



Supplementary Information for

Ultrafast end-to-end protein structure prediction enables high-throughput exploration of uncharacterised proteins

Shaun M. Kandathil, Joe G. Greener, Andy M. Lau and David T. Jones

Correspondence to: David T. Jones

Email: d.t.jones@ucl.ac.uk

This PDF file includes:

Figures S1 to S7
Tables S1 to S5

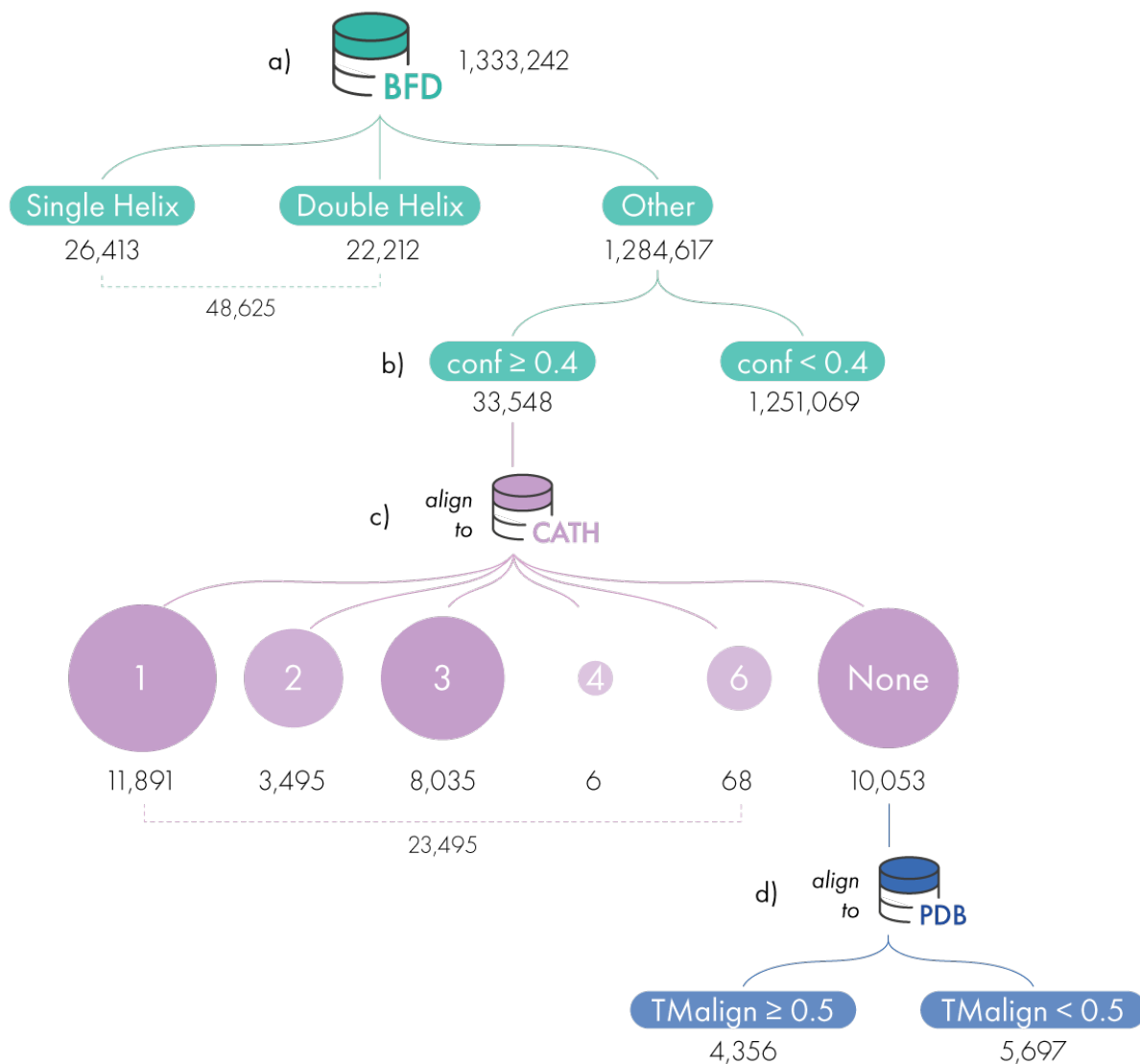


Figure S1. Filtering of BFD models. **a)** Models with folds corresponding to single and double helices were filtered from the ensemble to leave only complex folds. **b)** A confidence threshold of ≥ 0.4 was applied to isolate models likely to be correct. **c)** Each model was searched against known CATH domains. Models which matched CATH domains with a TM-score ≥ 0.5 calculated using TM-align were assigned to the corresponding CATH classes 1-4 and 6. **d)** The remaining ensemble not matched to CATH were searched against a representative set of PDB chains. The number of models at each stage of filtering is represented by the value proximal to each node.

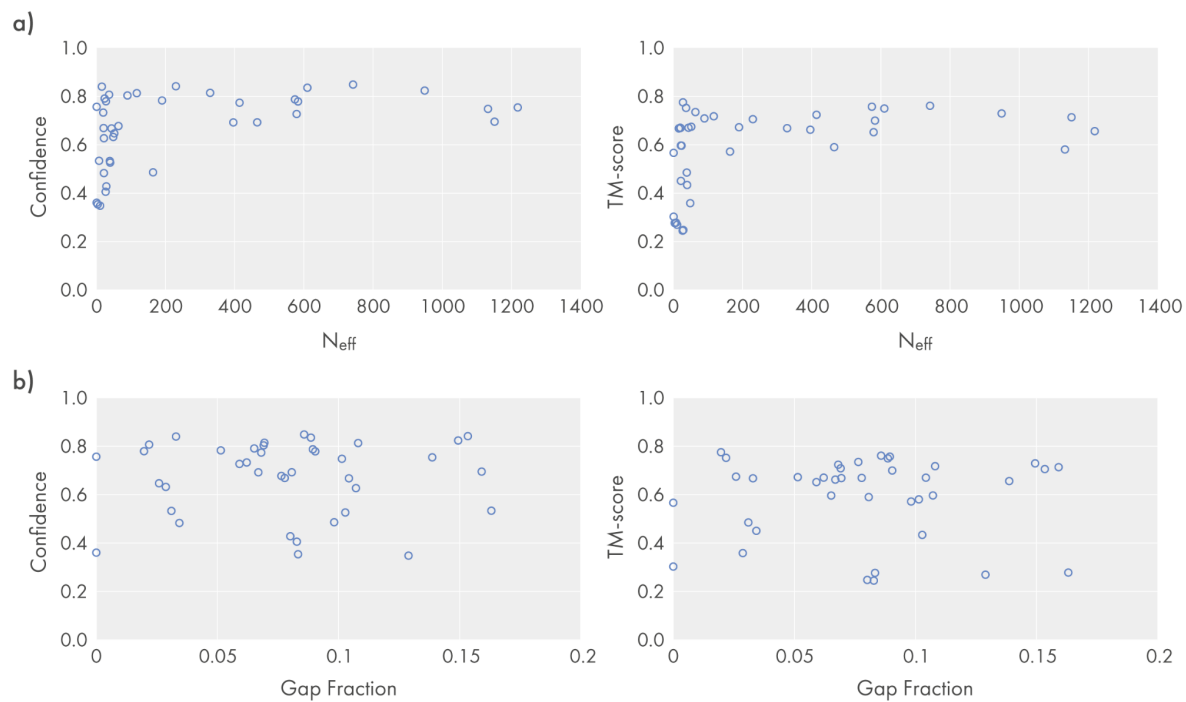


Figure S2. Effect of alignment quality metrics on network-predicted confidence score and TM-score. a) Effective sequence count (N_{eff}) and b) gap fraction is plot against confidence and TM-scores for 39 CASP13 FM and FM/TBM domains.

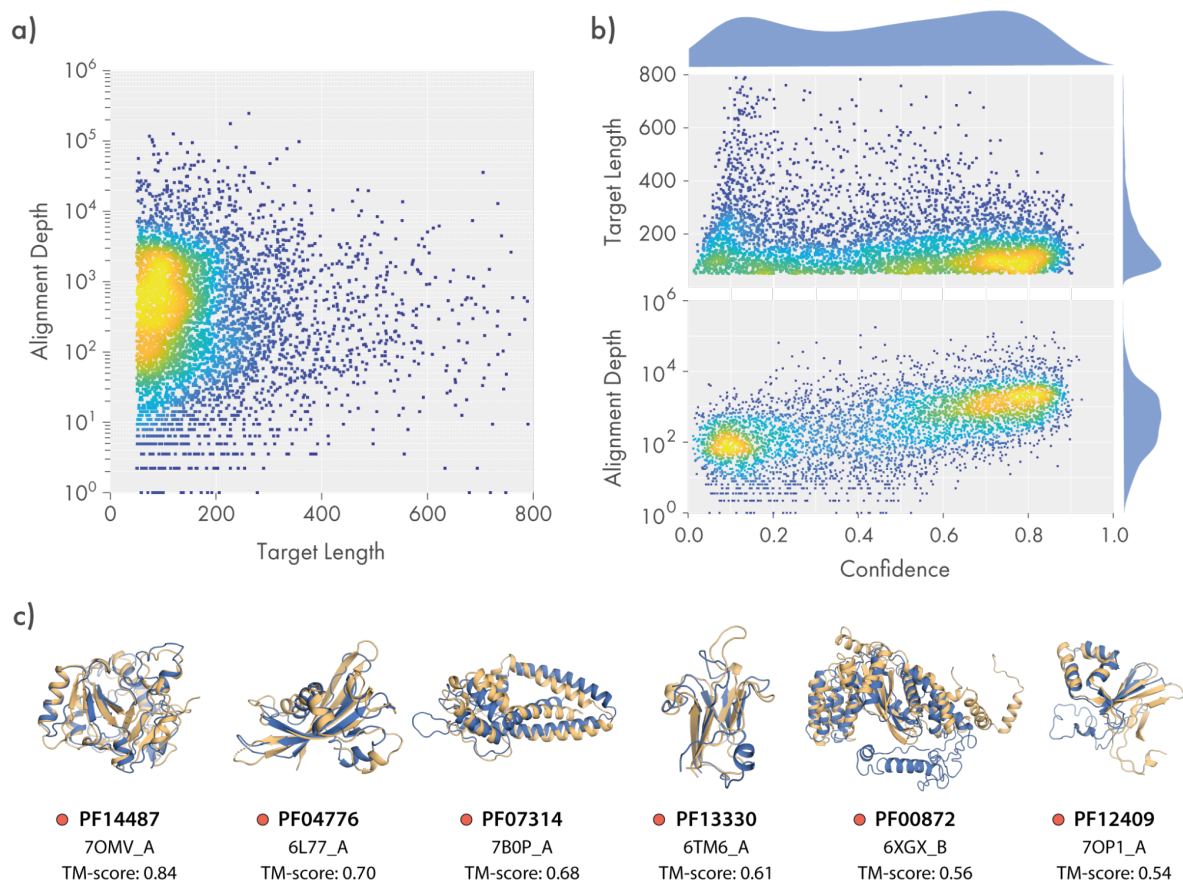


Figure S3. Pfam models generated by DMPfold2. Data represents 4,938 models from Pfam (version 32) generated from alignments used in DMPfold1. **a)** Summary of target length and alignment depth of 4,938 models. **b)** Distributions of model confidence against target length and alignment depth. Of the models generated, 1,988 (40.3%) and 2,950 (59.7%) have a confidence < 0.4 and ≥ 0.4 respectively. Color gradients indicate the density, with yellow and dark blue indicating highest and lowest respectively. **c)** Examples of DMPfold2 models for very hard targets. The mean TM-score across the shown targets are 0.655 for DMPfold2 and 0.564 for C-I-TASSER. Native structures and models are shown in orange and blue respectively.

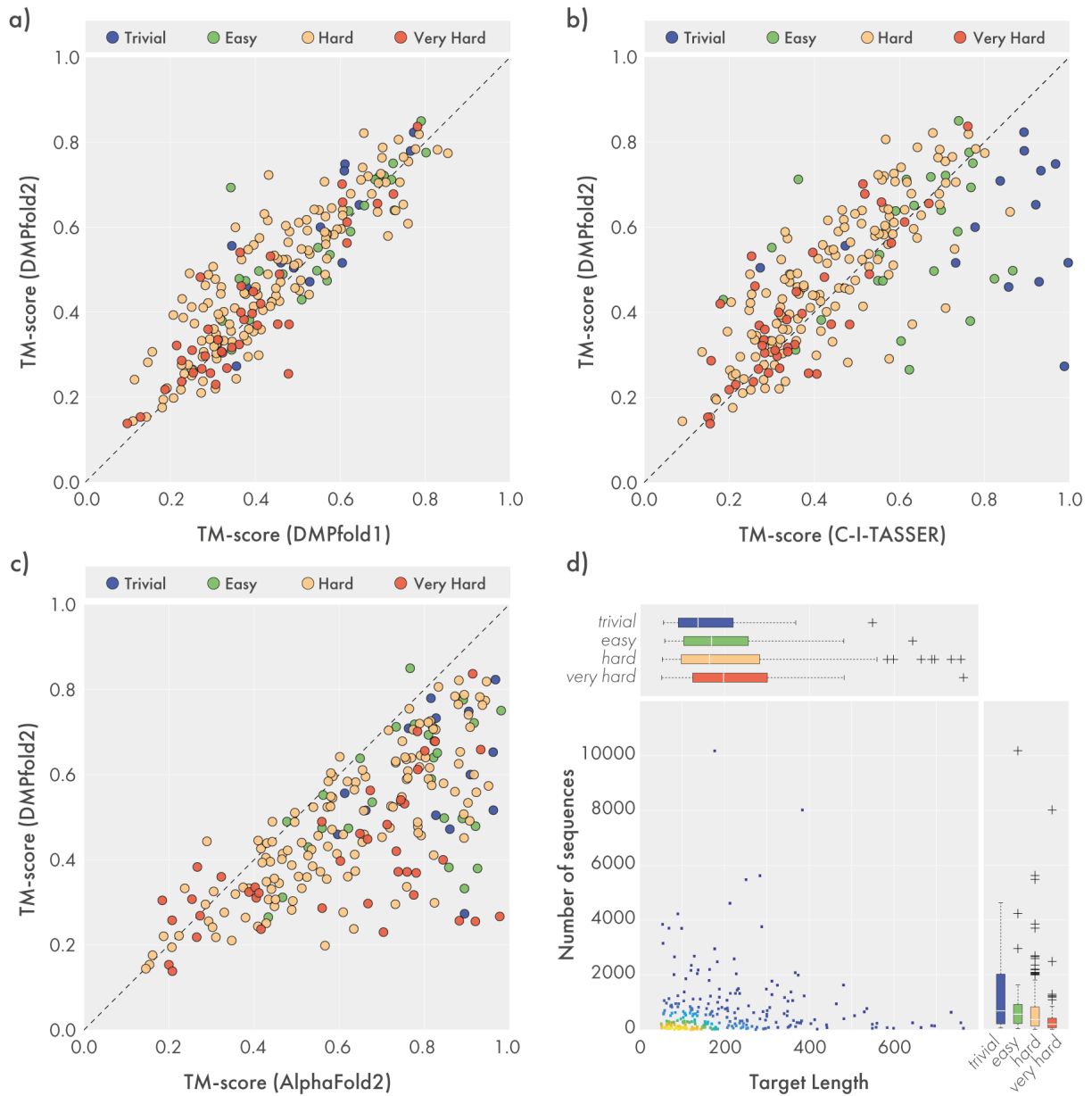


Figure S4. a-c) Comparison of Pfam families with recent structural annotations modelled by DMPfold1, DMPfold2, C-I-TASSER and AlphaFold. The colour of each data point represents the target difficulty as reported in the C-I-TASSER study (22) Only a common subset of 221 targets between all four methods were included. 34 targets were either not modelled in the original DMPfold1 publication or overlapped with the AlphaFold2 training set and were removed. A detailed list of TM-scores obtained by each method, difficulty and target can be found in Table S3 and S4. d) Distribution of target length and number of sequences in the alignment. Box plots show distributions when separated by difficulty classes.

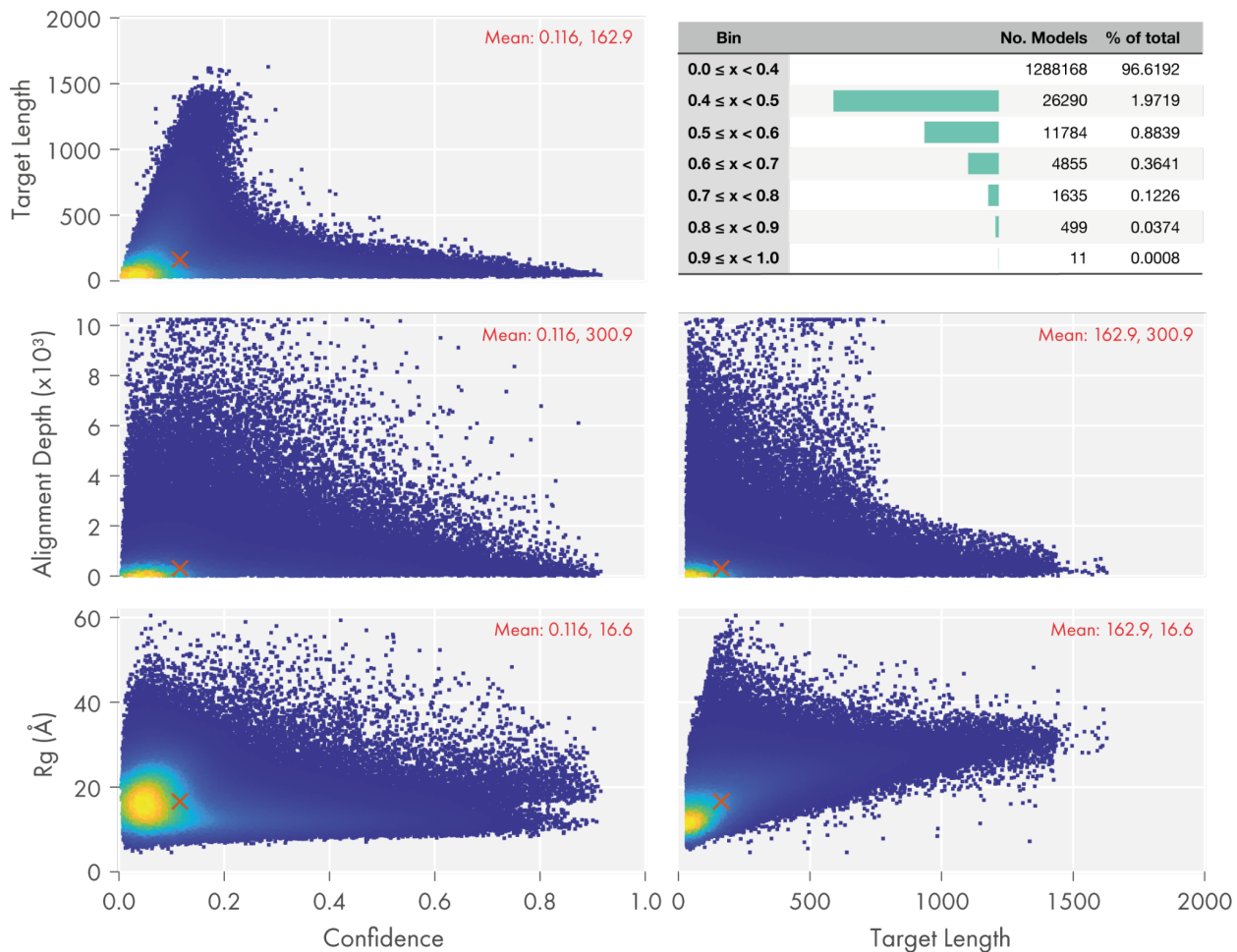


Figure S5. Summary of 1.3 million BFD models. Comparison of target length, alignment depth, radius of gyration (Rg) and model confidence for 1.3 million models. The red 'x' marks the mean of each variable, with the numerical values shown in the top right hand corner of each subplot. The table shows the number of models in each confidence bin. Color gradients indicate the density of datapoints, with yellow and dark blue indicating highest and lowest respectively.

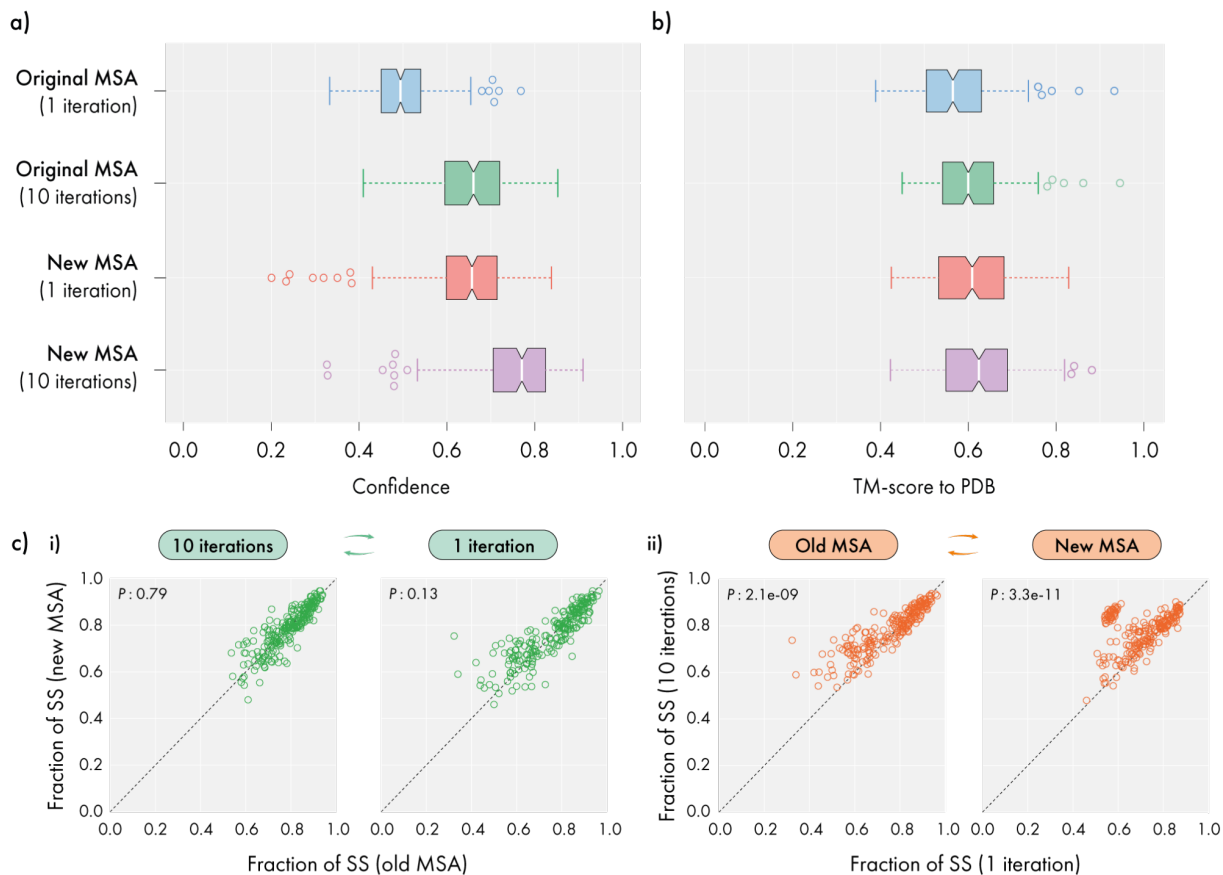


Figure S6. Effect of alignment regeneration and iterative modelling procedure. The models of 225 BFD targets were regenerated to assess the effect of improving the alignment and/or running the end-to-end procedure with 1 or 10 iterations. Box-and-whisker plot of **a)** confidence scores obtained from four different model generation schemes and **b)** the highest TM-score obtained between generated models and the closest match in a PDB representative set. **c)** Effect of modelling strategy on model secondary structure content. The fraction of secondary structure was calculated as the total number of residues assigned to the DSSP classes G, H, I, E, B, T or S (i.e. all classes excluding loop/irregular assignments), normalised by the target length. Values in each subplot represent the p-value as determined via a paired t-test.

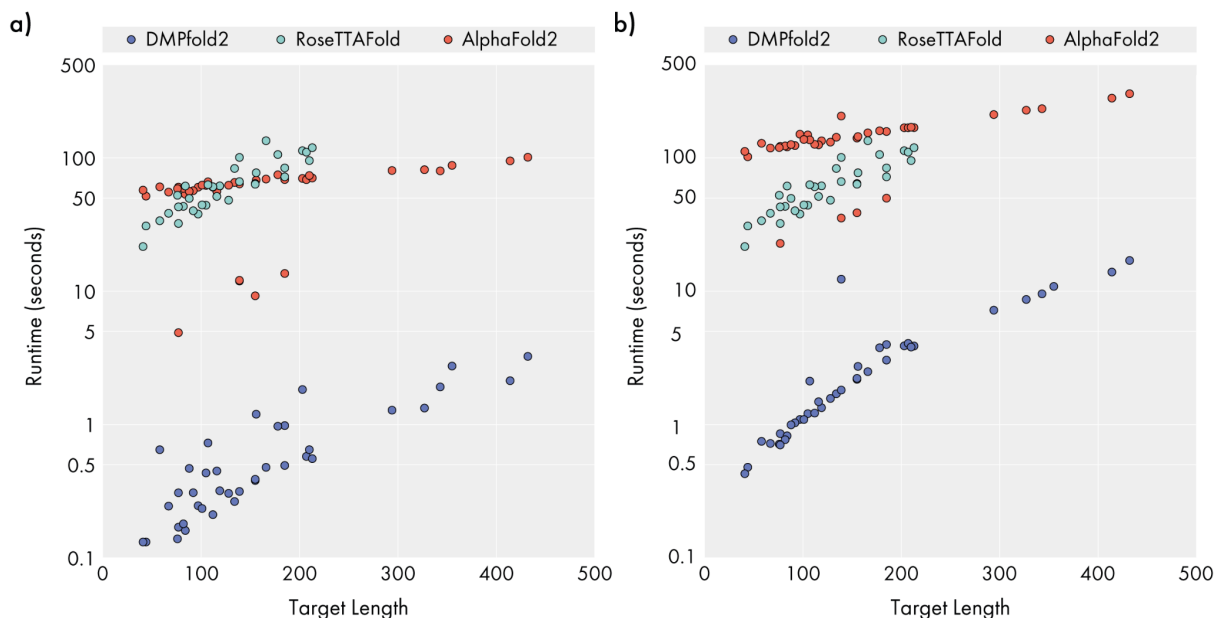


Figure S7. Runtimes on 39 CASP13 domains, comparing DMPfold2, RoseTTAFold and AlphaFold2, after a) 1 iteration and b) the default number of iterations of each method. DMPfold2 runs 10 iterations by default, RoseTTAFold's refinement module runs until convergence, and AlphaFold2 runs 3 iterations. All predictions were run using identical MSAs on a single NVIDIA GTX 1080Ti GPU. The largest six targets are omitted for RoseTTAFold due to memory limitations. DMPfold2 is roughly 2 orders of magnitude faster than AlphaFold2 when running a single iteration.

Table S1. Model accuracy on the 39 FM and FM/TBM CASP13 domains, comparing models built using either DMPfold1 or DMPfold2. In the case of DMPfold1, only the top-ranking model selected by each modelling run was evaluated for each target.

Target	DMPfold1			DMPfold2		
	GDT_HA	GDT_TS	TM-Score	GDT_HA	GDT_TS	TM-Score
T0949-D1	36.434	54.845	0.626	40.310	62.984	0.720
T0950-D1	24.342	40.863	0.610	19.956	35.526	0.578
T0953s1-D1	25.373	44.030	0.376	33.209	47.388	0.434
T0953s2-D1	23.295	34.659	0.256	45.455	65.909	0.451
T0955-D1	41.463	52.439	0.415	56.098	74.390	0.591
T0957s2-D1	36.774	58.226	0.690	31.129	51.613	0.615
T0958-D1	50.974	70.779	0.704	47.727	68.506	0.668
T0960-D2	38.690	57.143	0.564	15.179	24.405	0.249
T0963-D2	41.463	59.451	0.584	15.549	26.829	0.250
T0968s1-D1	46.186	67.373	0.742	41.102	62.924	0.718
T0968s2-D1	41.957	63.696	0.721	41.522	63.913	0.694
T0969-D1	19.986	37.641	0.633	25.989	45.056	0.697
T0970-D1	40.000	58.529	0.596	45.294	65.000	0.669
T0975-D1	19.217	33.185	0.507	24.555	43.683	0.653
T0978-D1	18.765	35.472	0.584	21.065	41.223	0.662
T0980s1-D1	18.750	30.288	0.310	21.875	33.654	0.356
T0981-D3	28.941	48.399	0.642	36.576	57.759	0.735
T0986s1-D1	47.011	68.750	0.726	39.402	61.141	0.662
T0986s2-D1	12.419	22.419	0.311	35.323	57.742	0.716
T0987-D1	22.568	40.676	0.547	31.216	53.108	0.687
T0987-D2	23.597	38.903	0.520	25.893	43.750	0.576
T0989-D1	24.440	37.500	0.443	9.701	19.590	0.244
T0989-D2	16.071	29.464	0.340	11.607	23.214	0.277
T0990-D1	44.408	65.461	0.611	52.961	74.013	0.745
T0990-D3	11.502	19.601	0.290	19.366	32.746	0.477
T0992-D1	50.467	72.430	0.773	48.364	69.860	0.761

T0997-D1	36.216	58.243	0.722	38.378	60.811	0.761
T0998-D1	10.843	18.825	0.269	13.554	21.988	0.282
T1000-D2	27.785	46.807	0.700	28.736	49.660	0.750
T1001-D1	28.237	46.763	0.560	35.072	57.374	0.672
T1005-D1	21.472	39.034	0.613	27.607	47.469	0.706
T1008-D1	19.481	27.597	0.266	24.026	37.338	0.304
T1010-D1	41.190	58.810	0.711	32.024	51.190	0.672
T1015s1-D1	27.273	42.898	0.452	34.659	54.261	0.559
T1017s2-D1	27.600	43.400	0.502	39.800	61.800	0.733
T1019s1-D1	33.621	53.879	0.433	49.569	70.259	0.652
T1021s3-D1	27.410	45.783	0.576	35.843	55.271	0.664
T1021s3-D2	13.918	26.031	0.287	50.258	72.423	0.775
T1022s1-D1	23.878	39.904	0.508	30.449	49.199	0.594
Mean	29.334	45.902	0.531	32.728	51.153	0.590

Table S2. Precision, recall and F1-score at different confidence thresholds. The confidence threshold that achieves the highest F1-score is shown in bold.

Confidence Threshold	Precision	Recall	F1-score
0.0	0.612	1.000	0.759
0.1	0.634	0.987	0.772
0.2	0.698	0.962	0.809
0.3	0.727	0.923	0.814
0.4	0.810	0.872	0.840
0.5	0.853	0.782	0.816
0.6	0.935	0.647	0.765
0.7	0.986	0.436	0.604
0.8	1.000	0.141	0.247

Table S3. Summary of model accuracy by target class for 221 Pfam targets comparing DMPfold1, DMPfold2, C-I-TASSER and AlphaFold2. Only targets which were common between all four methods were included. 34 targets were either not modelled by DMPfold1 or overlapped with the AlphaFold2 training set and were removed.

		Mean TM-score			
Difficulty	n	DMPfold1	DMPfold2	C-I-TASSER	AlphaFold2
Trivial	14	0.5578	0.5959	0.8192	0.8260
Easy	26	0.5410	0.5596	0.6037	0.7490
Hard	141	0.4402	0.4828	0.4412	0.6203
Very hard	40	0.3829	0.3954	0.3603	0.5987
Total	221	0.4491	0.4832	0.4696	0.6445

Table S4. TM-scores for 255 Pfam models generated by DMPfold1, DMPfold2, C-I-TASSER and AlphaFold2.

Target	Difficulty	TM-score			
		DMPfold1 ^a	DMPfold2	C-I-TASSER	AlphaFold2 ^b
PF00810	trivial	0.7404	0.7058	0.7326	0.8267
PF00811	trivial	0.8533	0.7740	0.8007	0.9432
PF01718	trivial	0.5690	0.4742	0.5496	0.5595
PF02405	trivial	0.3332	0.2684	0.3186	0.2730
PF04064	trivial	0.6160	0.6120	0.6124	0.7847
PF04727	trivial	0.5645	0.4885	0.5747	0.4952
PF05938	trivial	0.1437	0.1533	0.1549	0.1540
PF06995	trivial	0.3218	0.3795	0.7666	0.9251
PF08366	trivial	0.7302	0.6393	0.5121	0.7697
PF08742	trivial	0.7398	0.6902	0.9980	-
PF09447	trivial	0.7484	0.6534	0.6948	0.8844
PF09564	trivial	0.6057	0.6280	0.6512	0.7328
PF09856	trivial	0.4355	0.4141	0.4491	0.4670
PF10664	trivial	0.1470	0.2821	0.4788	0.4883
PF10716	trivial	0.3531	0.5998	0.4711	0.9170
PF11909	trivial	0.3961	0.3547	0.1958	0.5707
PF12534	trivial	0.3789	0.4594	0.8568	0.5964
PF13839	trivial	0.5499	0.6402	0.4654	0.8031
PF14065	trivial	0.8289	0.7824	0.6948	0.9305
PF14416	trivial	0.3745	0.3816	0.4419	-
PF02049	easy	0.6449	0.6528	0.9206	0.9618
PF03155	easy	0.4904	0.5045	0.2728	0.8271
PF03189	easy	0.6158	0.5628	0.5810	0.6729
PF03542	easy	0.5088	0.5486	0.7298	0.5803

PF04766	easy	0.2757	0.2954	0.2260	0.7319
PF05018	easy	0.5851	0.5822	0.5634	0.6386
PF05107	easy	0.2421	0.2766	0.2818	0.6093
PF05273	easy	0.4528	0.3718	0.4396	0.7387
PF05332	easy	0.3555	0.2730	0.9878	0.8946
PF05982	easy	0.2075	0.1981	0.1670	0.5666
PF06086	easy	0.3960	0.4138	0.5145	0.5946
PF06478	easy	0.6819	0.7123	0.3624	0.7339
PF06886	easy	0.2774	0.4107	0.3904	0.8220
PF08345	easy	0.3926	0.3970	0.3711	0.6027
PF08906	easy	0.4046	0.3691	0.2705	0.7809
PF09345	easy	0.7729	0.8232	0.8929	0.9673
PF09767	easy	-	0.2861	0.3145	0.2903
PF12031	easy	0.6568	0.6509	0.6345	0.8304
PF12333	easy	0.2528	0.2650	0.6234	0.4332
PF12509	easy	-	0.4378	0.3640	0.4887
PF13280	easy	0.6999	0.7876	0.6423	0.8939
PF13553	easy	0.3676	0.5239	0.3223	0.7263
PF14430	easy	0.6057	0.6587	0.5579	0.9326
PF14772	easy	0.2741	0.2666	0.2687	0.9774
PF15430	easy	0.3207	0.3072	0.3376	0.2538
PF15674	easy	0.5268	0.4414	0.4125	0.6541
PF16495	easy	0.7272	0.6399	0.6981	0.8219
PF17166	easy	0.3849	0.3603	0.3687	0.6202
PF01151	hard	0.1189	0.1382	0.1062	-
PF01366	hard	0.5652	0.6254	0.5765	0.8809
PF01578	hard	0.4742	0.5238	0.5490	0.5718
PF01673	hard	0.4214	0.6313	0.4957	0.7328

PF01763	hard	0.2890	0.3598	0.2836	0.3227
PF01803	hard	0.0986	0.1381	0.1550	0.2077
PF02106	hard	0.2783	0.3066	0.2986	0.4338
PF03311	hard	0.3516	0.4098	0.7087	0.5267
PF03348	hard	0.5659	0.5735	0.3368	0.9509
PF03511	hard	0.6096	0.7327	0.9326	0.8276
PF03581	hard	0.2513	0.3105	0.3077	0.4048
PF03586	hard	0.3060	0.4780	0.3304	0.7847
PF03609	hard	0.3834	0.5475	0.4448	0.7471
PF03613	hard	0.6556	0.8215	0.6786	0.8813
PF03626	hard	0.7773	0.7840	0.7795	0.8721
PF03963	hard	0.4297	0.4893	0.2487	0.4429
PF04031	hard	0.3117	0.3352	0.2816	0.4021
PF04061	hard	0.6875	0.6557	0.6693	0.8011
PF04109	hard	0.7853	0.8187	0.7623	0.9463
PF04112	hard	0.4134	0.3591	0.3373	0.5241
PF04143	hard	0.6899	0.7041	0.7091	0.7159
PF04213	hard	-	0.4227	0.6465	0.6580
PF04277	hard	0.3077	0.3483	0.2528	0.4290
PF04463	hard	0.6897	0.7087	0.8371	0.7629
PF04571	hard	0.4086	0.4969	0.6812	0.8894
PF04577	hard	0.6043	0.7014	0.5143	0.7838
PF04660	hard	0.6656	0.7198	0.5641	0.7890
PF04718	hard	0.7244	0.7504	0.7727	0.9805
PF04819	hard	0.7031	0.7212	0.7086	0.9171
PF05241	hard	0.6879	0.6359	0.8603	0.7577
PF05320	hard	-	0.4178	0.5999	0.9469
PF05615	hard	0.2278	0.4313	0.3248	0.5732

PF05680	hard	0.5093	0.4295	0.1858	0.5272
PF05751	hard	0.3671	0.4614	0.2605	0.6487
PF05914	hard	0.1296	0.1532	0.1497	0.1995
PF06022	hard	0.3067	0.2298	0.2152	0.7040
PF06098	hard	0.2069	0.3934	0.3211	0.4191
PF06101	hard	0.3771	0.4423	0.2526	0.4424
PF06333	hard	0.1122	0.1439	0.0897	0.1454
PF06417	hard	-	0.0581	0.0658	0.0667
PF06466	hard	0.1156	0.2414	0.2809	0.3021
PF06578	hard	0.7008	0.7114	0.9152	-
PF06664	hard	0.3423	0.6934	0.7683	0.8096
PF06760	hard	0.5673	0.5838	0.5844	0.6193
PF06761	hard	0.5452	0.4978	0.8665	0.8468
PF07019	hard	0.1574	0.3067	0.1360	0.4503
PF07426	hard	0.2466	0.2476	0.2668	-
PF07580	hard	0.3663	0.3997	0.3163	0.8441
PF07812	hard	0.3091	0.2905	0.5760	0.4478
PF07923	hard	0.5226	0.5923	0.5421	0.7611
PF07946	hard	0.2451	0.4913	0.3603	0.9068
PF07984	hard	0.3622	0.4788	0.8234	0.9205
PF08070	hard	0.4293	0.6160	0.4319	0.8125
PF08310	hard	0.4098	0.5747	0.6416	0.8174
PF08381	hard	0.5040	0.4511	0.3667	0.8934
PF08401	hard	0.2266	0.2761	0.3065	0.3298
PF08403	hard	0.2285	0.2207	0.2301	-
PF08586	hard	0.5533	0.5999	0.7779	0.9082
PF08691	hard	0.6060	0.6408	0.6124	0.7547
PF08740	hard	0.7252	0.6780	0.5187	0.8254

PF09135	hard	0.1664	0.2516	0.1877	-
PF09356	hard	0.2504	0.2639	0.2015	0.3797
PF09369	hard	-	0.3192	0.3138	0.9772
PF09429	hard	0.2706	0.4826	0.4237	0.7126
PF09432	hard	0.4447	0.5336	0.4498	0.9027
PF09684	hard	0.6742	0.6784	0.6937	0.7482
PF09766	hard	0.4528	0.5250	0.4820	0.5794
PF09777	hard	-	0.2545	0.2515	0.3070
PF09931	hard	0.3963	0.5583	0.4198	0.9222
PF10074	hard	-	0.5422	0.1997	0.5018
PF10117	hard	0.3187	0.4431	0.3635	0.2889
PF10169	hard	0.7596	0.6084	0.5109	0.7615
PF10171	hard	0.3442	0.3114	0.3561	0.4667
PF10172	hard	0.6110	0.7485	0.9671	0.9045
PF10198	hard	0.3006	0.5071	0.5286	0.8928
PF10221	hard	0.5208	0.6175	0.4693	0.7887
PF10270	hard	0.3739	0.4450	0.2866	0.4094
PF10572	hard	0.2248	0.2178	0.2506	0.3100
PF10657	hard	0.2503	0.4122	0.3203	0.6363
PF10806	hard	0.2303	0.3872	0.3262	0.4885
PF11203	hard	0.7656	0.7794	0.8935	0.8156
PF11210	hard	0.6199	0.6382	0.5911	0.6491
PF11235	hard	0.3829	0.2950	0.2617	0.2854
PF11256	hard	0.8021	0.7757	0.7638	0.8915
PF11262	hard	0.7123	0.5794	0.5427	0.8452
PF11315	hard	0.4915	0.5105	0.5855	0.4957
PF11390	hard	0.4179	0.4048	0.2469	0.4276
PF11594	hard	0.2889	0.4218	0.2498	0.4786

PF11597	hard	0.3447	0.5562	0.4728	0.6129
PF11718	hard	0.4601	0.5164	0.9967	0.9622
PF11779	hard	0.5069	0.4751	0.4963	0.7192
PF11801	hard	0.2501	0.2507	0.2079	0.4286
PF11876	hard	0.3861	0.3819	0.4163	0.8578
PF11882	hard	0.2930	0.3364	0.3201	0.7587
PF11938	hard	0.6197	0.6383	0.9864	-
PF11957	hard	0.5629	0.7072	0.5898	0.8211
PF12017	hard	0.5249	0.5453	0.4351	0.5843
PF12115	hard	0.4010	0.4594	0.3603	0.6586
PF12293	hard	0.1833	0.1943	0.1696	0.2065
PF12295	hard	0.2823	0.4632	0.3753	0.7857
PF12308	hard	0.2328	0.2552	0.2152	0.2930
PF12309	hard	0.1807	0.1756	0.2082	0.1612
PF12350	hard	-	0.7045	0.4892	0.9294
PF12357	hard	0.2768	0.2374	0.3320	0.6342
PF12378	hard	0.4594	0.4710	0.3419	0.5831
PF12410	hard	0.6246	0.6787	0.6281	0.8235
PF12432	hard	-	0.2938	0.3009	0.4044
PF12495	hard	-	0.6417	0.9549	0.9050
PF12603	hard	0.6047	0.5161	0.7323	0.6627
PF12803	hard	0.2266	0.2862	0.1574	0.5597
PF12871	hard	0.3083	0.3030	0.3256	0.4910
PF12922	hard	0.3257	0.4011	0.4286	0.4652
PF12926	hard	0.6955	0.7260	0.5871	0.8144
PF13117	hard	0.3060	0.2201	0.3175	0.1872
PF13157	hard	0.6527	0.7404	0.6040	0.9039
PF13182	hard	0.5684	0.5259	0.4088	0.7217

PF13331	hard	-	0.2558	0.2670	0.5307
PF13339	hard	0.4060	0.4402	0.3963	0.5118
PF13827	hard	0.7371	0.8062	0.5678	0.8332
PF13843	hard	0.3081	0.3254	0.3141	0.3048
PF13862	hard	0.3392	0.3719	0.6303	0.5557
PF13864	hard	0.4570	0.4894	0.5294	0.5590
PF13868	hard	0.2148	0.3218	0.2777	0.4111
PF13992	hard	0.2738	0.3327	0.3002	0.2369
PF14051	hard	0.5634	0.5890	0.3314	0.7583
PF14260	hard	0.5282	0.5370	0.5284	0.7450
PF14393	hard	0.7489	0.7651	0.7246	0.8799
PF14423	hard	0.3025	0.3605	0.3258	0.5044
PF14484	hard	0.5282	0.4717	0.9284	0.8601
PF14585	hard	0.4659	0.5234	0.5126	0.5391
PF14687	hard	0.3261	0.3329	0.3577	0.5120
PF14738	hard	0.4014	0.4530	0.4826	0.7007
PF14777	hard	0.6242	0.5897	0.7371	0.8164
PF14798	hard	0.1920	0.2212	0.2843	0.2232
PF14837	hard	0.2726	0.2100	0.2315	0.3473
PF14838	hard	0.6020	0.5961	0.6336	0.5852
PF14960	hard	-	0.3143	0.3054	-
PF15017	hard	0.5024	0.6108	0.5871	0.6548
PF15019	hard	0.5755	0.5352	0.5533	0.6775
PF15104	hard	0.3454	0.3308	0.4428	0.4405
PF15184	hard	0.4936	0.4569	0.4576	0.7925
PF15300	hard	-	0.4169	0.3936	0.4409
PF15387	hard	0.3642	0.5405	0.3964	0.7447
PF15448	hard	-	0.3808	0.3350	0.5466

PF15865	hard	0.5389	0.6417	0.5740	0.6013
PF15933	hard	-	0.5461	0.4420	0.5681
PF16134	hard	0.3174	0.3323	0.6039	0.8944
PF16192	hard	0.2189	0.2712	0.7316	-
PF16234	hard	0.3735	0.3829	0.3351	0.2664
PF16327	hard	0.3429	0.3558	0.4155	0.4474
PF16498	hard	0.3171	0.3847	0.3613	0.5374
PF17081	hard	-	0.3015	0.3442	-
PF17250	hard	0.7072	0.7057	0.6179	0.7734
PF18145	hard	0.3733	0.3447	0.2641	0.4587
PF18755	hard	0.3199	0.4593	0.3757	0.6092
PF00872	very hard	0.4314	0.7228	0.5503	0.8092
PF01677	very hard	0.6906	0.7181	0.6738	0.7772
PF02107	very hard	0.3811	0.3329	0.2963	0.3874
PF02119	very hard	0.7819	0.8369	0.7608	0.9137
PF03113	very hard	-	0.4917	0.3035	0.4429
PF03257	very hard	0.2820	0.2970	0.3127	0.6676
PF03381	very hard	0.3951	0.4484	0.3579	0.6667
PF03396	very hard	0.3620	0.3240	0.3542	0.3879
PF03509	very hard	0.5927	0.5952	0.5688	0.6618
PF03700	very hard	0.4779	0.2551	0.4058	0.9200
PF03839	very hard	0.4389	0.4237	0.2045	0.5349
PF04081	very hard	0.4649	0.4896	0.3420	0.4771
PF04281	very hard	0.3028	0.3193	0.3018	0.3558
PF04295	very hard	0.4126	0.4200	0.1787	0.7342
PF04776	very hard	0.2895	0.4881	0.3444	0.7843
PF05348	very hard	0.4718	0.6035	0.5661	0.5803
PF05422	very hard	-	0.3145	0.3441	0.8370

PF05518	very hard	0.2268	0.2367	0.2567	0.4161
PF05831	very hard	0.4794	0.3714	0.4830	0.7596
PF06138	very hard	0.4774	0.5646	0.4848	0.6357
PF07314	very hard	0.2944	0.2566	0.3864	0.8824
PF07365	very hard	0.3544	0.3506	0.2929	-
PF07774	very hard	0.4079	0.2988	0.4048	0.8239
PF07982	very hard	-	0.2829	0.2502	0.3892
PF08039	very hard	0.4485	0.4564	0.4722	0.5505
PF10216	very hard	-	0.2744	0.1987	0.7174
PF10744	very hard	0.3445	0.3172	0.3375	0.7752
PF11289	very hard	0.3128	0.3350	0.2602	0.3743
PF11704	very hard	0.4487	0.4390	0.4020	0.4301
PF11711	very hard	0.3129	0.3954	0.3413	0.4380
PF12409	very hard	0.3770	0.4738	0.5610	0.6216
PF12546	very hard	-	0.5967	0.4753	0.4787
PF13299	very hard	0.5477	0.5520	0.3001	0.5623
PF13330	very hard	0.1875	0.2180	0.2001	0.2649
PF14487	very hard	-	0.1566	0.1360	0.1074
PF14637	very hard	0.5659	0.5598	0.5735	0.8721
PF14638	very hard	0.6489	0.7115	0.5592	0.9077
PF14665	very hard	0.2567	0.3732	0.3111	0.6377
PF14728	very hard	0.5799	0.6450	0.5635	0.7912
PF14782	very hard	0.6983	0.7635	0.6909	0.9366
PF14892	very hard	0.4806	0.5899	0.4564	0.8096
PF14939	very hard	0.7904	0.8499	0.7393	0.7667
PF15014	very hard	0.4570	0.4260	0.4146	0.4159
PF15097	very hard	0.3228	0.3044	0.2834	0.1840
PF15122	very hard	0.2539	0.2577	0.2923	0.2071

PF15146	very hard	0.4357	0.5317	0.2524	0.7537
PF15693	very hard	0.5007	0.5151	0.4302	0.6577
PF15694	very hard	0.7206	0.7120	0.6166	0.8197
PF16500	very hard	-	0.3584	0.3648	0.3846
PF16564	very hard	0.3551	0.2433	0.2343	0.4087
PF16932	very hard	0.7606	0.7547	0.7080	0.7647
PF17229	very hard	0.5695	0.5841	0.5524	0.5744

a 24 families with missing scores were not modelled in the original DMPfold1 publication.
b 12 families had overlap with the AlphaFold2 training set and were excluded.

Table S5. CATH assignment of BFD models. The number of models assigned to each CATH class and not assigned to a CATH class are shown, along with the number of superfamilies present in each class in CATH.

Class	No. Superfamilies	% of Total	No. Models Assigned	% of Total
	945	39.46	11,891	35.44
2	451	18.83	3,495	10.42
3	962	40.17	8,035	23.95
4	4	0.17	6	0.02
6	32	1.34	68	0.20
None			10,053	29.97
Total	2,395		33,548	