Table 1. Benchmark set for multi-functional enzymes.

Family ID	Family name	Nb of modeled sequences	ECs	PDB “holo ”	Bound ligand equivalent to natural substrate/product
PF00108	Thiolase_N	2806	2.3.1.9 2.3.1.16 2.3.1.176	1NL7	Coenzyme A
PF00128	Alpha-amylase	17	2.4.1.4 2.4.1.7 3.2.1.10 3.2.1.20 3.2.1.70 3.2.1.98 3.2.1.93 3.2.1.141 5.4.99.16 5.4.99.15	2D3N	Glucose
PF00135	COesterase	386	3.1.1.1 3.1.1.3 3.1.1.7 3.1.1.8 3.1.1.13 3.1.1.59	1P0M	Choline ion
PF00215	OMPdecase	83	4.1.1.23 4.1.1.85	2CZE	Uridine-5’-monophosphate
PF00278	Orn_DAP_Arg_deC	343	4.1.1.17 4.1.1.18 4.1.1.19 4.1.1.20	1TWI	Lysine
PF00293	NUDIX	2670	2.7.7.1 3.6.1.13 3.6.1.17 3.6.1.52 3.6.1.52 5.3.3.2	2DSC	Adenosine-5-diphosphoribose
PF00348	Polyprenyl_synt	499	2.5.1.10 2.5.1.29	2F8Z	Zoledronic acid, 3-methylbut-3-enyl trihydrogen diphosphate
PF00351	Biopterin_H	268	1.14.16.1 1.14.16.2 1.14.16.4	1MMK	5,6,7,8-tetrahydrobiopterin, beta(2-thienyl)alanine
PF00579	tRNA-synt_1b	251	6.1.1.1 6.1.1.2	1WQ4	Tyrosine
PF00583	Acetyltransf_1	3381	2.3.1.1 2.3.1.4 2.3.1.48 2.3.1.57 2.3.1.59 2.3.1.82 2.3.1.87 2.3.1.88 2.3.1.128	1TIQ	Coenzyme A

PF00590	TP_methylase	315	2.1.1.98 2.1.1.107 2.1.1.130 2.1.1.131 2.1.1.132 2.1.1.133 2.1.1.152 2.1.1.151 4.2.1.75 4.99.1.4	1S4D	S-adenosyl-L-homocysteine
PF00755	Carn_acyltransf	228	2.3.1.6 2.3.1.7 2.3.1.21 2.3.1.137	1NDI	Coenzyme A
PF00871	Acetate_kinase	189	2.7.2.1 2.7.2.7 2.7.2.15	1TUY	Adenosine-5’-diphosphate
PF00962	A_deaminase	14	3.5.4.4 3.5.4.6	1NDZ	1-((1r)-1-(hydroxymethyl)-3-(6-((3-(1-methyl- 1h-benzimidazol-2-yl)propanoyl)amino)-1h- indol-1-yl)propyl)-1h-imidazole-4-carboxamide
PF01048	PNP_UDP_1	69	2.4.2.1 2.4.2.3 2.4.2.28 3.2.2.4 3.2.2.9	1PK7	Adenosine
PF01112	Asparaginase_2	417	3.5.1.1 3.5.1.26	1SEO	Aspartic acid
PF01135	PCMT	655	2.1.1.77 2.1.1.36	1R18	S-adenosyl-L-homocysteine
PF01202	SKI	174	2.7.4.3 2.7.1.12 2.7.4.14 2.7.1.71 4.2.3.4	1WE2	Adenosine-5’-diphosphate
PF01234	NNMT_PNMT_TEMT	54	2.1.1.1 2.1.1.28 2.1.1.49	2AN4	S-adenosyl-L-homocysteine
PF01467	CTP_transf_2	548	2.7.7.1 2.7.7.3 2.7.7.14 2.7.7.15 2.7.7.18 2.7.7.39	1N1D	[Cytidine-5’-phosphate] glycerylphosphoric acid ester

PF01712	dNK	229	1.6.99.3 2.7.1.21 2.7.1.74 2.7.1.76 2.7.1.113 2.7.1.145	2A2Z	Uridine-5’-diphosphate, 2’-deoxycytidine
PF02274	Amidinotransf	294	2.1.4.1 3.5.3.6 3.5.3.18	2A9G	Arginine
PF03171	2OG-FeII_Oxy	269	1.14.11.2 1.14.11.4 1.14.11.7 1.14.11.9 1.14.11.11 1.14.11.13 1.14.11.19 1.14.11.20 1.14.11.23 1.14.11.26 1.14.17.4 1.14.20.1 1.21.3.1	2FDJ	Succinic acid
PF03414	Glyco_transf_6	82	2.4.1.87 2.4.1.40	1LZJ	Alpha-L-Fucose

Table 2. Benchmark set for uni-functional enzymes

Family ID	Family name	Nb of modeled sequences	ECs	PDB “holo ”	Bound ligand equivalent to natural substrate/product
PF00303	Thymidylat_synt	84	2.1.1.45	2G8O	2’-deoxyuridine 5’-monophosphate, 10-propargyl-5,8-dideazafolic acid
PF00693	Herpes_TK	39	2.7.1.21	1VTK	Adenosine-5’-diphosphate, thymidine-5’-phosphate
PF00925	GTP_cyclohydro2	142	3.5.4.25	2BZ0	Phosphomethylphosphonic acid guanylate ester
PF01014	Uricase	127	1.7.3.3	2FXL	1-(2,5-dioxo-2,5-dihydro-1h-imidazol-4-yl)urea
PF01227	GTP_cyclohydro2	144	3.5.4.16	1A8R	Guanosine-5’-triphosphate cyclohydroI
PF01293	PEPCK_ATP	452	4.1.1.49	1YTM	Adenosine-5’-triphosphate, oxalic acid
PF01583	APS_kinase	63	2.7.1.25	1M7G	Adenosine-5’-phosphosulfate, adenosine-5’-di-phosphate-2’,3’-vanadate
PF01656	CbiA	348	6.3.3.3	1A82	Adenosine-5’-triphosphate, 7,8-diamino-nonanoic acid
PF01702	TGT	147	2.4.2.29	1Q2S	9-deazaguanine
PF01747	ATP-sulfurylase	147	2.7.7.4	1G8H	Adenosine-5’-phosphosulfate, pyrophosphate 2-sulfurylase
PF02110	HK	399	2.7.1.50	1ESQ	Adenosine-5’-triphosphate, 4-methyl-5-hydroxy-ethylthiazole phosphate
PF02223	Thymidylate_kin	1061	2.7.4.9	1E9E	Adenosine-5’-diphosphate, thymidine-5’-phosphate
PF02277	DBI_PRT	512	2.4.2.21	1L5L	7-alpha-d-ribofuranosyl-purine-5’-phosphate, nicotinic acid
PF02353	CMAS	497	2.1.1.79	1KPI	S-adenosyl-Ll-homocysteine
PF02569	Pantoate_ligase	841	6.3.2.1	2A86	Adenosine monophosphate, beta-alanine
PF02898	NO_synthase	130	1.14.13.39	1Q2O	L-n(omega)-nitroarginine-2,-4-L-diaminobutyric amide
PF02901	PFL	380	2.3.1.54	1MZO	Pyruvic acid
PF03332	PMM	139	5.4.2.8	2FUE	Alpha-d-mannose 1-phosphate

Table 3. Comparison of ASMC and SDPsite methods. The average distance between predicted functional residues (dist_CPs+SDPs for ASMC method and dist_λ=1.0 for SDPsite method) and the bound substrate is indicated in the first two columns. In SDPsite method (Kalinina *et al.*, 2009), λ=1.0 means that SDPs and CPs have the same weight in distance calculation to the ligand(s).

Family ID	ASMC dist_CPs+SDPs	SDPsite dist_λ=1.0	Comparison dist_Δ
PF00108	5.92	14.51	+8.59
PF00128	7.85	7.39	-0.45
PF00135	4.83	26.45	+21.62
PF00215	3.42	6.78	+3.36
PF00278	6.07	24.85	+18.78
PF00293	0.00	5.56	nan
PF00303	3.27	15.19	+11.92
PF00348	5.01	6.11	+1.10
PF00351	4.63	10.51	+5.88
PF00579	4.88	6.92	+2.04
PF00583	4.82	4.67	-0.14
PF00590	4.42	6.40	+1.98
PF00693	3.12	4.67	+1.55
PF00755	3.86	7.40	+3.54
PF00871	7.24	9.32	+2.08
PF00925	3.70	11.55	+7.85
PF00962	4.24	8.81	+4.57
PF01014	5.25	6.75	+1.50
PF01048	3.77	8.28	+4.51
PF01112	3.77	28.32	+24.55
PF01135	0.00	4.89	nan
PF01202	5.53	3.71	-1.81
PF01227	4.20	4.94	+0.74
PF01234	4.45	7.44	+2.99
PF01293	4.46	10.83	+6.37
PF01467	0.00	7.69	nan
PF01583	3.67	13.94	+10.27
PF01656	3.65	4.88	+1.23
PF01702	7.84	6.26	-1.57
PF01712	4.13	6.37	+2.24
PF01747	5.42	5.90	0.48
PF02110	3.73	7.18	+3.45
PF02223	3.91	4.38	0.47
PF02274	5.25	6.04	0.79
PF02277	6.95	8.82	+1.87
PF02353	3.75	5.99	+2.24
PF02569	6.51	8.67	+2.16
PF02898	6.50	12.26	+5.76
PF02901	3.47	6.81	+3.34
PF03171	5.18	6.29	+1.11
PF03332	7.64	11.84	+4.20
PF03414	5.77	8.78	+3.01

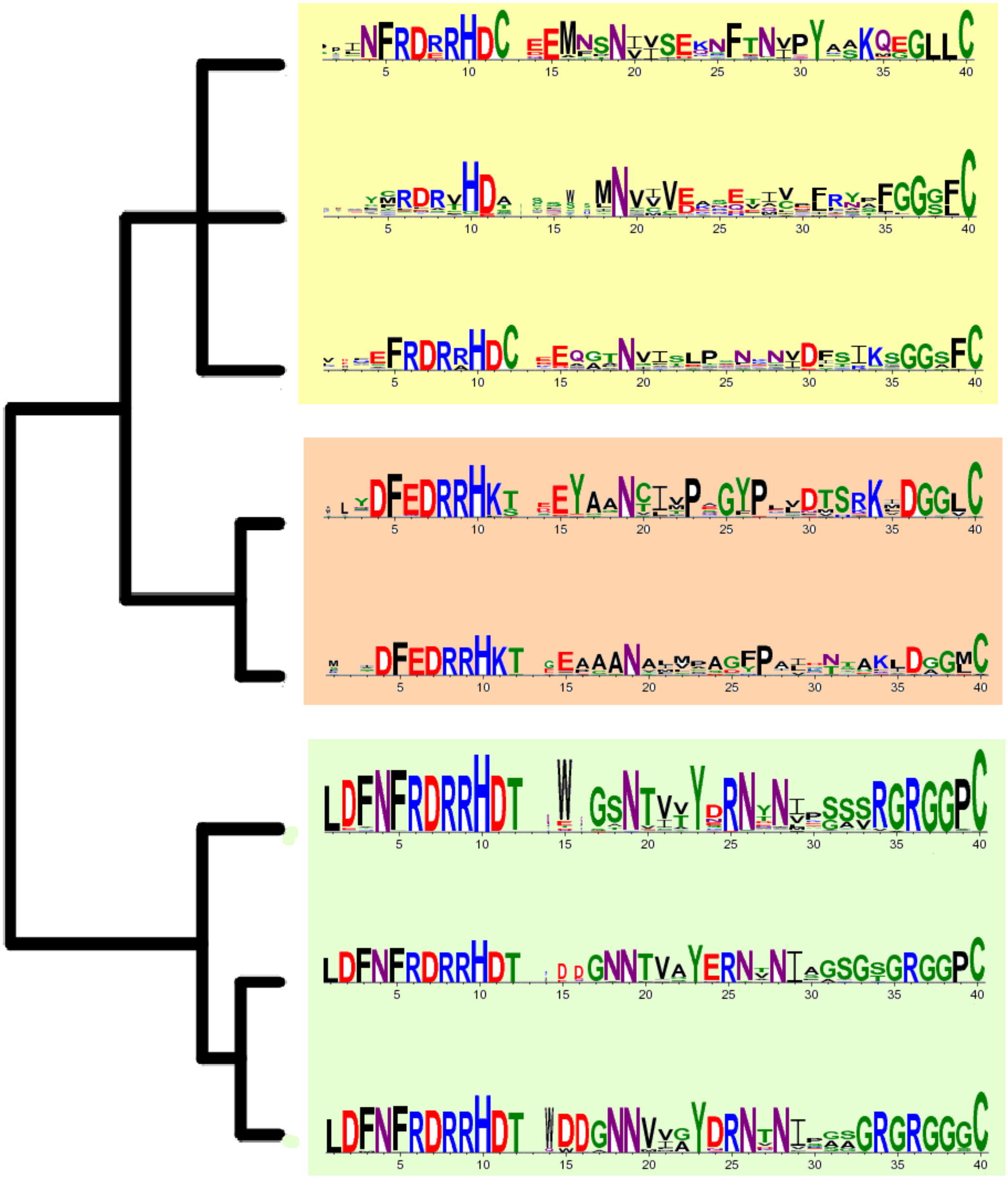

Fig. 1. ASMC tree of Family PF02274 (Amidinotransferase). This family contains glycine (EC:2.1.4.1) and inosamine (EC:2.1.4.2) amidinotransferases, enzymes involved in creatine and streptomycin biosynthesis respectively. This family also includes arginine deiminases, EC:3.5.3.6. These enzymes catalyse the reaction: arginine + H2O to citrulline + NH3. ASMC has divided the superfamily in 3 main groups: ”yellow” group (that includes three clusters), ”orange group” (with two clusters) and ”green” group (with three clusters). ”Yellow”, ”orange” and ”green” groups contain respectively sequences that are associated to Enzyme Commission number EC:2.1.4.1, EC:2.1.4.2, EC:3.5.3.

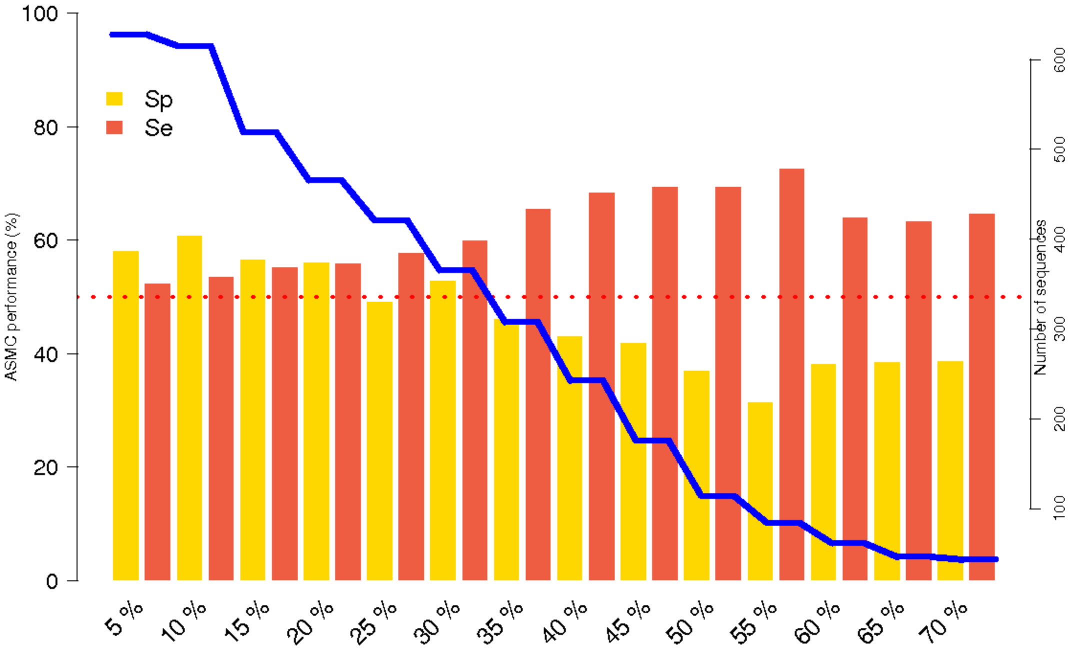

Fig. 2. Influence of the % of identity between targets and templates on ASMC performance. Evolution of Sp and Se in function of the threshold limit in term of % of identity between target and templates sequences. For instance, we performed ASMC on set of sequences with at least 45% of identity with one of the templates structures. Values indicated here are the mean over the 42 benchmark families. The blue line indicate the average number of sequences per family for each threshold range. The gain in sensibility with the increasing of sequences identities is limited by the drop of available sequences (less than 50 for high level of identity).

Table 4. Comparison of ASMC performance using a similar procedure but without any information on protein structures. This procedure is based on a multiple sequence alignment (called here MSeqA, see Material&Methods, section 2.4) For families PF00693, PF01135 and PF01234, Sp and Se could not be calculated (indicated with ”nan” in the table) as no cluster could emerged.

Family ID	Sp - ASMC	Se - ASMC	Sp - MSeqA	Se - MSeqA
PF00108	0	1	0	1
PF00128	0.69	0.5	0.56	0
PF00135	0.91	0.33	0.78	0.11
PF00215	1	0.4	1	0.4
PF00278	0.5	1	0.67	0.33
PF00293	1	0	0.77	0.46
PF00303	0.67	0.55	0.67	0.82
PF00348	0.43	0.86	0.86	0.11
PF00351	0.77	0.69	0.55	0.83
PF00579	0.77	0.4	0.43	0.47
PF00583	0.91	0.2	0.94	0
PF00590	0.6	0.41	0.1	1
PF00693-1	0.00	0.58	nan	nan
PF00693-2	1.00	0.00	0.25	0.62
PF00755	1	0.07	0.89	0.07
PF00871-1	0.52	0.86	0.32	0.90
PF00871-2	1.00	0.30	0.83	0.6
PF00925	0.5	0.9	0.5	0.95
PF00962	0.89	0.15	0.27	0.8
PF01014	0.67	1	0.33	0.67
PF01048	0.96	0.46	0.89	0.62
PF01112	0.5	0.83	0.5	0.83
PF01135	1	0	nan	nan
PF01202	0.89	0.19	0.85	0.44
PF01227	0	0.8	0.67	0.5
PF01234	0.76	0.8	nan	nan
PF01293	0	1	0.05	0
PF01467	1	0	0.6	0.54
PF01583-1	0.89	0.17	0.20	0.91
PF01583-2	1	0.45	0.4	0.55
PF01656	0.75	0.82	0.33	0.82
PF01702	0.08	0.92	0.04	1
PF01712-1	0.14	0.94	0.20	0.62
PF01712-2	0.00	0.46	0	0.54
PF01747	0.64	0.35	0.27	0.85
PF02110	0	0.71	0	0.68
PF02223-1	0.12	0.67	0.25	0.8
PF02223-2	0.00	0.93	0.50	0.27
PF02274	0.52	0.93	0.84	0.8
PF02277	0.25	1	0.12	0.88
PF02353	0.33	0.97	0.33	0.9
PF02569	0.4	0.78	0.4	0.74
PF02898	0.23	0.87	0.9	0.07
PF02901	0	1	0	0.8
PF03171	0.96	0.18	0.12	0.88
PF03332	0.18	0.62	0.05	1
PF03414	0.6	0.33	0.4	0.44