A comparative assessment and analysis of 20 representative sequence alignment methods for protein structure prediction

Renxiang Yan, Dong Xu, Jianyi Yang, Sara Walker, Yang Zhang

Department of Computational Medicine and Bioinformatics, University of Michigan, 100 Washtenaw Ave, Ann Arbor, MI 48109

SUPPLEMENTAL INFORMATION

Table S1. Performance of alignment methods of 137 Easy targets.

Methods[a]	TM-score[b]		RMSD (Å)[c]		Coverage[d]	
	First	Best in top10	First	Best in top10	First	Best in top 10
Profile-to-profile alignments						
HHsearch-I	0.707(0.736)	0.716(0.754)	5.0(8.4)	4.7(7.5)	0.896	0.894
HHsearch-II	0.699(0.732)	0.718(0.756)	5.1(8.5)	4.6(7.5)	0.890	0.895
PROSPECT	0.691(0.710)	0.705(0.740)	5.8(7.9)	5.2(7.0)	0.911	0.904
MUSTER	0.688(0.708)	0.716(0.747)	5.2(14.8)	4.4(8.8)	0.883	0.897
SP3	0.685(0.702)	0.714(0.746)	5.3(17.1)	4.4(8.8)	0.884	0.894
dPPAS	0.683(0.704)	0.712(0.742)	4.9(16.9)	4.6(8.8)	0.867	0.900
SPARKS	0.681(0.704)	0.714(0.746)	5.4(16.1)	4.4(8.7)	0.884	0.895
PPAS	0.679(0.703)	0.709(0.741)	5.2(16.6)	4.6(8.5)	0.872	0.896
PPA	0.676(0.699)	0.689(0.715)	5.4(16.5)	5.1(14.8)	0.872	0.876
PRC	0.666(0.692)	0.694(0.732)	5.3(14.1)	4.6(9.8)	0.867	0.869
FFAS	0.657(0.679)	0.695(0.724)	5.6(13.4)	4.8(9.6)	0.878	0.892
Average	**0.683(0.706)**	**0.707(0.740)**	**5.3(13.7)**	**4.7(9.1)**	**0.882**	**0.892**
Sequence-to-profile alignments						
SAM	0.638(0.654)	0.685(0.718)	5.1(25.0)	5.1(9.9)	0.823	0.876
PSI-BLAST	0.602(0.631)	0.652(0.688)	4.8(31.9)	4.2(19.6)	0.762	0.811
PSA	0.530(0.471)	0.493(0.513)	9.6(17.3)	11.0(12.6)	0.904	0.944
Average	**0.590(0.585)**	**0.610(0.640)**	**6.5(24.7)**	**6.8(14.0)**	**0.830**	**0.877**
Sequence-to-sequence alignments						
NW-align	0.567(0.588)	0.622(0.662)	6.5(29.0)	6.6(12.3)	0.813	0.868
SW-align	0.477(0.506)	0.555(0.596)	5.6(55.6)	5.1(30.7)	0.637	0.724
BLAST	0.459(0.486)	0.521(0.558)	5.0(62.8)	5.1(39.4)	0.598	0.678
Average	**0.501(0.527)**	**0.566(0.605)**	**5.7(49.1)**	**5.6(27.5)**	**0.683**	**0.757**
Other controls						
TM-align	0.779(0.809)	0.781(0.812)	2.5(5.4)	2.5(5.4)	0.901	0.899
MUSTER[SS+BTA+SA]	0.725(0.761)	0.731(0.770)	4.5(7.6)	4.3(7.6)	0.898	0.874
MUSTER[SS+BTA]	0.694(0.728)	0.697(0.746)	5.3(7.3)	4.8(7.4)	0.904	0.866
MUSTER[SS]	0.690(0.722)	0.694(0.737)	5.5(7.4)	5.1(7.2)	0.907	0.876

[a]Alignment methods as sorted by TM-score in each category.

[b]Average TM-score. Values in parentheses are for full-length models built by MODELLER. 'First' refers to the top-ranking model based on alignment score; 'Best in top10' to the best model of the highest TM-score among the top ten models with the highest alignment scores.

[c]RMSD to the native.

[d]Alignment coverage equals to the number of aligned residues divided by target length.

Table S2. Performance of alignment methods of 177 Medium targets.

Methods[a]	TM-score[b]		RMSD (Å)[c]		Coverage[d]	
	First	Best in top10	First	Best in top10	First	Best in top 10
Profile-to-profile alignments						
MUSTER	0.454(0.462)	0.508(0.532)	8.8(15.5)	7.5(13.5)	0.865	0.858
PPAS	0.443(0.455)	0.495(0.517)	8.4(18.7)	7.4(15.6)	0.819	0.830
SPARKS	0.438(0.439)	0.484(0.504)	9.6(14.8)	8.3(12.4)	0.886	0.871
dPPAS	0.437(0.446)	0.498(0.517)	7.6(24.0)	7.3(17.0)	0.792	0.827
HHsearch-II	0.436(0.446)	0.494(0.517)	8.7(19.7)	8.0(12.6)	0.784	0.832
SP3	0.434(0.440)	0.488(0.507)	9.7(14.0)	8.2(12.5)	0.870	0.867
HHsearch-I	0.430(0.445)	0.483(0.509)	8.5(19.0)	8.0(15.5)	0.769	0.815
PROSPECT	0.424(0.426)	0.488(0.503)	10.5(13.6)	8.8(10.9)	0.905	0.891
PPA	0.412(0.421)	0.471(0.491)	9.2(19.1)	7.8(16.3)	0.822	0.833
FFAS	0.408(0.415)	0.460(0.477)	7.9(24.8)	7.5(18.1)	0.749	0.783
PRC	0.386(0.395)	0.428(0.448)	6.9(32.0)	6.4(26.2)	0.665	0.687
Average	**0.427(0.435)**	**0.482(0.502)**	**8.7(19.5)**	**7.7(15.5)**	**0.811**	**0.827**
Sequence-to-profile alignments						
PSA	0.361(0.372)	0.413(0.436)	11.0(18.1)	9.7(15.9)	0.851	0.846
SAM	0.338(0.348)	0.416(0.429)	10.0(27.2)	8.6(21.8)	0.716	0.771
PSI-BLAST	0.280(0.293)	0.329(0.349)	6.8(51.1)	6.5(43.4)	0.485	0.530
Average	**0.326(0.338)**	**0.386(0.405)**	**9.3(32.1)**	**8.3(27.0)**	**0.684**	**0.716**
Sequence-to-sequence alignments						
NW-align	0.297(0.307)	0.361(0.381)	12.9(21.7)	11.0(16.8)	0.847	0.851
SW-align	0.254(0.269)	0.314(0.332)	9.5(46.6)	8.1(39.8)	0.580	0.602
BLAST	0.227(0.239)	0.275(0.291)	8.1(55.9)	7.3(47.4)	0.480	0.506
Average	**0.259(0.272)**	**0.317(0.335)**	**10.2(41.4)**	**8.8(34.7)**	**0.636**	**0.653**
Other controls						
TM-align	0.666(0.653)	0.668(0.680)	2.8(7.2)	2.8(6.7)	0.860	0.852
MUSTER[SS+BTA+SA]	0.491(0.511)	0.527(0.566)	6.5(17.9)	6.3(11.6)	0.774	0.802
MUSTER[SS+BTA]	0.471(0.494)	0.517(0.553)	8.0(14.3)	6.9(12.1)	0.813	0.817
MUSTER[SS]	0.465(0.486)	0.513(0.548)	8.2(14.1)	6.8(11.9)	0.820	0.821

[a]Alignment methods as sorted by TM-score in each category.

[b]Average TM-score. Values in parentheses are for full-length models built by MODELLER. 'First' refers to the top-ranking model based on alignment score; 'Best in top10' to the best model of the highest TM-score among the top ten models with the highest alignment scores.

[c]RMSD to the native.

[d]Alignment coverage equals to the number of aligned residues divided by target length.

Table S3. Performance of alignment methods of 224 Hard targets.

Methods[a]	TM-score[b]		RMSD (Å)[c]		Coverage[d]	
	First	Best in top10	First	Best in top10	First	Best in top 10
Profile-to-profile alignments						
MUSTER	0.264(0.280)	0.330(0.353)	14.6(15.4)	12.5(13.5)	0.878	0.875
dPPAS	0.260(0.268)	0.327(0.344)	14.1(20.2)	11.9(17.9)	0.810	0.823
SP3	0.258(0.274)	0.321(0.342)	15.0(16.1)	12.9(14.5)	0.869	0.864
HHsearch-II	0.257(0.278)	0.316(0.347)	13.0(31.3)	12.5(19.5)	0.679	0.765
PPAS	0.254(0.270)	0.311(0.337)	15.0(17.1)	12.8(14.8)	0.835	0.839
SPARKS	0.250(0.263)	0.308(0.330)	15.6(16.2)	13.4(14.0)	0.898	0.893
PROSPECT	0.246(0.257)	0.308(0.326)	16.0(16.5)	13.7(14.3)	0.923	0.911
HHsearch-I	0.242(0.264)	0.314(0.341)	13.3(28.7)	12.5(18.7)	0.676	0.771
FFAS	0.219(0.233)	0.277(0.296)	13.3(30.4)	11.8(25.2)	0.691	0.733
PPA	0.215(0.232)	0.281(0.301)	15.6(17.0)	14.1(15.4)	0.844	0.850
PRC	0.181(0.197)	0.239(0.261)	12.0(45.4)	11.4(31.7)	0.549	0.634
Average	**0.241(0.256)**	**0.303(0.325)**	**14.3(23.1)**	**12.7(18.1)**	**0.787**	**0.814**
Sequence-to-profile alignments						
PSA	0.203(0.218)	0.263(0.283)	16.4(17.3)	14.5(15.8)	0.863	0.850
SAM	0.168(0.185)	0.224(0.245)	14.5(29.4)	13.9(20.8)	0.653	0.723
PSI-BLAST	0.134(0.151)	0.168(0.190)	10.6(64.3)	10.2(55.0)	0.368	0.420
Average	**0.168(0.185)**	**0.218(0.239)**	**13.9(37.0)**	**12.9(30.5)**	**0.628**	**0.664**
Sequence-to-sequence alignments						
NW-align	0.191(0.204)	0.240(0.262)	16.6(17.3)	14.8(15.4)	0.871	0.877
SW-align	0.143(0.162)	0.191(0.210)	13.0(48.2)	12.7(35.8)	0.496	0.583
BLAST	0.131(0.146)	0.166(0.185)	11.0(60.9)	10.9(52.5)	0.382	0.456
Average	**0.155(0.171)**	**0.199(0.219)**	**13.5(42.1)**	**12.8(34.5)**	**0.583**	**0.639**
Other controls						
TM-align	0.586(0.583)	0.591(0.608)	3.6(9.0)	3.6(8.4)	0.826	0.810
MUSTER[SS+BTA+SA]	0.327(0.358)	0.379(0.424)	11.4(15.1)	9.6(13.0)	0.753	0.754
MUSTER[SS+BTA]	0.292(0.321)	0.355(0.394)	13.2(14.8)	11.0(13.1)	0.802	0.795
MUSTER[SS]	0.284(0.313)	0.349(0.385)	13.5(14.8)	11.3(13.3)	0.812	0.808

[a]Alignment methods as sorted by TM-score in each category.

[b]Average TM-score. Values in parentheses are for full-length models built by MODELLER. 'First' refers to the top-ranking model based on alignment score; 'Best in top10' to the best model of the highest TM-score among the top ten models with the highest alignment scores.

[c]RMSD to the native.

[d]Alignment coverage equals to the number of aligned residues divided by target length.