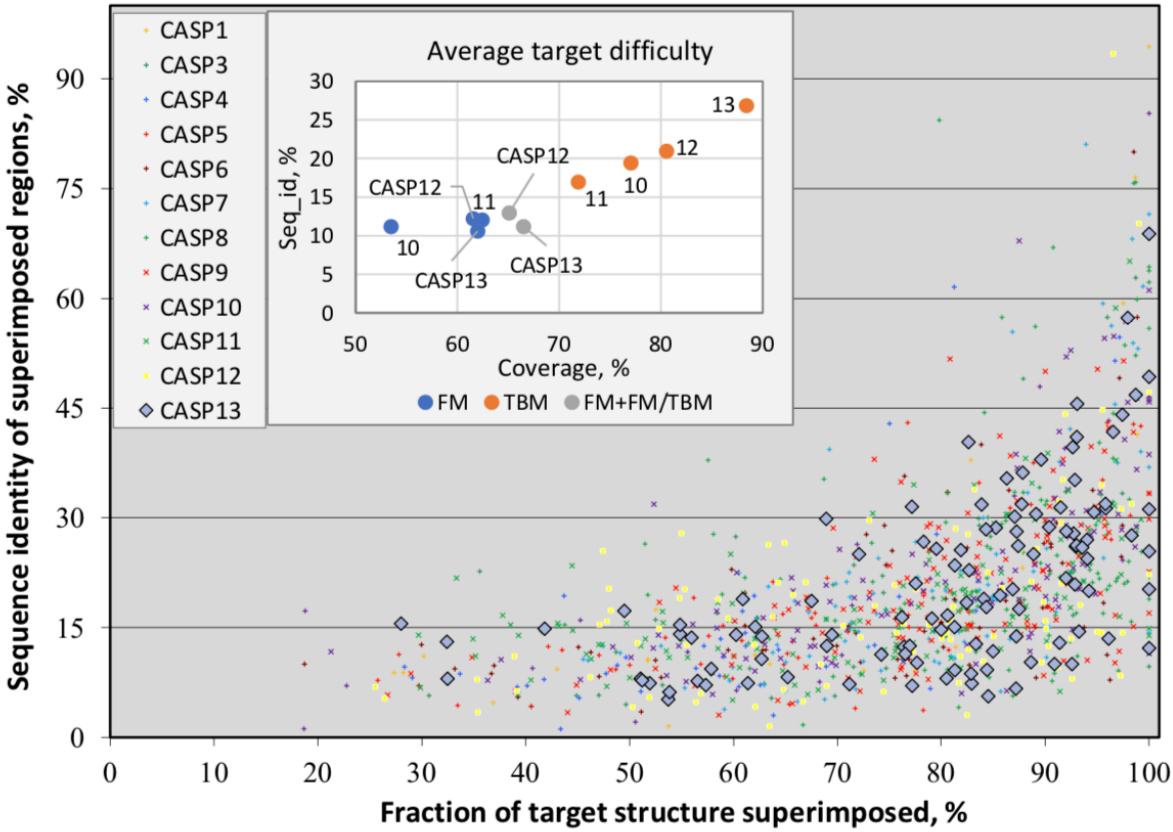
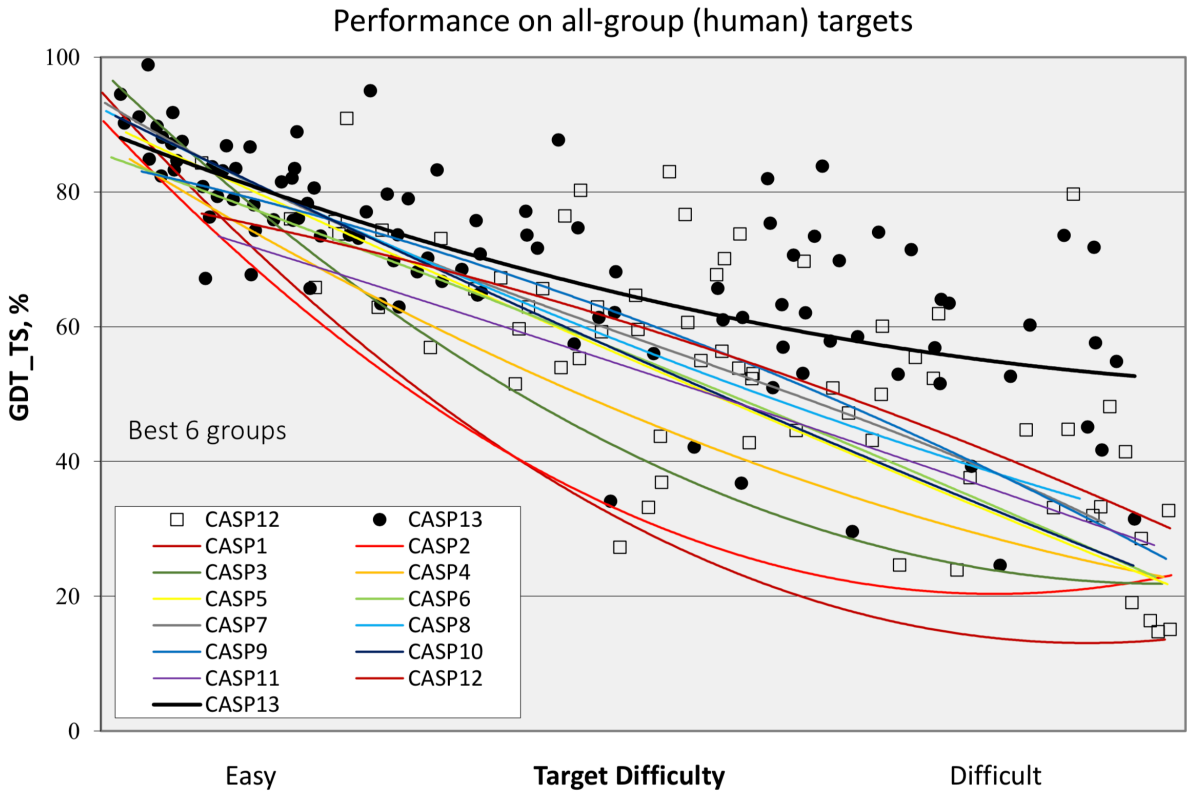
SUPPLEMENTARY FIGURES Critical Assessment of Methods of Protein Structure Prediction (CASP) – Round XIII

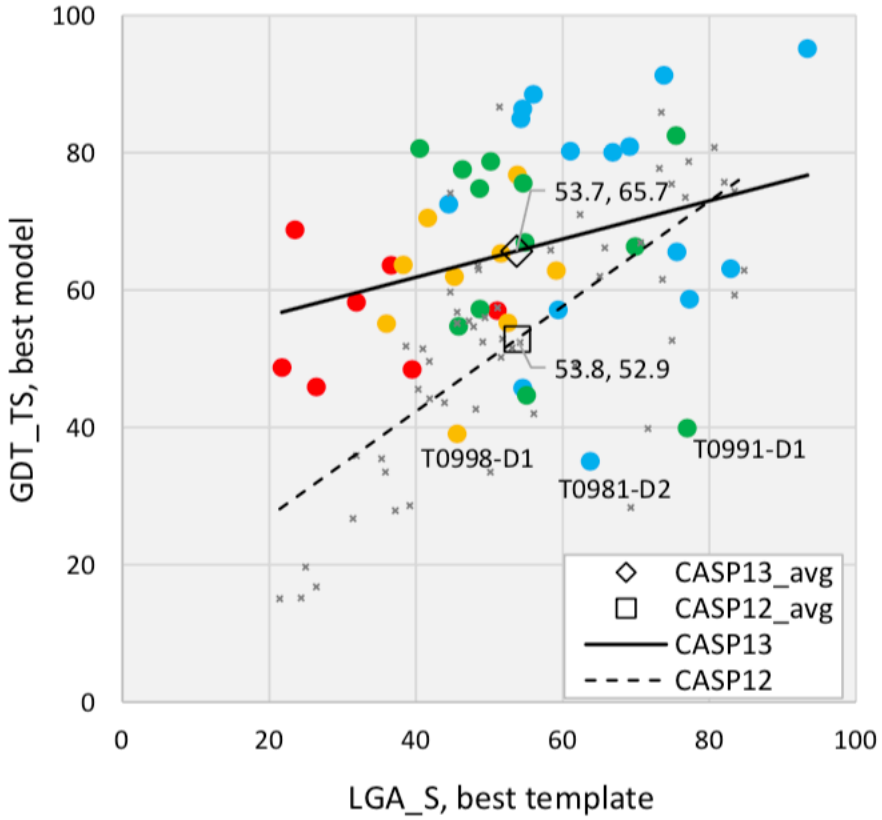


Supplementary Figure S1: Difficulty of each target in each CASP experiment, in terms of the sequence identity between the target and the closest available structural template and the fraction of the target that can be superposed on that template (see [45] for details). Overall target difficulty used in other figures is expressed as a combination of these two quantities [45]. The insert shows average target difficulty in recent CASPs, for different categories of target, using the same two variables. Note that by this measure, free modeling targets have been of very similar difficulty in last three CASPs, so

progress is not likely due to target differences. Supplementary figure S3 shows an additional comparison of target difficulty across the two most recent CASPs.

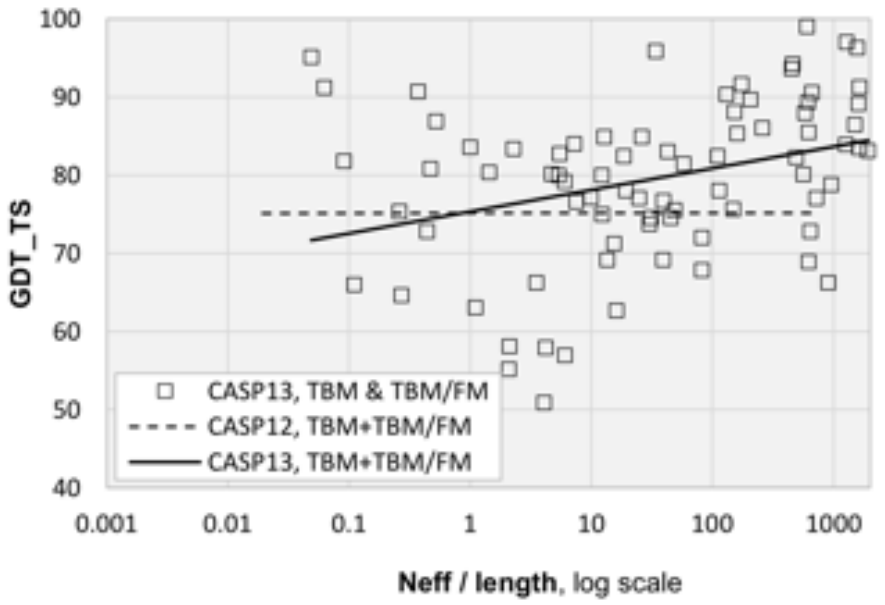


Supplementary Figure S2: Trend lines of average backbone accuracy (best six groups on each target) in each of the 13 CASP experiments. Individual target points are shown for the two most recent experiments. The accuracy metric, GDT_TS, is a multiscale indicator of the closeness of the Cα atoms in a model to those in the corresponding experimental structure. Target difficulty is based on sequence and structure similarity to other proteins with known experimental structures (see [5] for details). As with the single best model plot (figure 1), there is a striking improvement in accuracy in CASP13 (top black line), particularly for the more difficult targets. The similarity of the two plots shows that substantial improvement in performance was achieved by a number of groups.



Supplementary Figure S3: Comparison of the best model accuracy obtained (GDT_TS) with highest similarity of the target structure to any already known one (LGA_S [7]) for difficult modeling targets (FM and FM/TBM). CASP13 targets are colored according to their length: blue (<100 residues), green (100-149), orange (150-199), red (200+). CASP 12 targets are small grey crosses. There is a tendency for more accurate models to be produced when there is a good (but undetectable by sequence) template available, suggesting contact driven template searches were helpful for some targets. The trend is much less pronounced for CASP13 than for CASP12, where deep learning

methods had become so powerful that template search was comparatively less effective. Both the averages and distributions of LGA_S values are similar for CASPs 12 and 13, showing that the overall quality of possible structural templates was similar.



Supplementary Figure S4: Best model backbone accuracy for comparative modeling targets as a function of normalized alignment depth for the two most recent CASPs. CASP13 accuracy shows a shallow dependence on alignment depth, suggesting contact prediction played some role for this category of targets.