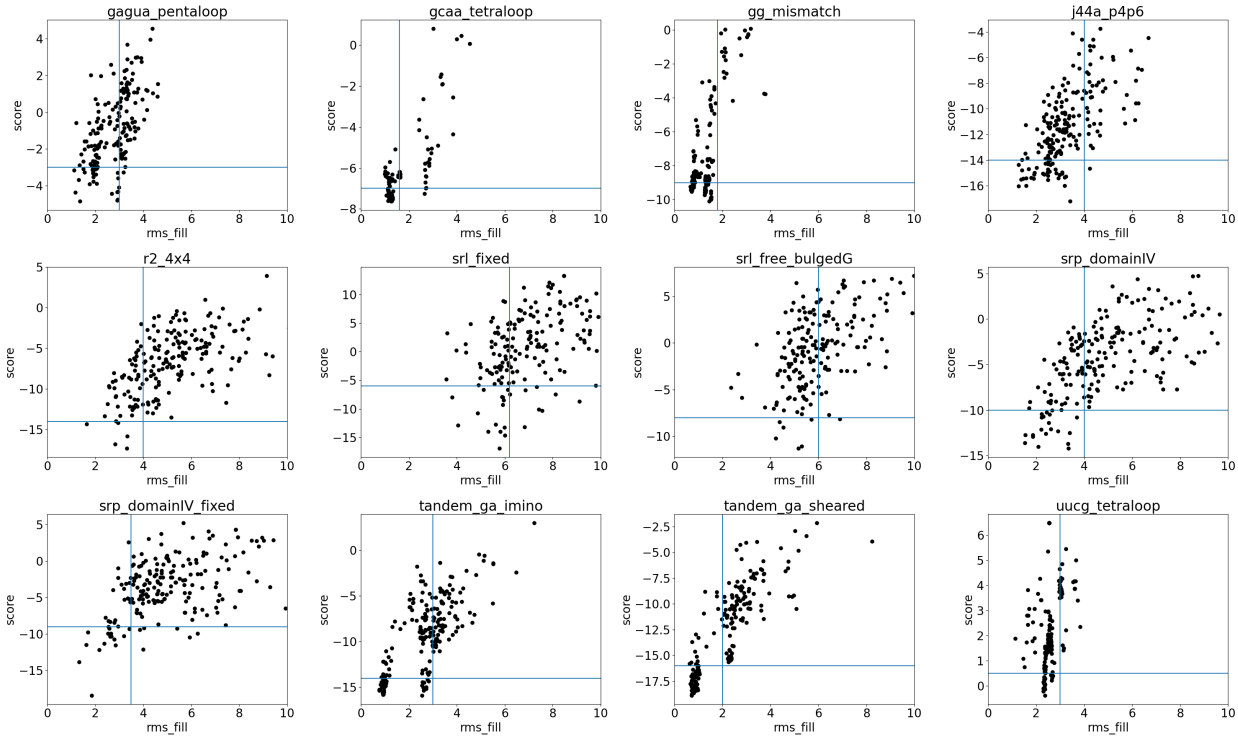
Scientific test: stepwise_rna_favorites

FAILURES

None

RESULTS



AUTHOR AND DATE

Andrew Watkins (andy.watkins2@gmail.com); Rhiju Das (October 2018)

PURPOSE OF THE TEST

This test judges the stability of the stepwise Monte Carlo algorithm's performance.

BENCHMARK DATASET

The benchmark set contains the 12 loop modeling problems from favorites.txt, a benchmark first established in Watkins et al., Science Advances 2018. The input files are ideal A-form RNA helices, fasta files, and portions of crystal structures.

PROTOCOL

See Watkins et al., Science Advances 2018 for a description of the SWM protocol. The benchmark takes 6 hours on the test server (~240 CPU-hours).

PERFORMANCE METRICS

Benchmark performance is based on the minimum RMSD sampled and minimum energy sampled; each case has a separate threshold. Since sampling is limited for each test, cutoffs are determined conservatively through repeated runs. Generally speaking, SWM should be able to achieve near-atomic accuracy in one of its top 5 lowest energy cluster centers if run for long enough on nearly every one of these cases; as is, we're just tracking some best-effort characteristics to see if the performance changes a *lot*.

KEY RESULTS

The majority of cases should display some strong funnel-like properties on the 'lower edge' of the score-versus-RMSD plot, with some outliers. That is, the lowest-energy structures for most models above the RMSD threshold should be much worse in energy than the lowest-energy structures for most models below the threshold. There are two expected "false minima": tandem_ga_imino should have a cluster of structures at about 2.9A RMSD; that cluster is sometimes slightly lower in energy than the near-native 0.9A RMSD cluster. Similarly, the gcaa_tetraloop should have a cluster just above the threshold that's only slightly worse in energy than the nativelike cluster.

DEFINITIONS AND COMMENTS

n/a

LIMITATIONS

The benchmark could be expanded to include favorites2.txt, challenges.txt, followups.txt, and more of the benchmarking challenges already developed for stepwise Monte Carlo. More CPU power would permit better sampling and therefore more precise cutoffs.