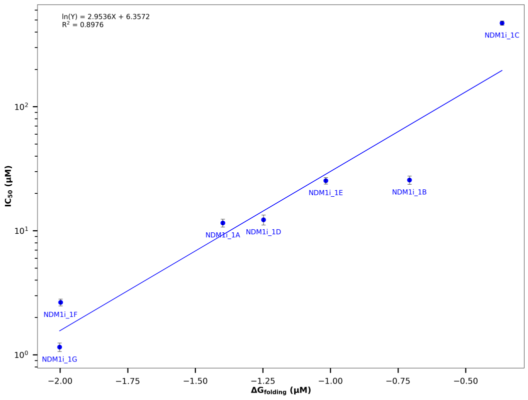
Scientific test: peptide_pnear_vs_ic50

FAILURES

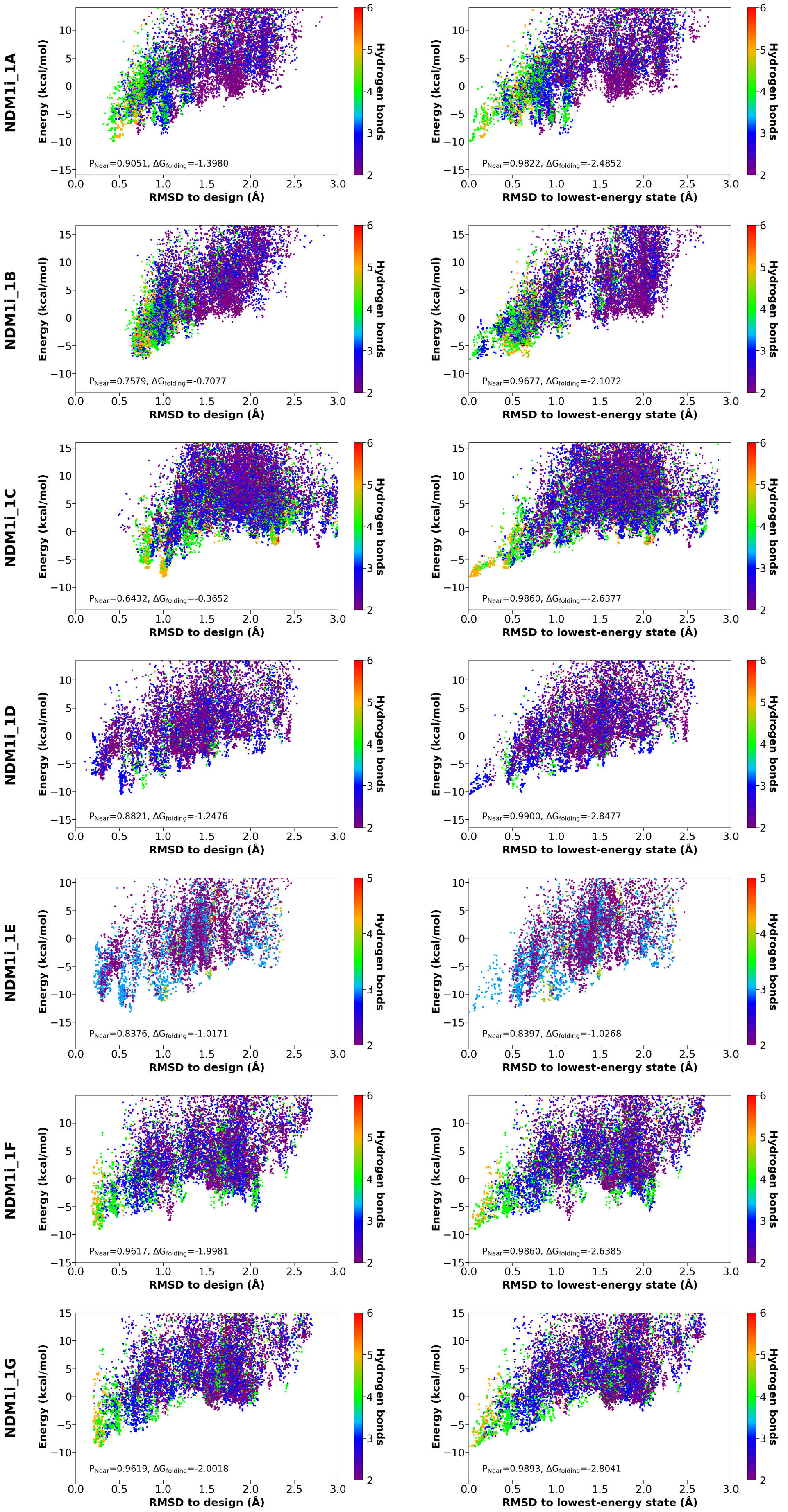
None.

RESULTS SUMMARY

Analysis of correlation of predicted folding propensity with experimentally-measured inhibition values (which should be linear):



Folding funnels:


NDM1i_1A
Energy (kcal/mol)
RMSD to design (Å)
RMSD to lowest-energy state (Å)
Hydrogen bonds
P_{Near}=0.9051, $\Delta G_{folding}$=-1.3980
P_{Near}=0.9822, $\Delta G_{folding}$=-2.4852
NDM1i_1B
P_{Near}=0.7579, $\Delta G_{folding}$=-0.7077
P_{Near}=0.9677, $\Delta G_{folding}$=-2.1072
NDM1i_1C
P_{Near}=0.6432, $\Delta G_{folding}$=-0.3652
P_{Near}=0.9860, $\Delta G_{folding}$=-2.6377
NDM1i_1D
P_{Near}=0.8821, $\Delta G_{folding}$=-1.2476
P_{Near}=0.9900, $\Delta G_{folding}$=-2.8477
NDM1i_1E
P_{Near}=0.8376, $\Delta G_{folding}$=-1.0171
P_{Near}=0.8397, $\Delta G_{folding}$=-1.0268
NDM1i_1F
P_{Near}=0.9617, $\Delta G_{folding}$=-1.9981
P_{Near}=0.9860, $\Delta G_{folding}$=-2.6385
NDM1i_1G
P_{Near}=0.9619, $\Delta G_{folding}$=-2.0018
P_{Near}=0.9893, $\Delta G_{folding}$=-2.8041

```
Peptide NDM1i_1A:
    Total samples = 22770
    Computed PNear = 0.9051
    Computed PNear to lowest E = 0.9822
    Computed DG_folding = -1.3980
    Computed DG_folding to lowest = -2.4852
    Lowest energy = -9.9485 kcal/mol
    RMSD of lowest energy = 0.427 Angstroms
    Lowest RMSD = 0.312 Angstroms
    Highest RMSD = 2.872 Angstroms
    Energy gap (minE>1.5A - minE) = 6.2886 kcal/mol

    More than 18,000 samples? YES
    PNear value over 0.83? YES
    PNear value to lowest E over 0.83? YES
    Lowest energy under 0.45 A RMSD? YES
    Sampling below expected lower threshold RMSD(0.45 A)? YES
    Sampling beyond 1.5 A RMSD? YES
    Sampling beyond 2.5 A RMSD? YES
    3+ kcal/mol energy gap? YES
    OVERALL PASS? YES

Peptide NDM1i_1B:
    Total samples = 26955
    Computed PNear = 0.7579
    Computed PNear to lowest E = 0.9677
    Computed DG_folding = -0.7077
    Computed DG_folding to lowest = -2.1072
    Lowest energy = -7.3760 kcal/mol
    RMSD of lowest energy = 0.773 Angstroms
    Lowest RMSD = 0.570 Angstroms
    Highest RMSD = 3.027 Angstroms
    Energy gap (minE>1.5A - minE) = 6.6448 kcal/mol

    More than 18,000 samples? YES
    PNear value over 0.68? YES
    PNear value to lowest E over 0.68? YES
    Lowest energy under 0.85 A RMSD? YES
    Sampling below expected lower threshold RMSD(0.7 A)? YES
    Sampling beyond 1.5 A RMSD? YES
    Sampling beyond 2.4 A RMSD? YES
    3+ kcal/mol energy gap? YES
    OVERALL PASS? YES

Peptide NDM1i_1C:
    Total samples = 39545
    Computed PNear = 0.6432
    Computed PNear to lowest E = 0.9860
    Computed DG_folding = -0.3652
    Computed DG_folding to lowest = -2.6377
    Lowest energy = -8.0555 kcal/mol
    RMSD of lowest energy = 0.999 Angstroms
    Lowest RMSD = 0.492 Angstroms
    Highest RMSD = 3.134 Angstroms
    Energy gap (minE>1.5A - minE) = 5.2344 kcal/mol

    More than 18,000 samples? YES
    PNear value over 0.57? YES
    PNear value to lowest E over 0.57? YES
    Lowest energy under 1.05 A RMSD? YES
    Sampling below expected lower threshold RMSD(0.55 A)? YES
    Sampling beyond 1.5 A RMSD? YES
    Sampling beyond 2.9 A RMSD? YES
    3+ kcal/mol energy gap? YES
    OVERALL PASS? YES

Peptide NDM1i_1D:
    Total samples = 22368
    Computed PNear = 0.8821
    Computed PNear to lowest E = 0.9900
    Computed DG_folding = -1.2476
    Computed DG_folding to lowest = -2.8477
    Lowest energy = -10.4433 kcal/mol
    RMSD of lowest energy = 0.521 Angstroms
    Lowest RMSD = 0.108 Angstroms
    Highest RMSD = 2.637 Angstroms
    Energy gap (minE>1.5A - minE) = 5.0747 kcal/mol

    More than 18,000 samples? YES
    PNear value over 0.8? YES
    PNear value to lowest E over 0.8? YES
    Lowest energy under 0.55 A RMSD? YES
    Sampling below expected lower threshold RMSD(0.25 A)? YES
    Sampling beyond 1.5 A RMSD? YES
    Sampling beyond 2.4 A RMSD? YES
    3+ kcal/mol energy gap? YES
    OVERALL PASS? YES

Peptide NDM1i_1E:
    Total samples = 25523
    Computed PNear = 0.8376
    Computed PNear to lowest E = 0.8397
    Computed DG_folding = -1.0171
    Computed DG_folding to lowest = -1.0268
    Lowest energy = -13.1072 kcal/mol
    RMSD of lowest energy = 0.623 Angstroms
    Lowest RMSD = 0.205 Angstroms
    Highest RMSD = 2.595 Angstroms
    Energy gap (minE>1.5A - minE) = 5.8261 kcal/mol

    More than 18,000 samples? YES
    PNear value over 0.75? YES
    PNear value to lowest E over 0.75? YES
    Lowest energy under 0.65 A RMSD? YES
    Sampling below expected lower threshold RMSD(0.3 A)? YES
    Sampling beyond 1.5 A RMSD? YES
    Sampling beyond 2.2 A RMSD? YES
    3+ kcal/mol energy gap? YES
    OVERALL PASS? YES

Peptide NDM1i_1F:
    Total samples = 23408
    Computed PNear = 0.9617
    Computed PNear to lowest E = 0.9860
    Computed DG_folding = -1.9981
    Computed DG_folding to lowest = -2.6385
    Lowest energy = -9.0343 kcal/mol
    RMSD of lowest energy = 0.260 Angstroms
    Lowest RMSD = 0.178 Angstroms
    Highest RMSD = 2.794 Angstroms
    Energy gap (minE>1.5A - minE) = 3.2386 kcal/mol

    More than 18,000 samples? YES
    PNear value over 0.88? YES
    PNear value to lowest E over 0.88? YES
    Lowest energy under 0.32 A RMSD? YES
    Sampling below expected lower threshold RMSD(0.22 A)? YES
    Sampling beyond 1.5 A RMSD? YES
    Sampling beyond 2.5 A RMSD? YES
    3+ kcal/mol energy gap? YES
    OVERALL PASS? YES

Peptide NDM1i_1G:
    Total samples = 23864
    Computed PNear = 0.9619
    Computed PNear to lowest E = 0.9893
    Computed DG_folding = -2.0018
    Computed DG_folding to lowest = -2.8041
    Lowest energy = -8.9909 kcal/mol
    RMSD of lowest energy = 0.274 Angstroms
    Lowest RMSD = 0.192 Angstroms
    Highest RMSD = 2.997 Angstroms
    Energy gap (minE>1.5A - minE) = 6.4392 kcal/mol

    More than 18,000 samples? YES
    PNear value over 0.9? YES
    PNear value to lowest E over 0.9? YES
    Lowest energy under 0.32 A RMSD? YES
    Sampling below expected lower threshold RMSD(0.25 A)? YES
    Sampling beyond 1.5 A RMSD? YES
    Sampling beyond 2.6 A RMSD? YES
    5+ kcal/mol energy gap? YES
    OVERALL PASS? YES

Correlation analysis:
    R-squared value from fitting greater than 0.85? YES
```

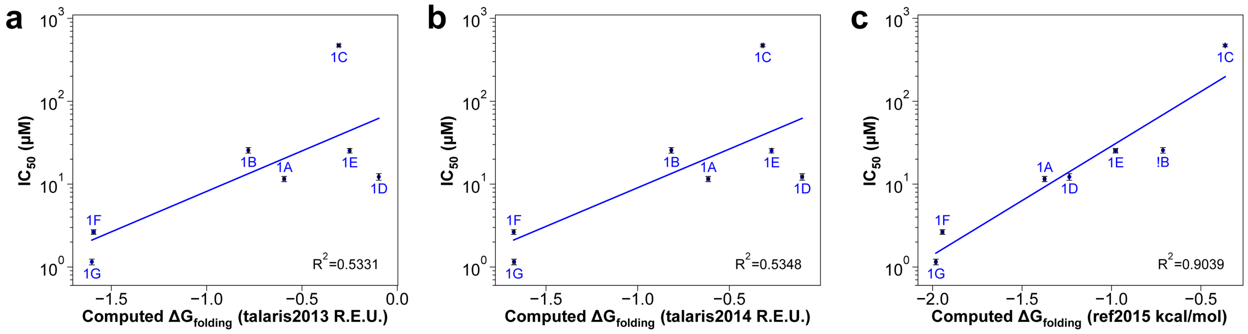
AUTHOR

Vikram K. Mulligan (vmulligan@flatironinstitute.org), Center for Computational Biology, Flatiron Institute, 5 June 2020.

PURPOSE OF THE TEST

When designing peptides to bind to targets, the flexibility of the peptide is a major impediment to binding. This is because there is an entropic cost associated with ordering a disordered molecule on binding. A rigidly-structured peptide that is pre-organized in the binding-competent conformation has a lower entropic cost and can bind more tightly. When interactions between peptide and target have been optimized with Rosetta, peptide rigidity becomes the major determininant of binding affinity. Rosetta predictions of peptide folding propensity, carried out with the simple_cycpep_predict application, correlate strongly with experimentally-measured binding affinity, at least with the ref2015 energy function. This correlation has improved from talaris2013 through talaris2014 to ref2015 due to improved training of the energy function against physical properties of fluids and against fluid MD simulation (see image below). This test exists to ensure that this correlation between prediction and experiment is maintained with future versions of the energy function and with future releases of Rosetta.

Improvements from talaris2013 through ref2015:



What does the benchmark test and why?

This test benchmarks seven peptides that were previously designed to bind to and inhibit the New Delhi metallo-beta-lactamase 1 (NDM-1). (For our purposes, IC50 values for these competitive inhibitors, measured at a constant concentration of the substrate of the reaction being inhibited, are proportional to inhibitor binding affinity.) The test carries out the following steps:

1. Large-scale conformational sampling for each of the seven peptides.

2. From the energy landscape sampled, calculation of the metrics PNear and estimated Delta-G of folding (which are closely-relarted measures of propensity to be rigid in the binding-competent conformation).

3. Plotting of the logarithm of experimentally-measured IC50 values against the computed Delta-G of folding values computed in the previous step.

4. Fitting of these data to a simple linear model, and extraction of the R-squared value for the fit.

The R-squared value should not drop below 0.85.

BENCHMARK DATASET

How many proteins are in the set?

- Seven peptides, named NDM1i-1A through NDM1i-1G. (These were previously called 8res1B, 8res2, 8res3, Holmes, Holmes_l8p, Adler, and Moriarty, respectively). All have been synthesized, the IC50 value for each binding to and inhibiting the NDM-1 enzyme's hydrolysis of nitrocefin have been measured.

What dataset are you using? Is it published? If yes, please add a citation.

- These peptides are described in Mulligan _et al._ (2020). Computationally-designed peptide macrocycle inhibitors of New Delhi metallo-beta-lactamase 1. Manuscript in preparation.

What are the input files? How were the they created?

- The input for each structure prediction run is the peptide, designed with Rosetta, in PDB format, along with its sequence in ASCII text format.

- The input for the correlation analysis is the set of output Delta-G of folding values from the structure prediction runs, plus an experimentally-measured set of IC50 values.

PROTOCOL

State and briefly describe the protocol.

The simple_cycpep_predict application uses the generalized kinematic closure algorithm (GenKIC) to rapidly sample closed conformations of a heteropolymer macrocycle built from any combination of alpha-amino acids, peptoids, or other related building-blocks. Each closure attempt is relaxed using the FastRelax protocol. For small (~8 to ~10 residue) peptide macrocycles, the application can usually sample close to the native state with less expense than protein _ab initio_.

The simple_cycpep_predict application supports hierarchical MPI-based job distribution and data reduction, as well as multi-threaded parallel job execution within a node. In MPI mode, statistics about the full sampled ensemble, including the funnel quality metric PNear, are computed automatically during the data collection and reduction phase.

Is there a publication that describes the protocol?

The simple_cycpep_predict application is described in the following publications:

1. Bhardwaj G, Mulligan VK, Bahl CD, _et al._ (2016). Accurate de novo design of hyperstable constrained peptides. _Nature_ 538(7625):329-35.

2. Hosseinzadeh P, Bhardwaj G, Mulligan VK, _et al._ (2018). Comprehensive computational design of ordered peptide macrocycles. _Science_ 358(6369): 1461-6.

3. Mulligan _et al._ (2020). Computationally-designed peptide macrocycle inhibitors of New Delhi metallo-beta-lactamase 1. Manuscript in preparation.

How many CPU hours does this benchmark take approximately?

Approximately 400 CPU-hours. The test runs on 4 nodes, 20 cores per node, for roughly 5 wall hours.

In debug mode, this test takes approximately 10 CPU-hours (1 node, 20 cores, for 0.5 wall hours).

PERFORMANCE METRICS

What are the performance metrics used and why were they chosen?

All of the following must be true for the test to pass:

- More than 18,000 successful samples (200 in debug mode).

- Lowest-RMSD sample within a threshold RMSD of the design model. (Varies by peptide.)

- Highest-RMSD sample outside of a thrshold RMSD of the design model. (Varies by peptide.)

- Lowest-energy sample within a threshold RMSD of the design model. (Varies by peptide.)

- Energy gap (gap between lowest-energy sample > 1.5 A and overall lowest) bigger than a threshold value. (Varies by peptide.)

- PNear greater than a threshold value. (Varies by peptide.)

- Most importantly, R-squared value for the linear relationship between ln(IC50) and Delta-G of folding is greater than 0.85.

How do you define a pass/fail for this test?

Failure of any of the above results in an overall failure.

How were any cutoffs defined?

These are based on Rosetta performance as of 5 June 2020. The observed R-squared value reported in Mulligan _et al._ (2020) is 0.9.

KEY RESULTS

What is the baseline to compare things to - experimental data or a previous Rosetta protocol?

Past iterations of this test and performance reported in Mulligan _et al._ (2020). (See Figure 3 in that paper.)

Describe outliers in the dataset.

NDM1i-1C should have the lowest PNear value and highest Delta-G of folding; NDM1i-1F and NDM1i-1G should have the highest PNear value and lowest Delta-G of folding. These are the worst and best binders, respectively.

DEFINITIONS AND COMMENTS

State anything you think is important for someone else to replicate your results.

For full reproducibility, see the 92-page supplement to Mulligan _et al._ (2020).

LIMITATIONS

What are the limitations of the benchmark? Consider dataset, quality measures, protocol etc.

The parameters of the correlation are expected to vary from target to target, and for any given target, there are very few data. (Chemically synthesizing a candidate peptide binder and testing its affinity are inherently low-throughput experimental techniques.)

How could the benchmark be improved?

More peptides, and more targets.

What goals should be hit to make this a "good" benchmark?

Well, I think it's pretty good as it is.