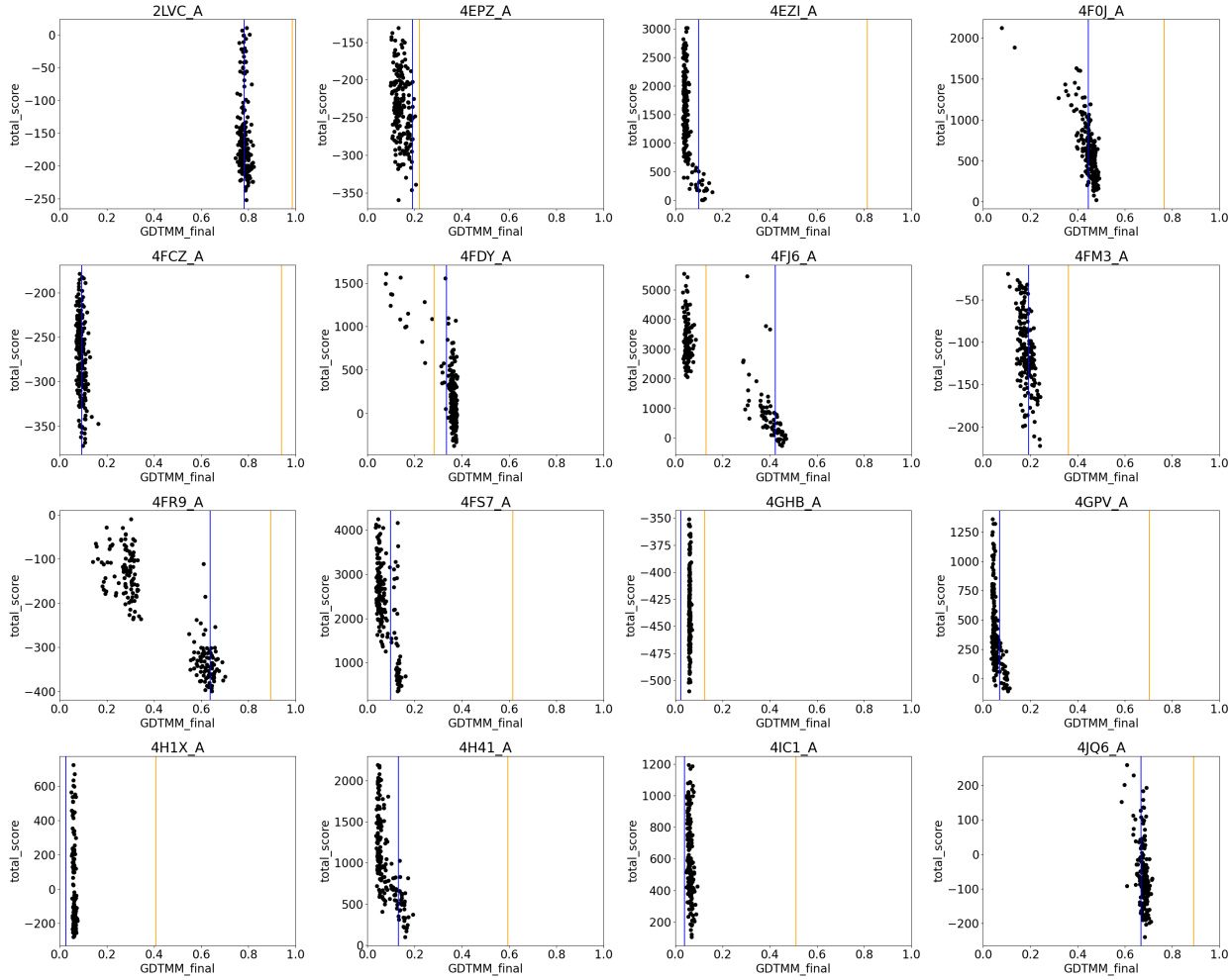
Scientific test: RosettaCM

FAILURES

None

RESULTS



AUTHOR AND DATE

Adapted for current benchmark framework by Jason Fell (jsfell@ucdavis.edu; Siegel Lab @ UC Davis), April 2020

PURPOSE OF THE TEST

This benchmark is meant to test the current performance of generating homology models with the RosettaCM protocol.

(Song, Y.; DiMaio, F.; Wang, R.Y.R.; Kim, D.; Miles, C.; Brunette, T.; et al. High-resolution comparative modeling with RosettaCM. Structure. 2013. 21:1735-1742.)

BENCHMARK DATASET

The dataset currently consists of 16 of 68 targets from the CASP10 (http://predictioncenter.org/casp10/), and were used in the original

RosettaCM paper. These 16 targets had pdb codes listed in RosettaCM paper, as well as have a variety of different sizes, folds, and fold complexities. The entire set contains several multi-domain proteins, we did not take them here. Each target uses a set of template structures, which were listed in the original RosettaCM article (Structure 2013), that are used to generate the homology models.

The targets used (by pdb code): 4EPZ 4H41 4JQ6 4FDY 4GHB 4FMZ 4GOQ 4FM3 4FJ6 4FGM 4EZI 4FR9 4AK1 4HES 2LVC 4FCZ

PROTOCOL

The Rosetta wiki describes the RosettaCM protocol (https://www.rosettacommons.org/docs/latest/application_documentation/structure_prediction/RosettaCM)

To run the RosettaCM protocol the required input files are:

*target fasta sequence

*threaded models

*hybridize.xml

The hybridize.xml runs a hybridize mover that samples structural combinations of multiple input models. The mover runs in three stages:

-Stage 1 is a stocastic procedure to generate a global superposition of aligned poritons of target and templates, which uses a centroid scoring method (score3).

-Stage 2 improves model geometry and explore conformational changes through a Monte Carlo sampling with two-step moves, and is scored by a centroid energy function

with smooth reparameterizations (score4_smooth_cart)

-Stage 3 side chains are built and optimized by Rosetta Monte Carlo sampling and is scored with a Rosetta full-atom energy function (ref2015_cart)

After the hybridize mover is called, Rosetta then performs a relax mover on the model (FastRelax; https://www.rosettacommons.org/docs/latest/application_documentation/structure_prediction/relax),

which is scored with the Stage 3 full-atom energy function (talaris2013_cart).

Once all of these files are ready run the command:

Rosetta/main/source/bin/rosetta_scripts.linuxclangrelease \

-database Rosetta/main/database \

-in:file:fasta target.fasta \

-parser:protocol hybridize.xml \

-default_max_cycles 200 \

-dualspace

-in:file:native (x-stal structure pdb if running test)

As a current estimate, generating 200 models for each target requires ~ 24 hours. 16 targets x 24 = ~ 384 CPU hours.

PERFORMANCE METRICS

The metric to use is the global distance test (GDT; Zemla, A., Venclovas, C., Moult, J., and Fidelis, K. (1999). Processing and analysis of CASP3 protein structure predictions.

Proteins (Suppl 3), 22-29.) to compare models to their respective crystal stucture. This test is meant to measure how well Rosetta can generate acurate homology models and tests how well the top target GDT changes compared to the original measures.

GDT values in orange are the top GDT obtained from the original RosettaCM paper (Structure 2013) - see Table in the Supplement. These values were generated from the average of the lowest 10% of energy models from several thousands, using an older scorefunction (likely talaris13). Here, we define the cutoffs as the top GDT values from the initial run, minus 0.05. If no model is generated above this cutoff, the test fails. We don't use the stdev here as it is meaningless with GDT-MM as a quality measure because there are many models created with all kinds of GDTs. Only the top GDT is meaningful to define model quality.

KEY RESULTS

This test is meant to measure changes in the overall performance of RosettaCM. As noted earlier, GDT was used as a measure of model accuracy. The original paper also

utilized alignment constraints generated by Robetta and used Fragment picker to generate fragments (a very robost method), where as this test uses the basic RosettaCM

protocol to generate models. Therefore the GDT's generated from this test may not perform as well as the original paper. This test can measure can be used as a means to

compare how well the curent protocol performs in relation to the more robust method.

DEFINITIONS AND COMMENTS

Potentially running these tests are time consuming, so more hours/time may be required.

LIMITATIONS

Independent of this protocl there are additional methods that can improve RosettaCM (i.e. to include additional evolutionary and catalytic constraints) which are not

currently standard for this protocol. Therefor, RosettaCM could include these other constraints when applicable.

Lastly, multiple tests could be run using other scorefunctions.