PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

CASP8 TARGET CLASSIFICATION

Target domain definition and classification in CASP8

Michael L. Tress,[1]* Iakes Ezkurdia,[1] and Jane S. Richardson[2]

[1] Structural and Computational Biology Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain
[2] Biochemistry Department, Duke University Medical Center, Durham, North Carolina 27710

ABSTRACT

In order to be successful CASP experiments require experimentally determined protein structures. These structures form the basis of the experiment. Structural genomics groups have provided the vast majority of these structures in recent editions of CASP. Before the structure prediction assessment can begin these target structures must be divided into structural domains for assessment purposes and each assessment unit must be assigned to one or more tertiary structure prediction categories. In CASP8 target domain boundaries were based on visual inspection of targets and their experimental data, and on superpositions of the target structures with related template structures. As in CASP7 target domains were broadly classified into two different categories: "template-based modeling" and "free modeling." Assessment categories were determined by structural similarity between the target domain and the nearest structural templates in the PDB and by whether or not related structural templates were used to build the models. The vast majority of the 164 assessment units in CASP8 were classified as template-based modeling. Just 10 target domains were defined as free modeling. In addition three targets were assessed in both the free modeling and template based categories and a subset of 50 template-based models was evaluated as part of the "high accuracy" subset. The targets submitted for CASP8 confirmed a trend that has been apparent since CASP5: targets submitted to the CASP experiments are becoming easier to predict.

Proteins 2009; 77(Suppl 9):10–17.

Key words: protein structure; domains; assessment units; structure prediction; structure classification.

INTRODUCTION

As part of the eighth round of the "Community Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction" (CASP8) protein structure prediction groups were asked to make blind predictions of 3D structure for 128 distinct target proteins over a 3-month period. The task of the CASP assessors is to evaluate these predictions, and the first step in the assessment process is to split the target structures into domains for the purposes of the assessment, to trim local segments where necessary, and to assign each target domain to a category based on the difficulty of the prediction and the methods used to predict the 3D structure.

The term domain is used here and throughout the assessment process to denote a self-contained segment of the 3D structure of the target that has been defined for the purposes of the assessment of the tertiary structure predictions. The domains chosen for assessment purposes are based on, but may not always coincide with, the generally accepted evolutionary or structural views of domains. The term "assessment unit" was introduced in CASP7[1] and is used interchangeably with the term "target domain" in the context of this article.

The classification of assessment units was unchanged from CASP7.[1] Most targets were assigned to one of the two assessed categories, "template-based modeling" or "free modeling,"

Additional Supporting Information may be found in the online version of this article.
The authors state no conflict of interest.
Grant sponsor: BioSapiens Network of Excellence; Grant number: LSHG-CT-2003-503265 (MLT); Grant sponsor: NIH; Grant number: GM073930 (JSR).
*Correspondence to: Dr. Michael Tress, Structural and Computational Biology Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain. E-mail: mtress@cnio.es
Received 8 April 2009; Revised 22 May 2009; Accepted 29 May 2009
Published online 9 June 2009 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/prot.22497

though four targets were classified in both the free modeling and template-based modeling categories. A subset of the template-based modeling target domains were also categorized as "high accuracy"; for these target domains template detection was considered trivial and assessors were able to evaluate more than just the backbone accuracy of the models. The classification process was based on structural similarity to the nearest PDB[2] template found by LGA.[3]

EXPERIMENTAL TARGET STRUCTURES

Predictors were asked to predict tertiary structure for 128 target sequences (T0387 to T0514) during the 3-month prediction season (Table 1). All but seven of the target structures were provided by structural genomics groups. The shortest target sequence was just 55 residues long (T0480), while the longest target sequence was the mostly disordered structure T0500 (829 residues). A total of 107 target structures were determined by X-ray crystallography, 21 by NMR.

At the time of their release, CASP8 target sequences were annotated by the organizers as either "human/server" targets or "server only" targets. All groups were assessed over the human/server targets, server only targets could be predicted by all groups, but only server predictions would be officially assessed. Targets expected to be more difficult were usually placed in the human/server category. There were 57 human/server targets (five were cancelled) and 71 server only targets (two were cancelled).

Seven targets were cancelled either during or by the end of the prediction season, so predictors were evaluated over 121 targets. Two targets were excluded from the assessment because the structure was not released in time for the CASP meeting (T0403 and T0439) and a third target (T0387) was cancelled because of the premature availability of its structure. Two targets (T0484 and T0500) were not assessed for structure prediction because they were predominantly disordered. Target T0467 was not assessed because the initial structure that the Prediction Center received was a very loose NMR ensemble.

Target T0410 (PDB: 3D3L) has a violently altered C-terminal conformation resulting from an inadvertent read-through of a stop codon, and was cancelled by the assessors. The C-terminal residue isoleucine is replaced by 33 extra residues starting with a serine. In the nearest template, (2P0M, 60% identity) the C-terminal residue, a leucine, binds the active site iron with its terminal carboxyl group and is thoroughly buried. In T0410, the C-terminal iron ligand is no longer available and is replaced by a water molecule; the extra C-terminal 32 residues in T0410 are disordered, as are several other

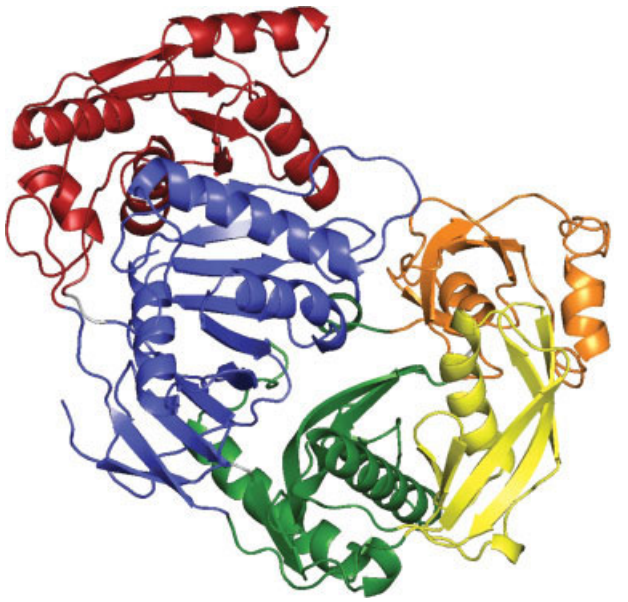

Figure 1

Domain definition for target T0487. Ribbon representation of target T0487 (PDB 3DLB),[5] the argonaute protein silencing complex, which was defined as having five domains. Here, the five assessment units are shown labeled blue to red along the chain.

nearby loops. In addition, the long N-terminal helix and its loops take up an entirely new orientation relative to all related enzymes; nearly 90° rotated and with some residues shifted more than 20 Å. There is evidence from the template that the N-terminal may undergo conformational changes,[4] but the assessors felt that the extra C-terminal residues may have played a large role in this case, making this target too unpredictable to evaluate.

The 121 targets were divided into 164 assessment units: the smallest had just 30 residues (T0480 again), the largest 561 (target T0450). One target (T0487) was divided into five target domains, the most domains of any target in any CASP (Fig. 1). Many of the assessment units were structurally related, 70 of 165 target domains having a common fold with at least one other target domain. Ten of the target domains had Rossman folds and there were eight Tudor domains. Supporting information Table 1 contains an overview of the CASP8 targets and further information is available on the CASP8 web pages (http://predictioncenter.org/casp8/index.cgi).

STRUCTURAL TEMPLATES

As in previous CASPs, a key part of partitioning the targets into domains was the search for structurally similar templates. We carried out template searches against the whole PDB database, both with whole target structures and with a range of potential assessment units. In a first pass, we used the structural alignment program

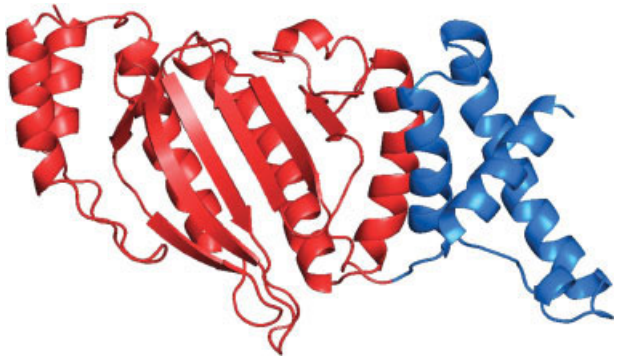

Figure 2

The domain definition of target T0405. Target T0405 (not deposited) was a free modeling target that was split into two domains for the structure prediction evaluation. The all-helical domain 1, shown in blue, was the better predicted domain.

Mammoth[6] and a 97% nonredundant set of chains from the PDB. In the second pass, we used LGA to superimpose the target onto the top scoring 1500 chains by Mammoth z-score and all similar chains previously removed by clustering at 97%. For those target domains where no clear template was found despite running both Mammoth and LGA, we ran LGA again against the whole PDB database.

The top 20 structural templates for each assessment unit can be found on the CASP8 web pages (http://predictioncenter.org/download_area/CASP8/list_of_templates.txt).

TARGET DOMAIN DEFINITION

Assessment units were generally assigned based both on structural integrity and on evolutionary relationship. Visual inspection of the targets and the target-template superpositions allowed us to make domain boundary definitions for the majority of the targets. Final domain definitions were made by consensus between the assessors.

When making domain boundary decisions, assessors took other factors into account beyond structural integrity and evaluation. Targets with two structural domains that had identical orientations relative to their closest structural templates were not split into separate domains as it was felt that predictors should be able to predict the domains with the correct orientation. One side effect of this decision was that the domain definitions for targets T0430, T0456, T0483, and T0494 were incongruous. Although all four targets have the same SCOP[7] family (the protein kinase catalytic subunit), T0430 and T0456 were split into two assessment units while T0494 and T0483 were not.

In a number of cases, assessment units were defined based on what the assessors felt would most help predictors. For example, the target T0405 did not have a template that covered the whole structure. It was split into two assessment units (Fig. 2) because the three N-terminal helices formed a predictable domain that did have useful templates. T0443 was another target for which there were not templates for the entire structure and it does not have clear domain boundaries in the structural sense. It was split into three assessment units because templates were available for domains 1 and 3. Indeed predictors made relatively successful template-based predictions for these target domains. The remainder of the target was defined as the second assessment unit (Fig. 3) but should not be considered a domain in the structural sense. It was particularly difficult to predict as it is made up of two nonconsecutive sections.

The domain definition process is best illustrated with examples. Target T0472 is unusual in that it was an NMR target that was split into different assessment units. Although the target is clearly a single domain in the structural sense, there were no template structures that included both halves, which meant that there was no indication of the correct orientation of the subdomains. The target was evaluated as two separate assessment units, and there were templates for both subdomains. In fact, the templates for the second target domain were so similar that it was classified as part of the high accuracy subset. For this target, the prediction of the correct relationship of the two subdomains was assessed as a special case in the TBM section.[8]

Target T0446 [Fig. 4(a)] was also split into two assessment units. Target T0446 really only makes sense as a pentamer, but the nearest template, 3B77 [Fig. 4(b)], has 12-fold symmetry in a conical dish shape, with the related structural regions out at the ends of long helical hairpins. The contact between the chains in T0446 and 3B77 is very different, and the relative orientation of the two subdomains in each chain is also somewhat different. Therefore, T0446 was split into two assessment units.

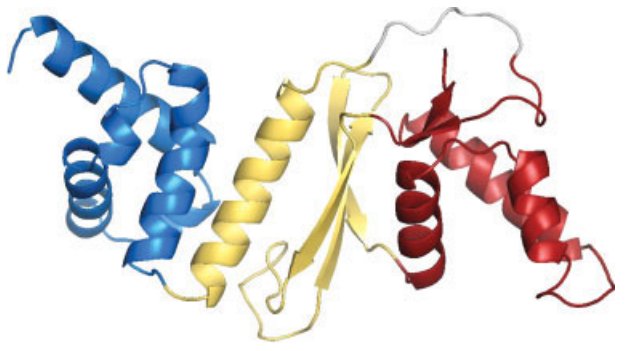

Figure 3

The domain definition of target T0443. Target T0443 (PDB 3DEE) was split into three assessment units despite the fact that the central target domain (in yellow) is not a domain in the structural sense. Domains 1 and 3 have templates and were in fact better predicted than domain 2.

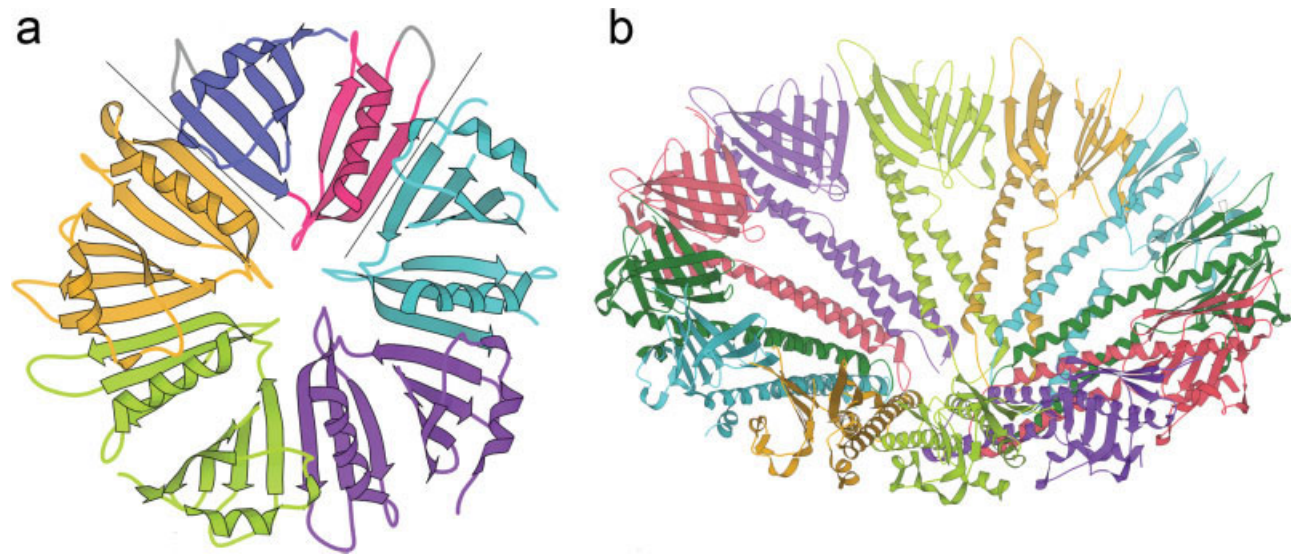

Figure 4

The domain definition of target T0446. In (a) target T0446 (PDB 3CDX) is shown as a pentamer, with four of the chains colored gold, green, purple, and cyan, while the chain at the top is colored as per the domain definition—domain 1 in blue, domain 2 in red, and trimmed loops in grey. Those grey loops are geometrically required to have different crystal contacts for each chain in the pentamer, since fivefold symmetry cannot repeat in space; those differences can be appreciated around the edges. It is apparent from the diagram that the stronger contacts are between the chains, through a continuous b-sheet, rather than within the chains. Section (b) shows the nearest template, 3B77, in which the regions homologous to T0466 form a 12-mer, joined in a dish shape by a long helical hairpin.

Although the official domain definitions were those used by the assessors as part of the main evaluation track, the predictions for all targets that were split into domains were also assessed automatically by the Prediction Center as full targets and these results are available on the CASP8 web pages.

DOMAIN EXTENSIONS, CRYSTAL CONTACTS, AND LINKERS

As in CASP8, we excluded residues from the official assessment units where we felt that the local conformation of the target structure was either uncertain or was likely to be influenced by factors other than intrachain interactions. Regions that we judged were strongly influenced by crystal packing were left out from the official definitions, for instance the C-terminal tail in T0434 past proline 195 (proline 176 in PDB structure 3DAL), which sticks far out to make a crystal contact. Cloning artifacts that appeared in the final structure were left out, for example, target T0392 where the N-terminal extension that forms a dimer is a cloning artifact. Also omitted were regions that show significant differences (>3.5 Å) between the chains in the deposited structure, such as the outer loops in target T0446 that can be seen to differ even in the ribbon representation [Fig. 4(a)].

The assessors identified and left out of the assessment local regions in target structures whose positions were highly uncertain (and therefore not appropriate as a reference for judging predictions). For structures solved by NMR we made the decision to select what we felt was the most representative model from the ensemble and then trimmed those residues where there was a large dif-

Figure 5

Trimming an NMR target. Target T0466 (PDB 2K5D) is an NMR ensemble of 20 models, shown with ribbons colored blue to red from N- to C-terminus. The N-terminus and the orange loop are clearly disordered, and short sections at the C-terminus and in the two green loops differed by more than the 3.5 Å cutoff. The domain definition thus contained four residue ranges.

ference between the models. In previous CASPs this was done by eye; here we left out residues where there were differences of greater than 3.5 Å between the Cαs of any two models, although single deviating loop residues were not trimmed. For example, for the T0466 NMR ensemble in Figure 5, the N-terminus and orange loop are completely disordered (where there was little or no experimental data), and the C-terminus and the two green loops differed by more than the 3.5 Å cutoff after superposition. For crystal structures uncertain regions were identified and removed from the assessment if they had very weak or absent electron density (for example loop 214–218 in T0442 and loop 92–100 in T0493); if maps were not available then a judgment was made by the assessors based on regions with very high crystallographic B-factors relative to the rest of the protein.

Many of the target structures had extensions beyond the compact globular core, which in many cases could not be built directly from any structural template. In deciding whether or not to include an extension as part of an official assessment unit, the assessors considered many factors, and each case was treated on its own merits. If the extension could be predicted directly from templates, it was always included in the official target domain. For example, target T0408 has a four-helix-bundle core with a two-helix N-terminal extension that forms a noncompact monomer. Six chains interlock into a compact hexamer in the biological molecule [Fig. 6(a)]. This arrangement also occurs in the homologous templates, so we retained the whole chain of T0408 as a single assessment unit.

If the extension formed a separate core and might be considered as a domain on its own, it was defined as a separate assessment unit. The C-terminus of T0510 cannot be built from a homologous template – the C-termini of all similar structures snake across the surface of the templates. However, the T0510 C-terminus forms a small domain with a hydrophobic core [Fig. 6(b)], so it was retained as a separate assessment unit for free modeling.

In target T0409 there is a 30-residue region at the N-terminus that forms a domain-swap interaction with the other chain in the crystal structure, tight enough to be the presumed biological unit. There is no template for this extension in the PDB and the predictors were not told expressly that this target would be a dimer. There was nothing at all to suggest that predictors might have been able to predict the extension so it was removed from the official target definition [Fig. 6(c)].

Target T0395 has a 60-residue C-terminal section of mixed helix and non-repetitive structure that wraps itself around the surface of the protein. There is no PDB template for this section, though the undeposited target T0320 from CASP7 is highly similar. Without a template the C-terminal region could only have been evaluated in the free modeling category, but it does not form a hydro-

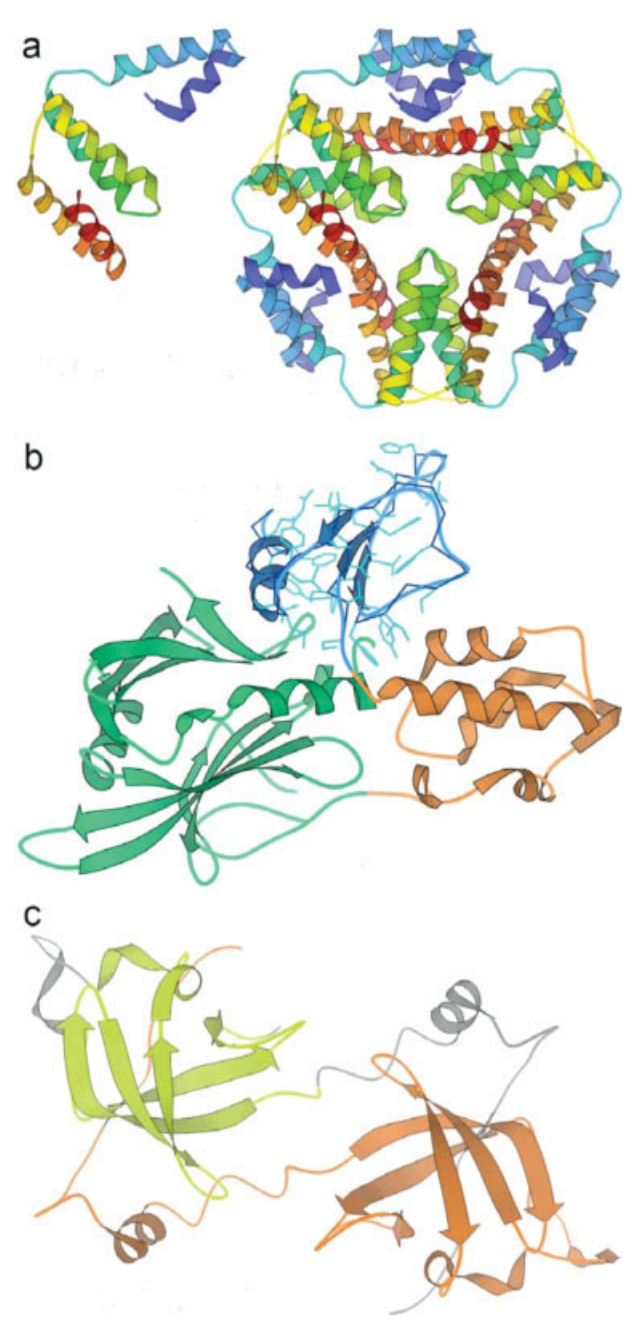

Figure 6

Extensions from the core. Three examples of domain definitions for targets that had extensions beyond the core of the principal structural domain. Target T0408 (PDB 3D7I) in (a) has a two-helix extension that is only compact in the hexamer. There are similar templates, however, so the extension was left in the domain definition. Part (b) shows the domain definition for target T0510 (PDB 3DOA), which has an extension without a homologous template structure (shown in blue); in this case it folds with a hydrophobic core and was assessed as a separate domain. In (c), the domain extension in target T0409 (PDB 3D0F) forms a tight domain-swap dimer, chain A shown in green and gray, chain B in orange. In this case, there is no similar template and the extension was left out of the core domain definition (in lime green).

phobic core and could only be modeled in the context of its interactions with the rest of the chain. As a result it was trimmed from the official assessment unit. Both the T0395 C-terminal region and a compact, unswapped monomer version of target T0409 were considered separately as special cases outside the main evaluation track and are reported in the template-based modeling assessment paper.[8]

Although we trimmed residues from several targets, we did not trim all regions without a template. Knowing when to use a template and when to deviate from it is an important part of the prediction of harder targets. For example, target T0513 has N- and C-terminal helices that could not be modeled from homologous templates. Although we trimmed two residues from the N-terminus because of differences between the chains, the two helices were left in the assessment because they pack to the surface of the structure and are not dominated by crystal contacts.

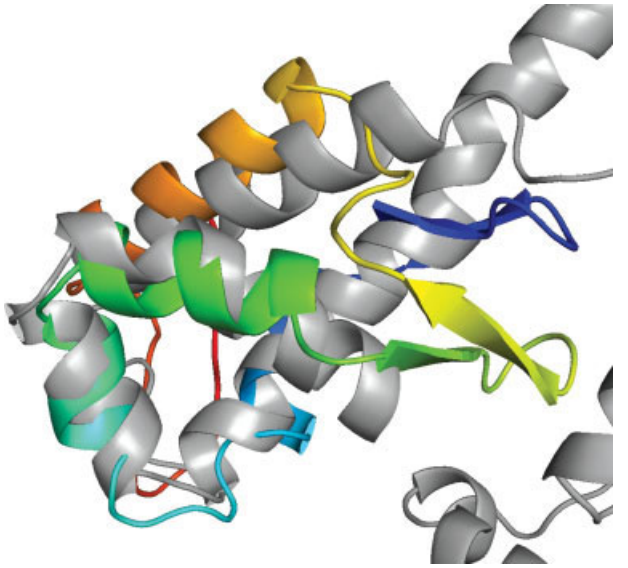

Figure 7

Target T0476 and its two structural templates. Target T0476 (PDB 2K5C) superimposed on the nearest structural template (PDB 2ZBZ). T0476 is colored blue to red along the sequence, and the template is shown in grey. The template covers more than half the target structure, but the yellow and blue strands on the right of the figure form a zinc-binding motif that could have been predicted from rubredoxin-like zinc ribbon proteins unrelated to the major template structure. Much of the structure of 2ZBZ has been hidden to make the figure clearer.

TARGET DOMAIN CATEGORIZATION

For evaluation purposes, the target domains were classified as template-based modeling or free modeling. A few target domains were classified as both free modeling and template-based modeling. A subset of the template-based modeling target domains were also classified as "high accuracy" to allow the modeling of detailed structural features such as side chains to be assessed.

When making decisions on the category of the target domains, the assessors generally made use of just two pieces of information: the complete set of structurally similar template structures detected from Mammoth and LGA searches of the PDB, and the list of parent structures that the predictors used to build their models. Not all predictors provide this information in their submitted models, but the information is vital, since the assessors can instantly know when a predictor has used a relevant template, or much of a relevant template, to model a target structure. Unfortunately, the converse is not true - the absence of explicit information on the use of templates does not indicate that a prediction was "template-free", it may also be that the predictors do not know which templates have been used in a fragment-based assembly, or they have not chosen to provide this input for other reasons.

Assessment units were categorized as template-based modeling if LGA detected one or more structurally similar templates that covered the core of the assessment unit and if predictors had explicitly used one or more of the structurally similar templates during the modeling process.

Target domains were classified as free modeling targets either because we were unable to find a similar template structure to cover the whole assessment unit (just targets T0397-D1 and T0496-D2), or because the predictors did not use the structurally similar template structures that we found with LGA (for example, T0416-D2 and T0443-D2).

T0513-D2 was a difficult target domain to classify because there were many good predictions for the target, yet no predictor had explicitly used any of the available good templates. The 26 models with the highest GDT-TS scores were nearly identical to one another and were more similar to the real structure of the domain than the best template structure. We reasoned that the good predictions for this assessment unit could be traced back to a single good server model and that the predictors had recognized the value of this model. T0513-D2 was thus defined as a free modeling target.

Three assessment units were considered to have both template-based and free modeling characteristics and were assessed in both categories. T0476 has remotely similar templates for two sections that were difficult to combine. Although some predictors had used a template for target T0443-D1, their models were not among the best. Target T0460 was reclassified from the template-based modeling category into the template-based/free modeling overlap category after the Cagliari meeting. While one of the best predictors had predicted a good

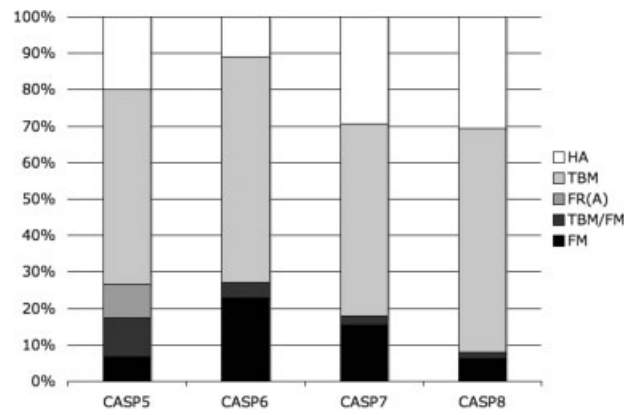

Figure 8

The changing distribution of CASP targets. CASP target domains from CASPs 5 to 7 were classified in the same way as CASP8, and the proportion of target domains in the four classification categories (free modeling, FM/TBM overlap, template-based modeling and high accuracy) are plotted against successive CASP experiments. Targets were reclassified for this exercise; for example even though CASP7 has the same categories, three targets that were classified as template-based in CASP7 (T0283, T0299-D1 and T0299-D2) would have been classified as free modeling in CASP8. CASP5 and 6 had a different categorization system from CASP7 and 8. All the CASP6 targets have been reclassified as per CASP8, but the 7 targets categorized as "fold recognition (analogy)," abbreviated FR(A), in CASP5 have not been reclassified. It is highly likely that most or all of these targets would have been categorized as free modeling targets in CASP8.

model with a template, it was clear that the remainder of the good models (and in particular the outstanding models for the target) were predicted by methods best suited to the free modeling section.

Template-based modeling targets were also classified as high accuracy if highly similar template structures were available with LGA-S scores greater than 80, and if there were a substantial number of first models with a GDT-TS of over 80.

In total 11 target domains were defined as free modeling targets, 151 were defined as template-based targets, and three as free modeling/template-based modeling overlap targets. Fifty template-based target domains were also defined as high-accuracy targets.

DISCUSSION

The CASP categorization system was last changed after CASP6 and the "new fold" category was replaced by the free modeling category, in part because there were few real new folds in CASP6. It is clear that the change was needed, since in CASP8 only two of the 10 free modeling targets could be described as new folds (T0397-D1 and T0496-D2). In addition, there were several target domains where it would only have been possible to build a tempate-based model for the core of the structure. These targets included two free modeling targets (T0405-D2, T0482) and one free modeling/template-based overlap target (T0476), for which it was possible to bolt together a model using two unrelated templates (Fig. 7). T0443 would have been a new fold if it had not been split into separate assessment units.

The lack of new folds illustrates a trend that is clear from the last four CASP experiments (Fig. 8)—the nature of the CASP targets is changing. While the number of targets has increased with every new CASP—total assessment units have more than doubled between CASP5 and CASP8—there has been no similar increase in the number of target domains evaluated in the free modeling/ab initio sections (FM targets). Comparisons between CASP experiments, using the CASP8 criteria to categorize the targets where possible, show that the opposite has happened. There are now half as many FM targets than there would have been in CASP6,[9] despite the best efforts of the Prediction Center to find free modeling targets. In terms of proportion, the FM targets in CASP8 are a third of what they would have been in CASP6.[10] At the same time, the proportion of targets in the high accuracy section has increased. Fifteen CASP5 target domains fit the criteria used to define the CASP8 high accuracy targets. This had risen to 50 assessment units by CASP8, in part due to a conscious decision to include more high accuracy targets after CASP6.

There are many possible reasons for these changes. It is certainly true that the PDB now covers more of the structural space. Between 2002 (the time of CASP5) and 2008, the number of deposited PDB structures rose from 19,000 to more than 55,000. It may also be that the number of new folds being deposited in the PDB has slowed dramatically.[11] Changes in the priorities of the structural genomics initiative consortia may also have contributed to the lack of new folds in CASP. Finally, methods for the detection of templates are also clearly more sensitive than they were in 2002.[12]

Whatever the reason, the trend seems likely to continue in CASP9. This will have several consequences. For example, the subset of free modeling targets is likely to become too small to allow any conclusions at all to be made in the free modeling category, and assessing free modeling predictions may require some sort of continuous assessment system.[13,14] In addition if the proportion of high-accuracy targets grows, more emphasis will have to be placed on evaluating more than just the accuracy of predicting Cα positions. The changing equilibrium between the "hardest" and "easiest" targets also suggests that the system of target categorization may need to be revisited.

ACKNOWLEDGMENTS

The authors thank the Prediction Center at UC Davis and especially Andriy Kryshtafovych for processing the targets and for providing comprehensive and timely sup-

port and materials for the categorization process. The authors would also thank the experimental groups that provided the target structures for the CASP8 experiments and we are grateful to the previous assessors for sharing their expertise and experience prior to the assessment. Finally, the authors would like to thank the developers of the structural analysis programs that have greatly aided the analysis and categorization of the targets in the last three CASPs, Adam Zemla and Angel Ortiz.

REFERENCES

1. Clarke ND, Ezkurdia I, Kopp J, Read RJ, Schwede T, Tress ML. Domain definition and target classification for CASP7. Proteins 2007;69(Suppl 8):10–18.
2. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. Nucleic Acids Res 2000;28:235–242.
3. Zemla A. LGA: a method for finding 3D similarities in protein structures. Nucleic Acids Res 2003;31:3370–3374.
4. Choi J, Chon JK, Kim S, Shin W. Conformational flexibility in mammalian 15S-lipoxygenase: reinterpretation of the crystallographic data. Proteins 2008;70:1023–1032.
5. Wang Y, Sheng G, Juranek S, Tuschl T, Patel DJ. Structure of the guide-strand-containing argonaute silencing complex. Nature 2008;456:209–213.
6. Ortiz AR, Strauss CE, Olmea O. MAMMOTH (matching molecular models obtained from theory): an automated method for model comparison. Protein Sci 2002;11:2606–2621.
7. Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 1995;247:536–540.
8. Keedy DA, Williams CJ, Arendall WB III, Chen VB, Kapral GJ, Gillespie RA, Zemla A, Richardson DC, Richardson JS. The other 90% of the protein: assessment beyond the Cas for CASP8 template-based models. Proteins 2009; this issue.
9. Kinch LN, Qi Y, Hubbard TJ, Grishin NV. CASP5 target classification. Proteins 2003;53(Suppl 6):340–351.
10. Tress M, Tai CH, Wang G, Ezkurdia I, Lopez G, Valencia A, Lee B, Dunbrack RL, Jr. Domain definition and target classification for CASP6. Proteins 2005;61(Suppl 7):8–18.
11. Levitt M. Growth of novel protein structural data, Proc Natl Acad Sci USA 2007;104:3183–3188.
12. Battey JN, Kopp J, Bordoli L, Read RJ, Clarke ND, Schwede T. Automated server predictions in CASP7. Proteins 2007;69(Suppl 8):68–82.
13. Bujnicki JM, Elofsson A, Fischer D, Rychlewski L. LiveBench-1: continuous benchmarking of protein structure prediction servers. Protein Sci 2001;10:352–361.
14. Koh IY, Eyrich VA, Marti-Renom MA, Przybylski D, Madhusudhan MS, Eswar N, Grana O, Pazos F, Valencia A, Sali A, Rost B. EVA: evaluation of protein structure prediction servers. Nucleic Acids Res 2003;31:3311–3315.