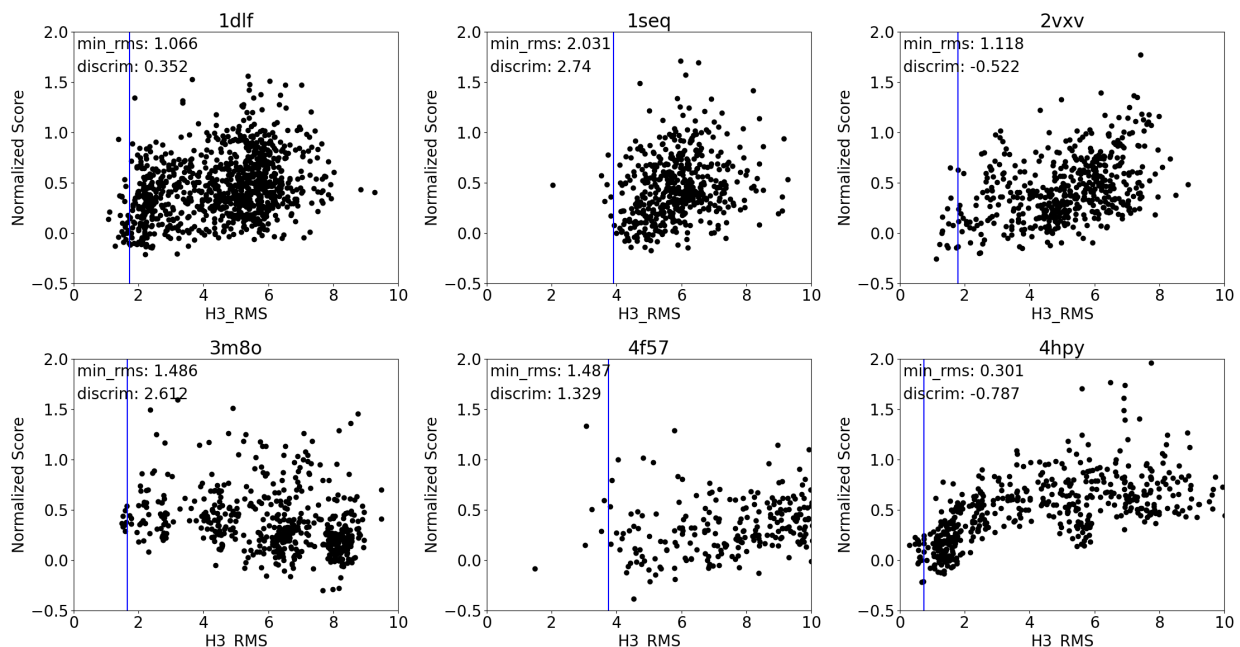# Scientific test: antibody_h3_modeling

FAILURES

None

RESULTS

AUTHOR AND DATE

Antibody CDR-H3 loop modeling scientific test first implemented June, 2019 by Jeliazko Jeliazkov (Gray Lab, JHU, jeliazkov@jhu.edu).

PURPOSE OF THE TEST

This test seeks to evalute the executable antibody_h3. This application models the CDR-H3 loop (typically on homology frameworks, but also on crystal structures). Here we test its ability to predict loop conformation on crystal structures by evaluating minimum rmsd and discrimination. Minimum RMSD indicates whether or not we are sampling native-like loop conformations. Discrimination indicates whether the score function can distinguish native from non-native. Both metrics are compared to 99-th percentile value determined by bootstrap resampling of previous simulations.

BENCHMARK DATASET

The dataset comprises six antibody targets from Weitzner and Gray (J. Immunol. 2016). The targets are of varying difficulty and CDR H3 length. 1DLF (12 residues) and 4HPY (13 residues) are easy targets. 1SEQ (16 residues) and 4F57 (18 residues) are hard targets. 2VXV (14 residues) and

3M8O (10 resiudes) are intermediate targets. Native structures are un-relaxed crystals. Inputs for modeling are relaxed with constraints (no ramping).

PROTOCOL

The Rosetta Antibody CDR-H3 loop modeling protocol is describe in our publication (Weitzner, Jeliazkov, Lyskov, et al., Nat. Protocol. 2017):

Briefly, this protocol has three stages (1) de novo CDR-H3 loop modeling, (2) VH-VL orientation refinement, (3) CDR-H3 loop refinement. Throughout the protocol the CDR-H3 loops is constrained to occupy the kink conformation observed in ~90% of antibodies. The benchmark is rather time consuming, taking ~1 hour per model. We aim to produce 500 models for 6 antibodies, so the runtime is 3000 CPU hours.

PERFORMANCE METRICS

To consider a run successful, for antibody_h3 run, we would like to produce a low energy, low rmsd model. Thus we need to assess two measures: (1) minimum rmsd achieved (do we sample a native like state) and (2) discrimination (do low rmsd models have low scores) We use discrimintaiton (typically lower is better, ideally negative values) as defined by Conway et al. Prot. Sci 2014, although our rmsd bins have adjusted spacing at min(rmsd) + (0, 0.5, 1, 1.5, 2, 4). We derive cutoffs for these metric from bootstrap resampling previous simulations (see below). The discrimnation or minimum rmsd test for each antibody is "passed" if the value is lower than that observed in the 99th percentile of our resampled simulations.

H3 Discrimation Scores Computed Over 2000 Simulations
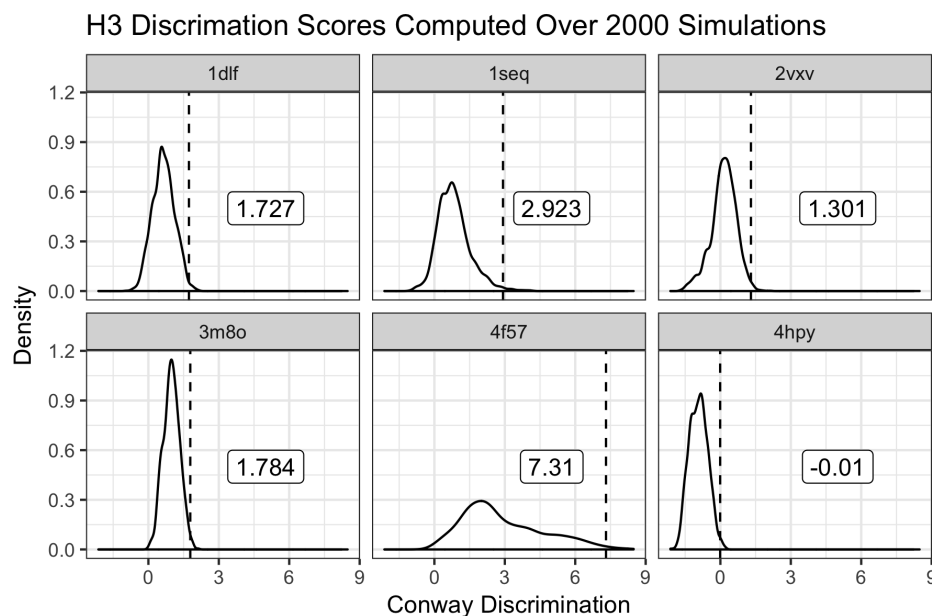
Minimum H3 RMSDs Computed Over 2000 Simulations
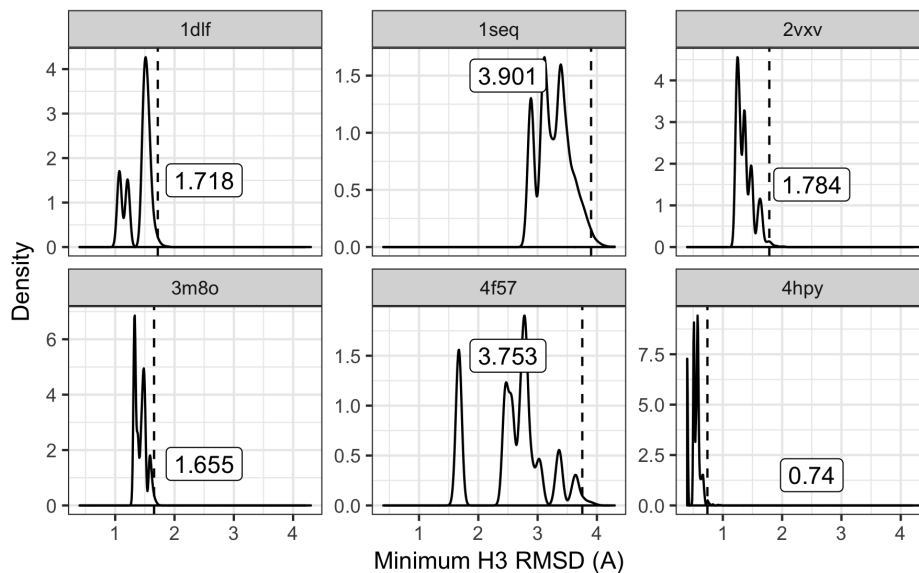
The exception is 3m8o whose discrimination score cutoff was adjusted after changes to the antibody code by Jeliazko in revision 61279 created a permanent scientific failure for this target. The cutoff was adjusted to 3.5 after looking at the discrimination score in the 12 consecutive failures of this test.

KEY RESULTS

Antibody CDR-H3 loop structure prediction is a challenging task. The current test assesses two easy (1DLF [12 residues] and 4HPY [13 residues]), two medium (2VXV [14 residues] and 3M8O [10 residues]), and two hard (1SEQ [16 residues] and 4F57 [18 residues]) targets. For the easy/medium cases we expect a minimum rmsd between 0 and 2 Angstroms, whereas for the hard cases the rmsd will be greater than that. We expect discrimination to always be below zero for 4hpy. For 1dlf and 2vxv, discrimination should be below 2. For the others discrimination will be worse. Unfortunately, discrimination is not the most stable metric, so we are exploring alternatives.

DEFINITIONS AND COMMENTS

This test is a first stab at an antibody CDR-H3 loop modeling scientific test. More work should be done to determine better targets (other Abs? more Abs? different protocol?) and metrics. In particular, discrimination varies across simulations.

LIMITATIONS

The full assessments for Rosetta Antibody typically cover the entire protocol from homology modeling to H3 modeling and consist of ~50 targets. Due to time considerations, we obviously do not test the full protocol over 50 antibodies here. Instead we focus on the crucial H3 modeling stage.