957	75.000	75.000	3.780	3.780	0.176	0.176	4.330	4.330	13.521	13.521	5.802	5.802	66.424	66.424	3.143	3.143	49.207	49.207	75.3
2	013	MetaPSICOV	92	94	6.572	6.433	71.630	70.106	3.575	3.499	0.176	0.173	5.372	5.257	13.657	13.366	5.578	5.460	63.633	62.279	3.032	2.967	46.050	45.071	71.9
3	219	Deepfold-Contact	92	94	6.522	6.383	71.087	69.574	3.551	3.476	0.183	0.179	5.786	5.663	12.621	12.353	6.417	6.280	70.497	68.997	3.492	3.418	43.053	42.137	71.6
4	109	naive	92	94	6.372	6.236	70.761	69.255	3.469	3.395	0.204	0.200	6.457	6.320	14.272	13.969	6.257	6.123	70.073	68.583	3.404	3.331	42.352	41.451	71.3
5	432	Pcons-net	94	94	6.188	6.188	69.043	69.043	3.374	3.374	0.156	0.156	4.268	4.268	12.897	12.897	4.382	4.382	52.501	52.501	2.378	2.378	34.552	34.552	69.4
6	044	Yang-Server	92	94	6.023	5.895	68.478	67.021	3.259	3.190	0.182	0.178	5.105	4.996	12.360	12.097	4.608	4.510	55.828	54.641	2.485	2.432	34.114	33.388	68.8
7	020	RBO-Epsilon	94	94	6.581	6.581	67.872	67.872	3.609	3.609	0.170	0.170	3.769	3.769	14.128	14.128	5.291	5.291	57.202	57.202	2.899	2.899	36.173	36.173	68.6
8	236	MULTICOM-CONSTRUCT	94	94	5.950	5.950	67.021	67.021	3.230	3.230	0.162	0.162	4.619	4.619	12.301	12.301	4.792	4.792	57.154	57.154	2.595	2.595	37.798	37.798	67.5
9	079	iFold_1	66	94	6.670	4.683	66.970	47.021	3.605	2.531	0.155	0.109	4.528	3.179	12.620	8.861	6.617	4.646	66.582	46.749	3.577	2.511	34.566	24.270	66.9
10	097	PconsC31	94	94	5.669	5.669	66.596	66.596	3.065	3.065	0.150	0.150	4.087	4.087	11.696	11.696	3.986	3.986	50.275	50.275	2.148	2.148	31.935	31.935	67.0
11	348	FALCON_COLORS	94	94	5.703	5.703	62.447	62.447	3.099	3.099	0.153	0.153	4.424	4.424	11.833	11.833	4.366	4.366	51.013	51.013	2.365	2.365	28.957	28.957	62.8
12	320	raghavagps	66	94	6.532	4.586	62.424	43.830	3.577	2.511	0.206	0.144	6.045	4.245	12.844	9.018	5.547	3.895	54.661	38.379	3.033	2.130	25.005	17.556	62.7
13	310	IGBteam	94	94	5.254	5.254	59.894	59.894	2.859	2.859	0.151	0.151	4.528	4.528	12.233	12.233	4.562	4.562	53.254	53.254	2.485	2.485	20.850	20.850	60.1
14	389	AkbAR	62	94	5.671	3.740	58.387	38.511	3.065	2.021	0.176	0.116	4.710	3.107	14.545	9.593	3.917	2.584	42.877	28.281	2.117	1.396	21.279	14.035	58.6
15	287	MULTICOM-CLUSTER	94	94	4.865	4.865	57.766	57.766	2.623	2.623	0.148	0.148	4.925	4.925	10.184	10.184	4.233	4.233	52.910	52.910	2.275	2.275	22.637	22.637	58.0

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Table S1: List of targets with average precision value 50% or above when list size is L/5 and contact range is Long.

Target	Precision Score (%)
T0866-D1	100.00
T0912-D1	100.00
T0894-D1	100.00
T0944-D1	100.00
T0896-D2	97.50
T0947-D1	97.14
T0942-D1	97.14
T0948-D1	93.33
T0886-D2	92.31
T0909-D1	91.04
T0892-D2	90.91
T0946-D2	90.70
T0907-D1	88.89
T0918-D3	87.50
T0868-D1	87.50
T0899-D1	86.54
T0913-D1	83.82
T0895-D1	83.33
T0896-D1	83.33
T0887-D1	81.82
T0911-D1	81.71
T0901-D1	80.00
T0892-D1	78.57
T0864-D1	78.00
T0912-D2	76.47
T0876-D1	75.00
T0942-D2	74.42
T0890-D1	70.59
T0907-D2	68.75
T0882-D1	68.75
T0900-D1	66.67
T0915-D1	61.29
T0894-D2	54.55
T0905-D1	51.02
T0872-D1	50.00

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Table S2: List of targets with average precision value 50% or above when list size is Top10 and contact range is Long.

Target	Precision Score (%)
T0896-D1	100.00
T0866-D1	100.00
T0946-D2	100.00
T0944-D1	100.00
T0942-D1	100.00
T0918-D3	100.00
T0886-D2	100.00
T0912-D1	100.00
T0892-D2	100.00
T0894-D1	100.00
T0901-D1	90.00
T0912-D2	90.00
T0896-D2	90.00
T0907-D1	90.00
T0899-D1	90.00
T0868-D1	90.00
T0948-D1	90.00
T0909-D1	90.00
T0892-D1	90.00
T0947-D1	90.00
T0913-D1	90.00
T0911-D1	80.00
T0915-D1	80.00
T0864-D1	80.00
T0882-D1	80.00
T0907-D2	80.00
T0895-D1	80.00
T0904-D1	70.00
T0918-D1	70.00
T0894-D2	60.00
T0890-D1	60.00
T0888-D1	60.00
T0898-D1	60.00
T0887-D1	60.00
T0876-D1	60.00
T0900-D1	60.00
T0898-D2	50.00
T0870-D1	50.00
T0942-D2	50.00

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S3. List of targets with average precision value 50% or above when list size is L/5 and contact range is Long+Medium.

Target	Precision Score (%)
T0866-D1	100.00
T0912-D1	100.00
T0944-D1	100.00
T0896-D1	100.00
T0894-D1	100.00
T0896-D2	97.50
T0947-D1	97.14
T0942-D1	97.14
T0892-D2	95.45
T0882-D1	93.75
T0948-D1	93.33
T0946-D2	93.02
T0918-D3	91.67
T0909-D1	91.04
T0907-D1	88.89
T0913-D1	88.24
T0868-D1	87.50
T0892-D1	85.71
T0911-D1	85.37
T0942-D2	83.72
T0912-D2	82.35
T0898-D2	81.82
T0887-D1	81.82
T0894-D2	81.82
T0918-D1	81.82
T0886-D2	80.77
T0899-D1	78.85
T0864-D1	78.00
T0901-D1	75.56
T0907-D2	75.00
T0895-D1	70.83
T0898-D1	68.18
T0872-D1	66.67
T0876-D1	66.67
T0890-D1	64.71
T0915-D1	64.52
T0900-D1	61.90
T0904-D1	54.90
T0878-D1	52.17
T0905-D1	51.02

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table S4. List of targets with average precision value 50% or above when list size is Top10 and contact range is Long+Medium.

Target	Precision Score (%)
T0894-D1	100.00
T0946-D2	100.00
T0892-D2	100.00
T0942-D1	100.00
T0866-D1	100.00
T0944-D1	100.00
T0942-D2	100.00
T0912-D1	100.00
T0896-D2	100.00
T0896-D1	100.00
T0909-D1	90.00
T0912-D2	90.00
T0864-D1	90.00
T0868-D1	90.00
T0947-D1	90.00
T0948-D1	90.00
T0882-D1	90.00
T0886-D2	90.00
T0892-D1	90.00
T0918-D3	90.00
T0899-D1	90.00
T0918-D1	90.00
T0901-D1	90.00
T0913-D1	90.00
T0907-D1	90.00
T0894-D2	80.00
T0898-D2	80.00
T0911-D1	80.00
T0907-D2	80.00
T0907-D3	80.00
T0890-D1	80.00
T0878-D1	80.00
T0915-D1	80.00
T0895-D1	70.00
T0918-D2	70.00
T0898-D1	70.00
T0941-D1	70.00
T0904-D1	70.00
T0872-D1	60.00
T0888-D1	60.00

bioRxiv preprint doi: https://doi.org/10.1101/192120; this version posted September 21, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

T0876-D1	60.00
T0887-D1	60.00
T0870-D1	60.00
T0900-D1	50.00
T0874-D1	50.00