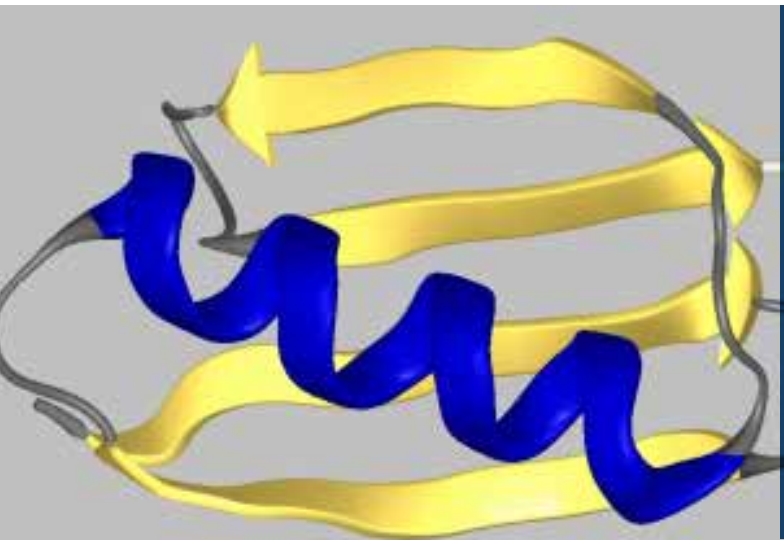
DNASTAR
SOFTWARE FOR LIFE SCIENTISTS

608.258.7420 tel
866.511.5090 U.S, toll free
0.808.234.1643 U.K.
0.800.182.4747 Germany

3801 Regent Street
Madison, WI 53705

www.dnastar.com
info@dnastar.com



Benchmarks

HOT SPOT SCANNING IN LASERGENE PROTEIN

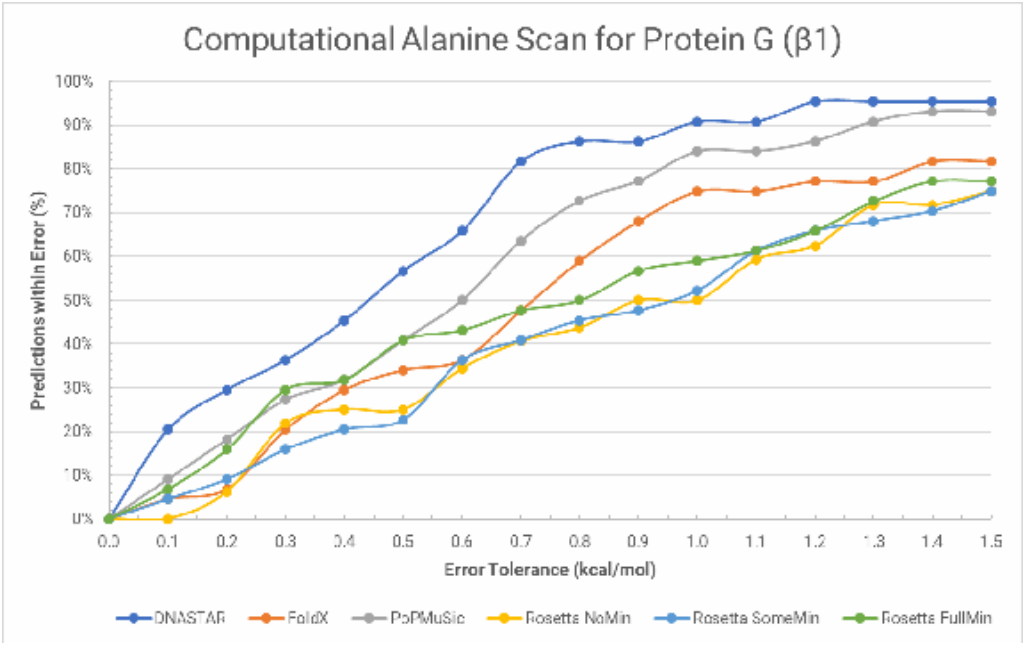
The best performing tool for identifying critical residues in protein folding

Overview

Alanine hot spot scanning can be used to identify important residues in protein folding, and is an important first step in many protein design experiments. In this analysis we evaluate six software tools for computational alanine scanning, including the protein design functionality in DNASTAR's Lasergene Protein.

By comparing experimentally determined thermodynamic stability data[1] for alanine substitution mutations at nearly every position in the β1 domain of Streptococcal protein G (G β1) to in silico calculations from each tool, we demonstrate that tools vary widely in the accuracy of hot spot detection, especially at low error tolerances.

The Lasergene Protein alanine hotspot scanning method provides the most accurate prediction of energy change in the G β1 protein, with the tightest tolerance of any tool studied.



The graph above describes the percent of correct hotspot predictions in the dataset within a given error tolerance. DNASTAR predictions from Lasergene Protein (top in blue), are the most accurate across all tolerances, even at the lowest error thresholds.

Study Highlights

- **Based on our error tolerance analysis, Lasergene Protein predictions are the most accurate in terms of error of predicting the real change in energy value.** This error analysis considers absolute error (the magnitude of the difference between the predicted and actual change in fold stability).
- For a set of 44 alanine variants in the G β1 data set, Lasergene Protein predictions have a Pearson linear correlation coefficient of 0.72 for predicted versus actual changes in fold stability, well ahead of FoldX and three Rosetta methods (at 0.47, 0.30, 0.49, and 0.61, respectively) and comparable to PopMusic at 0.75.
- Lasergene Protein is also shown to have the lowest error at the hot spots with the largest energy changes, making it a reliable predictor of true hot spots.

			Absolute Error in Predicted Energy Change						Experimental
Position	Variant	Accessible	DNASTAR	FoldX	PoPMuSic	Rosetta MoMin	Rosetta ScreenMin	Rosetta FullMin	ΔΔG
25	T25A	Surface	0.7	1.6	0.9	1.2	1.6	1.9	-0.9
42	D42A	Surface	0.4	0.6	0.7		0.5	0.4	-0.7
35	N35A	Surface	0.4	0.8	0.5	0.1	0.3	0.1	-0.6
28	K28A	Surface	0.3	0.8	1.0	0.6	0.6	0.7	-0.6
14	G14A	Surface	0.1	2.6	2.3	20.1	1.5	2.5	-0.5
41	G41A	Surface	0.3	5.7	0.8		5.2	5.3	-0.5
21	V21A	Surface	0.6	0.4	0.2	0.2	0.6	1.1	0.4
15	E15A	Surface	0.1	0.8	0.3	0.7	1.5	0.4	-0.4
27	E27A	Surface	0.0	0.9	0.4	0.7	0.7	0.7	-0.4
36	D36A	Surface	0.3	0.6	0.1	0.2	0.0	0.1	0.4
44	T44A	Surface	0.5	0.6	0.7		0.6	0.3	-0.4
37	N37A	Surface	0.1	1.4	0.7	1.3	1.0	1.6	-0.4
16	T16A	Surface	0.1	0.9	1.2	2.6	2.6	2.4	-0.3
11	I11A	Surface	0.2	0.6	0.6	0.8	0.9	0.2	-0.2
47	D47A	Surface	0.0	1.8	0.5		1.0	0.6	0.2
18	T18A	Boundary	0.0	0.3	1.6	1.1	1.5	0.3	-0.2
10	K10A	Surface	0.6	0.4	0.7	0.6	1.1	0.7	-0.2
13	K13A	Surface	0.0	1.0	0.7	1.3	1.4	1.4	0.2
32	Q32A	Surface	0.5	0.5	0.2	0.4	0.1	0.3	-0.2
1	M1A	Surface	0.3	0.2	0.7	0.3	0.2	0.5	-0.1
29	V29A	Surface	0.7	0.3	0.5	1.1	1.1	1.4	0.1
40	D40A	Surface	0.3	0.0	0.1	1.1	0.4	0.0	-0.1
19	F19A	Surface	0.6	0.9	0.4	0.3	0.3	0.1	0.0
56	D56A	Surface	0.7	0.5	0.2		1.4	1.3	0.2
50	K50A	Surface	0.0	1.2	1.1		0.5	0.2	0.2
17	T17A	Surface	0.3	0.2	0.2	1.0	0.1	0.2	0.2
9	G9A	Core	0.7	0.0	0.1	0.5	1.0	0.5	0.3
12	L12A	Surface	0.7	1.6	1.3	2.3	2.3	1.7	0.3
8	N8A	Surface	0.6	0.3	0.3	0.3	0.5	0.6	0.3
4	K4A	Surface	0.7	1.0	0.5	0.9	1.6	1.0	0.3
31	K31A	Surface	0.4	0.3	0.2	0.3	1.2	0.4	0.4
49	T49A	Surface	0.5	0.3	0.2		0.8	0.1	0.5
55	T55A	Surface	0.7	0.7	0.4		0.2	0.2	0.7
22	D22A	Surface	1.6	4.6	0.0	0.5	1.1	1.2	1.1
51	T51A	Boundary	1.1	0.5	0.5		0.4	0.2	1.1
39	V39A	Boundary	0.6	0.8	1.0	2.5	2.1	3.4	1.2
53	T53A	Surface	1.1	0.7	0.1		0.7	0.3	1.2
6	I6A	Surface	0.1	0.0	0.0	1.5	1.1	1.6	1.3
7	L7A	Core	0.4	1.9	1.4	1.6	1.5	1.3	1.4
33	Y33A	Surface	0.1	0.2	0.6	1.4	1.1	1.9	1.8
45	D45A	Surface	1.8	1.4	1.0		0.7	0.8	1.8
54	V54A	Core	1.0	0.9	0.7		0.3	1.1	1.9
5	L5A	Core	0.7	2.5	1.5	4.0	4.1	3.5	2.0
30	F30A	Core	0.9	0.6	0.5	4.9	3.5	3.7	3.0

Most Stabilizing

Most Destabilizing

Workflow

The following steps for hot spot scanning and protein design can be completed on virtually any Mac or Windows computer in just a few minutes

01

Open structure

Load structure file in Protean 3D

02

Scan for hot spots

Choose from alanine or serine scanning in our Protein Design module

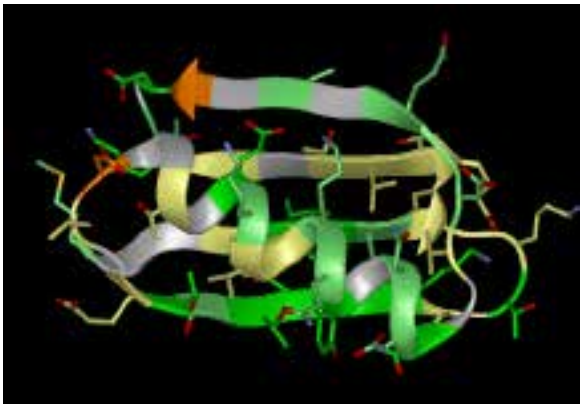
03

Test Variants

Model additional variants at positions of interest to test other hypotheses in silico

Variant Analysis

In the chart above, the experimental energy change (ΔΔG) between the wild type and variant is compared to the calculated energy change for six different methods. Alanine variants at each of 44 positions within the G β1 are sorted by the experimental energy change value, with the most stabilizing mutations at the top. The magnitude of absolute error for each of the scanning tools is indicated by color, green being lowest error and red being the highest error. The color for absolute error for DNASTAR hotspot predictions is also mapped onto the G β1 structure file at the right.



Free Trial

Obtain a fully functional, free trial version of Lasergene to try this workflow on your laptop or desktop computer.

www.dnastar.com/freetrial

References

1. Nisthal A, et al. (2019) PNAS. 116(33):16367-16377. https://www.protabank.org/study_analysis/gwoS2haU3

DNASTAR

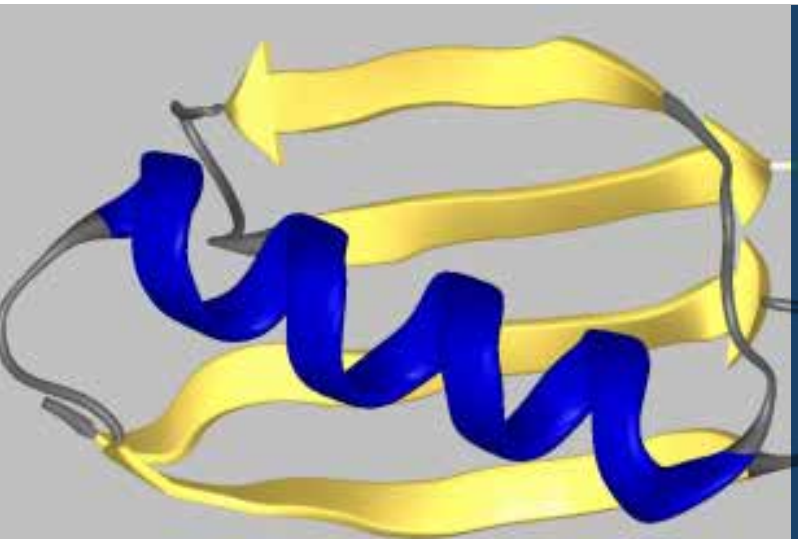
SOFTWARE FOR LIFE SCIENTISTS

DNASTAR
SOFTWARE FOR LIFE SCIENTISTS

608.258.7420 tel
866.511.5090 U.S, toll free
0.808.234.1643 U.K.
0.800.182.4747 Germany

3801 Regent Street
Madison, WI 53705

www.dnastar.com
info@dnastar.com



Benchmarks

HOT SPOT SCANNING IN LASERGENE PROTEIN

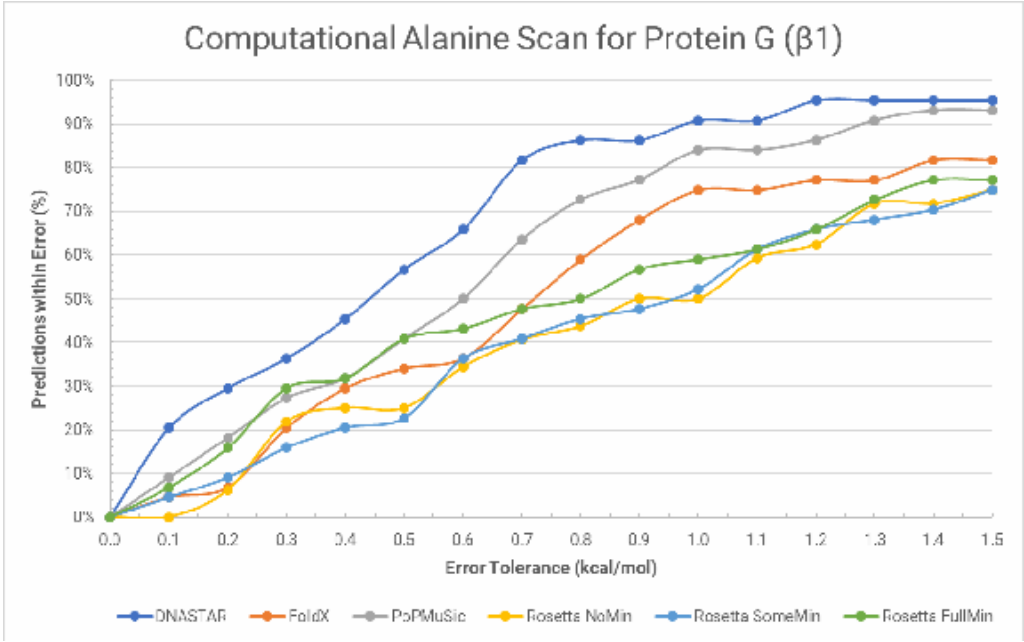
The best performing tool for identifying critical residues in protein folding

Overview

Alanine hot spot scanning can be used to identify important residues in protein folding, and is an important first step in many protein design experiments. In this analysis we evaluate six software tools for computational alanine scanning, including the protein design functionality in DNASTAR's Lasergene Protein.

By comparing experimentally determined thermodynamic stability data[1] for alanine substitution mutations at nearly every position in the β1 domain of Streptococcal protein G (G β1) to in silico calculations from each tool, we demonstrate that tools vary widely in the accuracy of hot spot detection, especially at low error tolerances.

The Lasergene Protein alanine hotspot scanning method provides the most accurate prediction of energy change in the G β1 protein, with the tightest tolerance of any tool studied.



The graph above describes the percent of correct hotspot predictions in the dataset within a given error tolerance. DNASTAR predictions from Lasergene Protein (top in blue), are the most accurate across all tolerances, even at the lowest error thresholds.

Study Highlights

- **Based on our error tolerance analysis, Lasergene Protein predictions are the most accurate in terms of error of predicting the real change in energy value.** This error analysis considers absolute error (the magnitude of the difference between the predicted and actual change in fold stability).
- For a set of 44 alanine variants in the G β1 data set, Lasergene Protein predictions have a Pearson linear correlation coefficient of 0.72 for predicted versus actual changes in fold stability, well ahead of FoldX and three Rosetta methods (at 0.47, 0.30, 0.49, and 0.61, respectively) and comparable to PopMusic at 0.75.
- Lasergene Protein is also shown to have the lowest error at the hot spots with the largest energy changes, making it a reliable predictor of true hot spots.

		Absolute Error in Predicted Energy Change						Experimental
Position	Variant	Accessible	DNASTAR	FoldX	PoPMuSic	Rosetta MoMin	Rosetta SomeMin	Rosetta FullMin	ΔΔG
25	T25A	Surface	0.7	1.6	0.9	1.2	1.6	1.9	-0.9
42	D42A	Surface	0.4	0.6	0.7		0.5	0.4	-0.7
35	N35A	Surface	0.4	0.8	0.5	0.1	0.3	0.7	-0.6
28	K28A	Surface	0.3	0.8	1.0	0.6	0.6	0.7	-0.6
14	G14A	Surface	0.1	2.0	2.3	20.1	1.5	2.5	-0.5
11	G11A	Surface	0.3	3.7	0.8		3.2	3.3	-0.5
21	V21A	Surface	0.6	0.4	0.2	0.2	0.6	1.1	0.4
15	E15A	Surface	0.1	0.8	0.8	0.7	1.5	0.4	-0.4
27	E27A	Surface	0.0	0.9	0.4	0.7	0.7	0.7	-0.4
26	D26A	Surface	0.3	0.6	0.1	0.2	0.0	0.1	0.4
44	T44A	Surface	0.5	0.0	0.7		0.6	0.3	-0.4
37	N37A	Surface	0.1	1.4	0.7	1.3	1.0	1.6	-0.4
16	T16A	Surface	0.1	0.9	1.2	2.6	2.6	2.4	-0.3
11	T11A	Surface	0.2	0.0	0.6	0.8	0.9	0.2	-0.2
47	D47A	Surface	0.0	1.8	0.5		1.0	0.8	0.2
18	T18A	Boundary	0.0	0.3	1.5	1.1	1.5	0.3	-0.2
10	K10A	Surface	0.6	0.4	0.7	0.6	1.1	0.7	-0.2
13	K13A	Surface	0.0	1.0	0.7	1.3	1.4	1.4	0.2
32	Q32A	Surface	0.5	0.5	0.2	0.4	0.1	0.3	-0.2
1	M1A	Surface	0.3	0.2	0.7	0.3	0.2	0.5	-0.1
29	V29A	Surface	0.7	0.3	0.5	1.1	1.1	1.4	0.1
40	D40A	Surface	0.3	0.0	0.1	1.1	0.4	0.3	-0.1
19	F19A	Surface	0.6	0.9	0.4	0.3	0.3	0.1	0.0
56	E56A	Surface	0.7	0.3	0.2		1.4	1.3	0.2
50	K50A	Surface	0.0	1.2	1.1		0.5	0.2	0.2
17	T17A	Surface	0.3	0.2	0.2	1.0	0.1	0.2	0.2
9	G9A	Core	0.7	0.0	0.1	0.5	1.0	0.5	0.3
12	L12A	Surface	0.7	1.6	1.3	2.3	2.3	1.1	0.3
8	N8A	Surface	0.6	0.3	0.3	0.3	0.5	0.6	0.3
4	K4A	Surface	0.7	1.0	0.5	0.9	1.6	1.0	0.3
31	K31A	Surface	0.4	0.3	0.2	0.3	1.2	0.4	0.4
49	T49A	Surface	0.5	0.3	0.2		0.8	0.1	0.6
55	T55A	Surface	0.7	0.7	0.4		0.2	0.3	0.7
22	D22A	Surface	1.6	4.6	0.0	0.5	1.1	1.2	1.1
51	T51A	Boundary	1.1	0.5	0.5		0.4	0.2	1.1
39	V39A	Boundary	0.6	0.8	1.0	2.5	2.1	3.4	1.2
53	T53A	Surface	1.1	0.7	0.1		0.7	0.3	1.2
6	I6A	Surface	0.1	0.0	0.0	1.5	1.1	1.6	1.3
7	L7A	Core	0.4	1.9	1.4	1.8	1.5	1.3	1.4
33	Y33A	Surface	0.1	0.2	0.6	1.4	1.1	1.9	1.8
45	D45A	Surface	1.8	1.4	1.0		0.7	0.8	1.8
54	V54A	Core	1.0	0.9	0.7		0.3	1.1	1.9
5	L5A	Core	0.1	2.5	1.5	4.0	4.1	3.5	2.0
30	F30A	Core	0.9	0.0	0.6	4.9	3.5	3.7	3.0

Most Stabilizing

Most Destabilizing

Workflow

The following steps for hot spot scanning and protein design can be completed on virtually any Mac or Windows computer in just a few minutes

01

Open structure

Load structure file in Protean 3D

02

Scan for hot spots

Choose from alanine or serine scanning in our Protein Design module

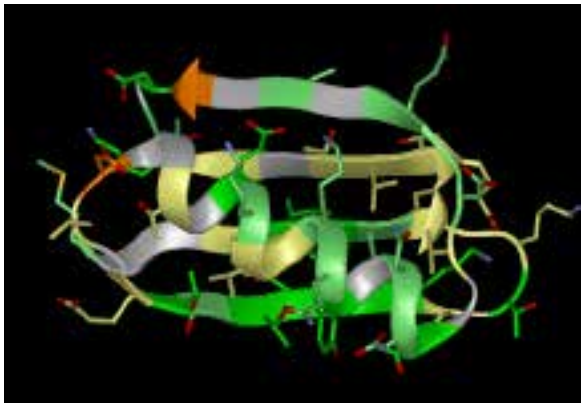
03

Test Variants

Model additional variants at positions of interest to test other hypotheses in silico

Variant Analysis

In the chart above, the experimental energy change (ΔΔG) between the wild type and variant is compared to the calculated energy change for six different methods. Alanine variants at each of 44 positions within the G β1 are sorted by the experimental energy change value, with the most stabilizing mutations at the top. The magnitude of absolute error for each of the scanning tools is indicated by color, green being lowest error and red being the highest error. The color for absolute error for DNASTAR hotspot predictions is also mapped onto the G β1 structure file at the right.



Free Trial

Obtain a fully functional, free trial version of Lasergene to try this workflow on your laptop or desktop computer.

www.dnastar.com/freetrial

DNASTAR

SOFTWARE FOR LIFE SCIENTISTS

References

1. Nisthal A, et al. (2019) PNAS. 116(33):16367-16377. https://www.protabank.org/study_analysis/gwoS2haU3