



CASP 9

EVALUATION OF RESIDUE-RESIDUE CONTACT PREDICTION

Bohdan Monastyrskyy

Andriy Kryshafovych

Krzysztof Fidelis



Bohdan Monastyrskyy



Evaluated domains:

- Analysis performed on FM and TBM/FM domains only
- 1 domain (T0537_D2) not considered due to its short length - 31 residues
- 28 domains total



Participating groups:

- 28
- 25 submitted predictions for the evaluated domains



Format of predictions:

- Residue pairs in contact
- Probability scores for contacts

Probability scores were used to rank and select predicted contacts



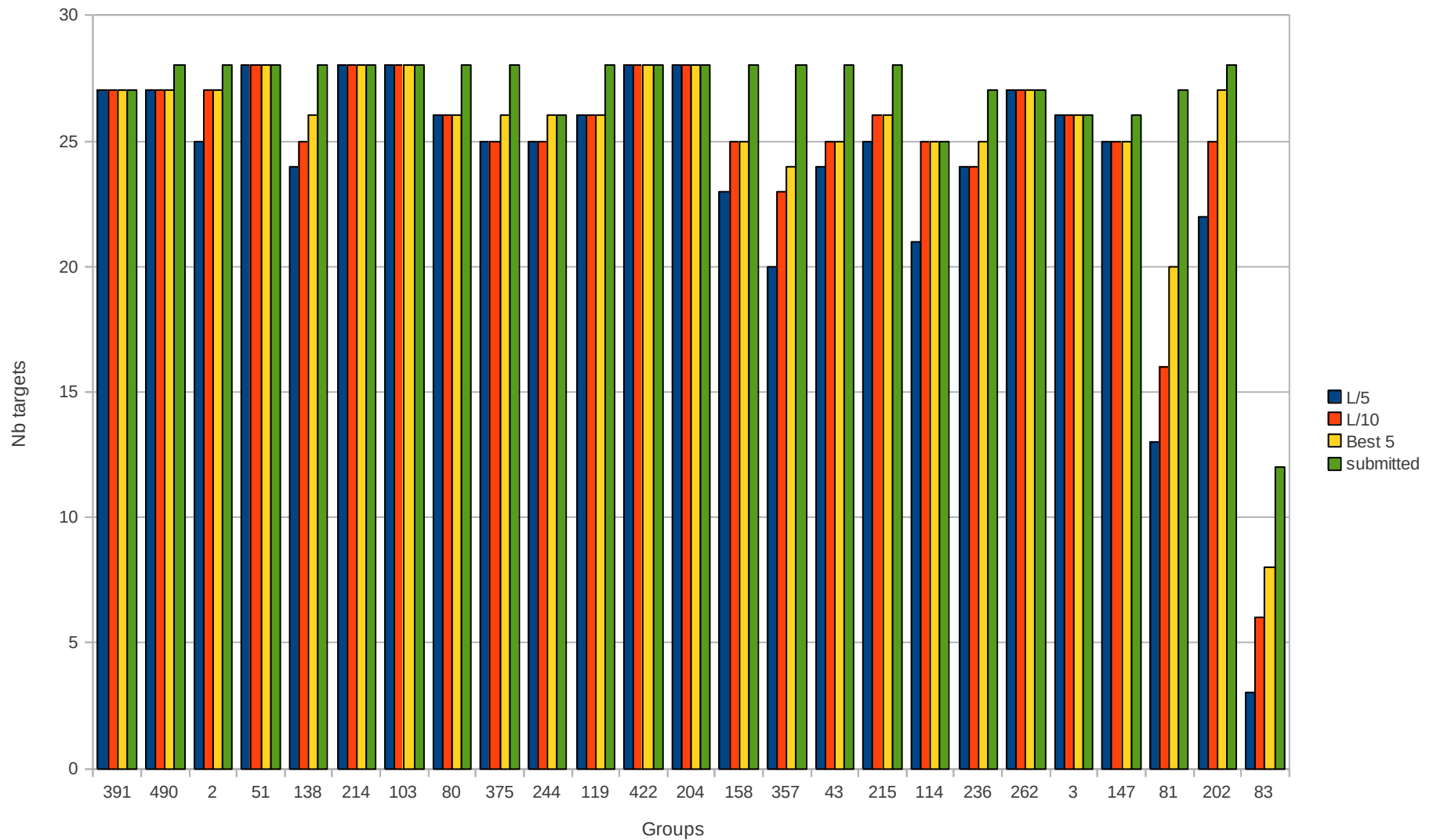
Evaluation:

- Evaluation performed on predictions ranked as #1
- Only long range contacts considered (separation along the sequence ≥ 24 residues)
- For each target domain of length L , 3 sets of top ranked contacts were considered, comprising $L/5$, $L/10$, and just 5 predictions

Not all groups submitted sufficient number of contacts to be evaluated



Numbers of predicted targets per group





Evaluation scores:

As in previous CASPs two scores for evaluation of contact predictions were used: Acc and Xd

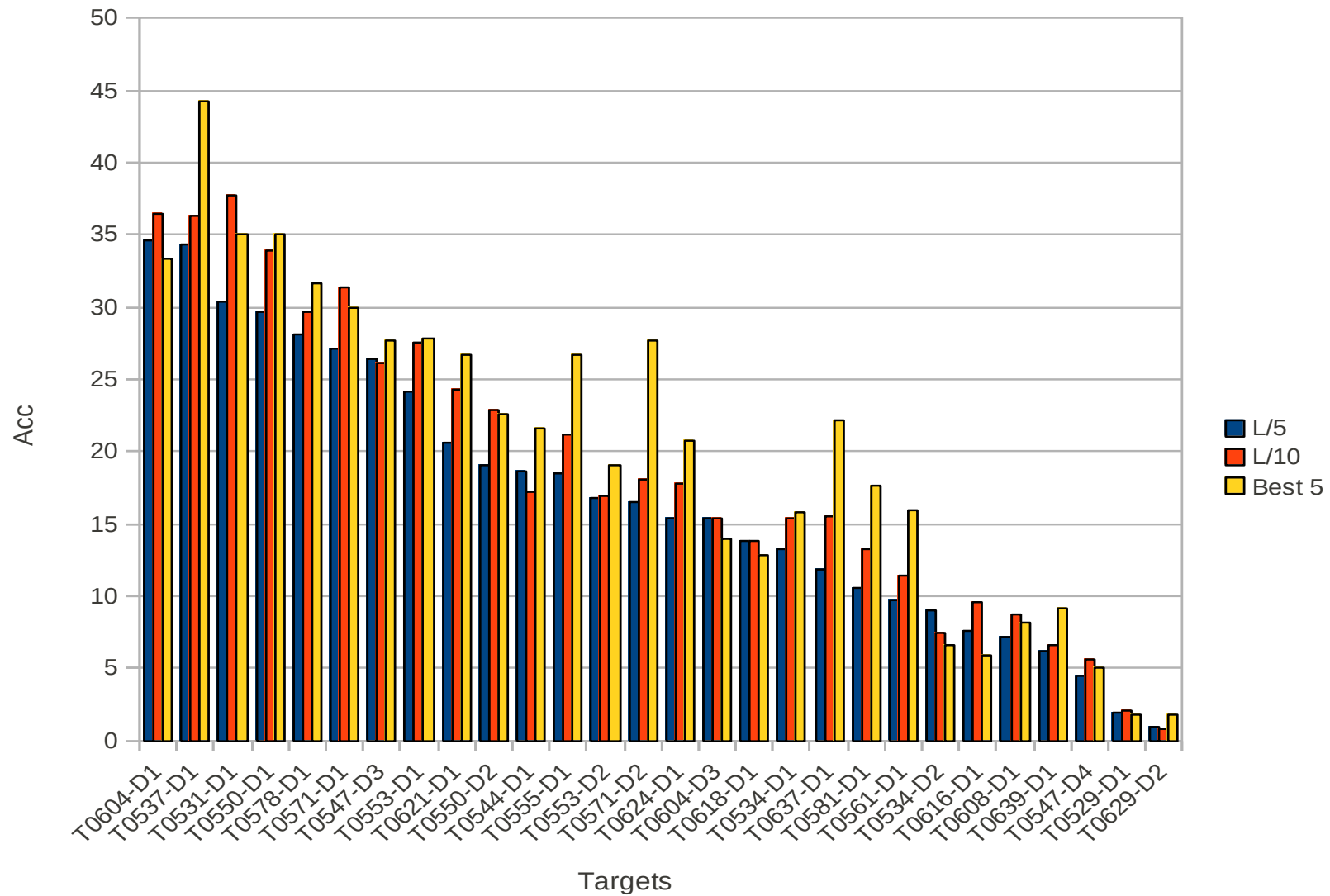
$$Acc = \frac{TP}{TP + FP}; \quad Xd = \sum_{i=1}^{15} \frac{(Pp_i - Pa_i)}{i},$$

Pp_i – percentage of predicted pairs
with distance $d \in [4(i-1), 4i]$;

Pa_i – percentage of all pairs of target
with distance $d \in [4(i-1), 4i]$;

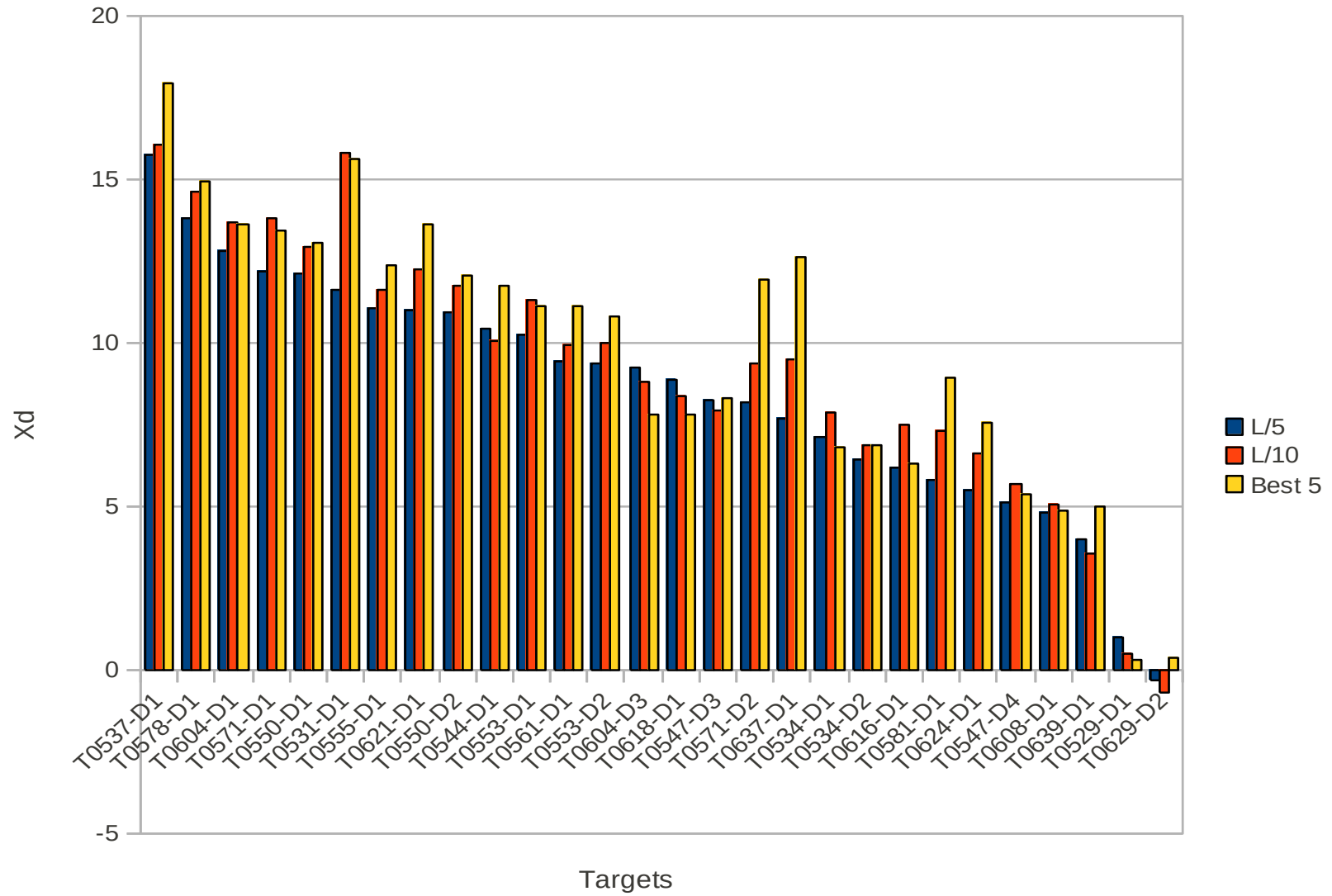


Acc per target



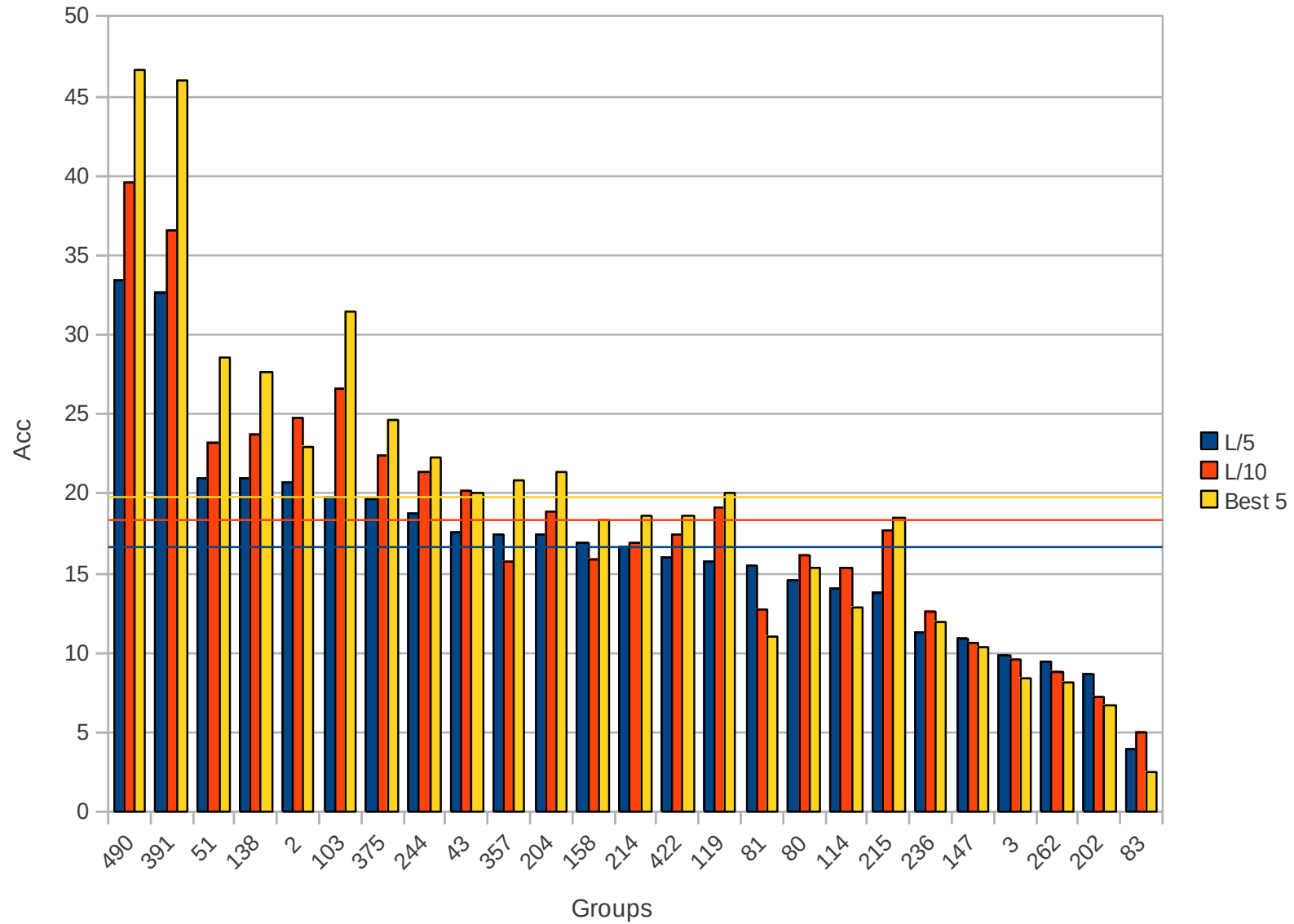


X_d per target



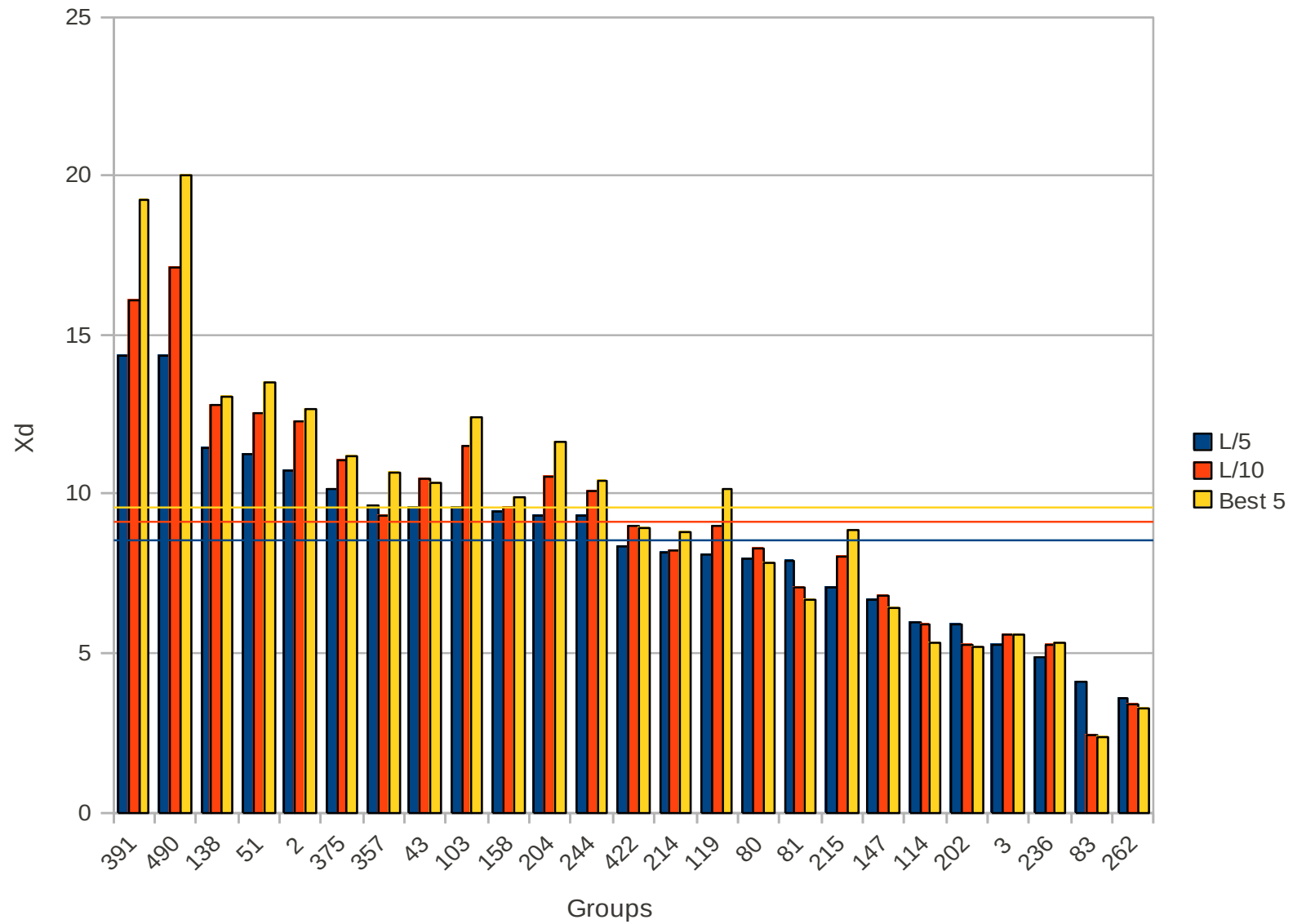


Acc per group





X_d per group



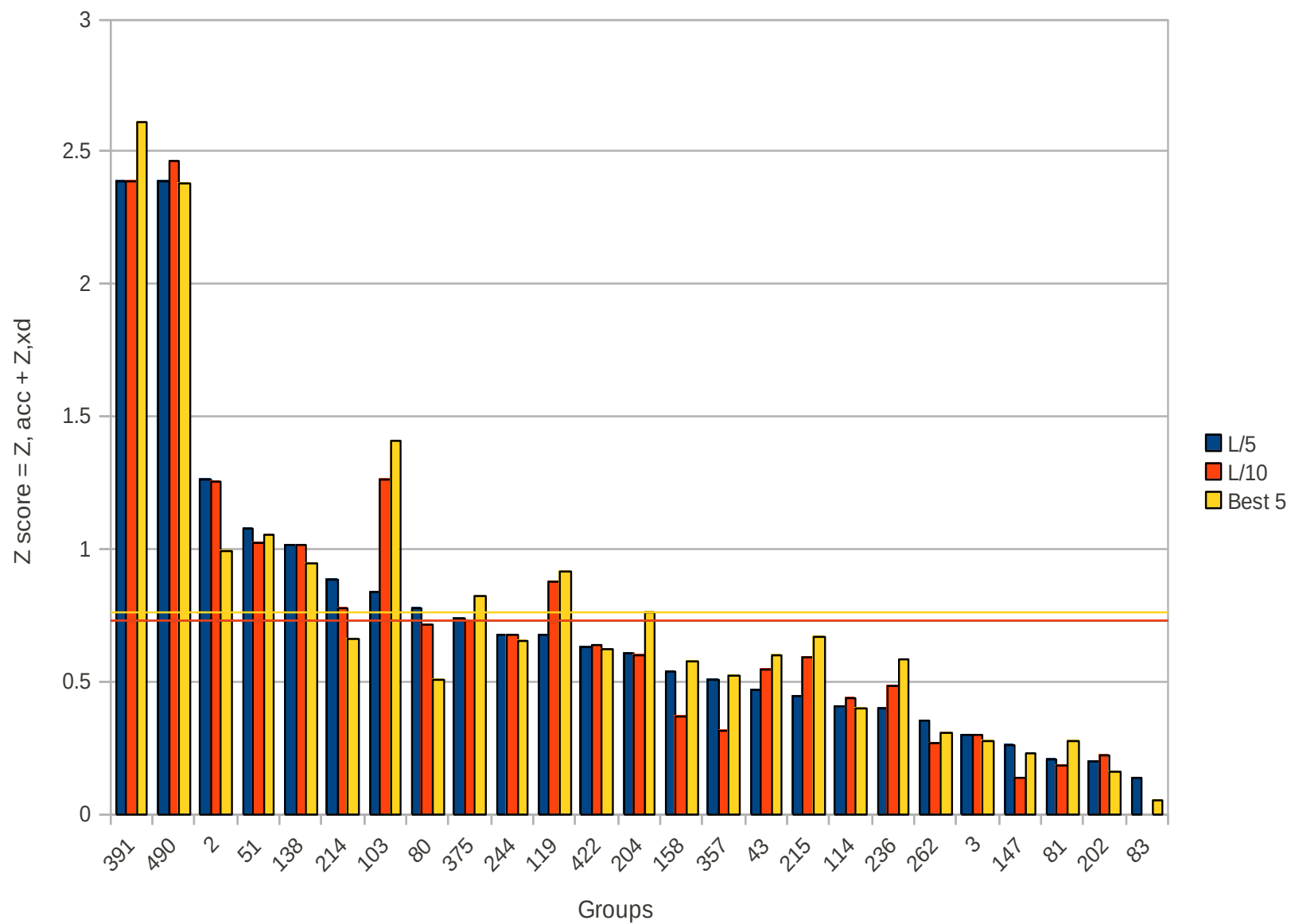


The groups were ranked according to the cumulative Z score (removing outliers):

$$Z_{\text{total}} = Z_{\text{Acc}} + Z_{\text{Xd}}$$



$Z, total$





T test for groups (Acc score, L/5)

\	391	490	2	51	138	214	103	80	375	244	119	422
391	NA	0.449	0.006	0.006	0.004	0.002	0.009	0.001	0.004	0.001	0.001	0.001
490	0.449	NA	0.006	0.005	0.005	0.001	0.009	0	0.004	0.001	0.001	0.001
2	0.006	0.006	NA	0.468	0.443	0.23	0.478	0.026	0.438	0.313	0.05	0.16
51	0.006	0.005	0.468	NA	0.261	0.072	0.315	0.055	0.189	0.094	0.085	0.004
138	0.004	0.005	0.443	0.261	NA	0.258	0.469	0.112	0.44	0.299	0.161	0.08
214	0.002	0.001	0.23	0.072	0.258	NA	0.171	0.285	0.282	0.398	0.371	0.415
103	0.009	0.009	0.478	0.315	0.469	0.171	NA	0.112	0.377	0.23	0.151	0.042
80	0.001	0	0.026	0.055	0.112	0.285	0.112	NA	0.131	0.223	0.161	0.355
375	0.004	0.004	0.438	0.189	0.44	0.282	0.377	0.131	NA	0.322	0.187	0.055
244	0.001	0.001	0.313	0.094	0.299	0.398	0.23	0.223	0.322	NA	0.293	0.164
119	0.001	0.001	0.05	0.085	0.161	0.371	0.151	0.161	0.187	0.293	NA	0.451
422	0.001	0.001	0.16	0.004	0.08	0.415	0.042	0.355	0.055	0.164	0.451	NA



T test for groups at (Xd score, L/5)

\	391	490	2	51	138	214	103	80	375	244	119	422
391	NA	0.27	0.004	0.01	0.005	0	0.001	0.002	0.002	0	0.001	0.001
490	0.27	NA	0.01	0.017	0.007	0	0.002	0.001	0.003	0	0.001	0.001
2	0.004	0.01	NA	0.323	0.477	0.09	0.284	0.058	0.425	0.173	0.031	0.108
51	0.01	0.017	0.323	NA	0.159	0.006	0.036	0.018	0.071	0.017	0.012	0.001
138	0.005	0.007	0.477	0.159	NA	0.043	0.128	0.014	0.224	0.037	0.018	0.011
214	0	0	0.09	0.006	0.043	NA	0.157	0.386	0.175	0.354	0.408	0.439
103	0.001	0.002	0.284	0.036	0.128	0.157	NA	0.159	0.363	0.251	0.141	0.066
80	0.002	0.001	0.058	0.018	0.014	0.386	0.159	NA	0.099	0.245	0.437	0.366
375	0.002	0.003	0.425	0.071	0.224	0.175	0.363	0.099	NA	0.182	0.087	0.041
244	0	0	0.173	0.017	0.037	0.354	0.251	0.245	0.182	NA	0.249	0.282
119	0.001	0.001	0.031	0.012	0.018	0.408	0.141	0.437	0.087	0.249	NA	0.383
422	0.001	0.001	0.108	0.001	0.011	0.439	0.066	0.366	0.041	0.282	0.383	NA



Top12 groups:

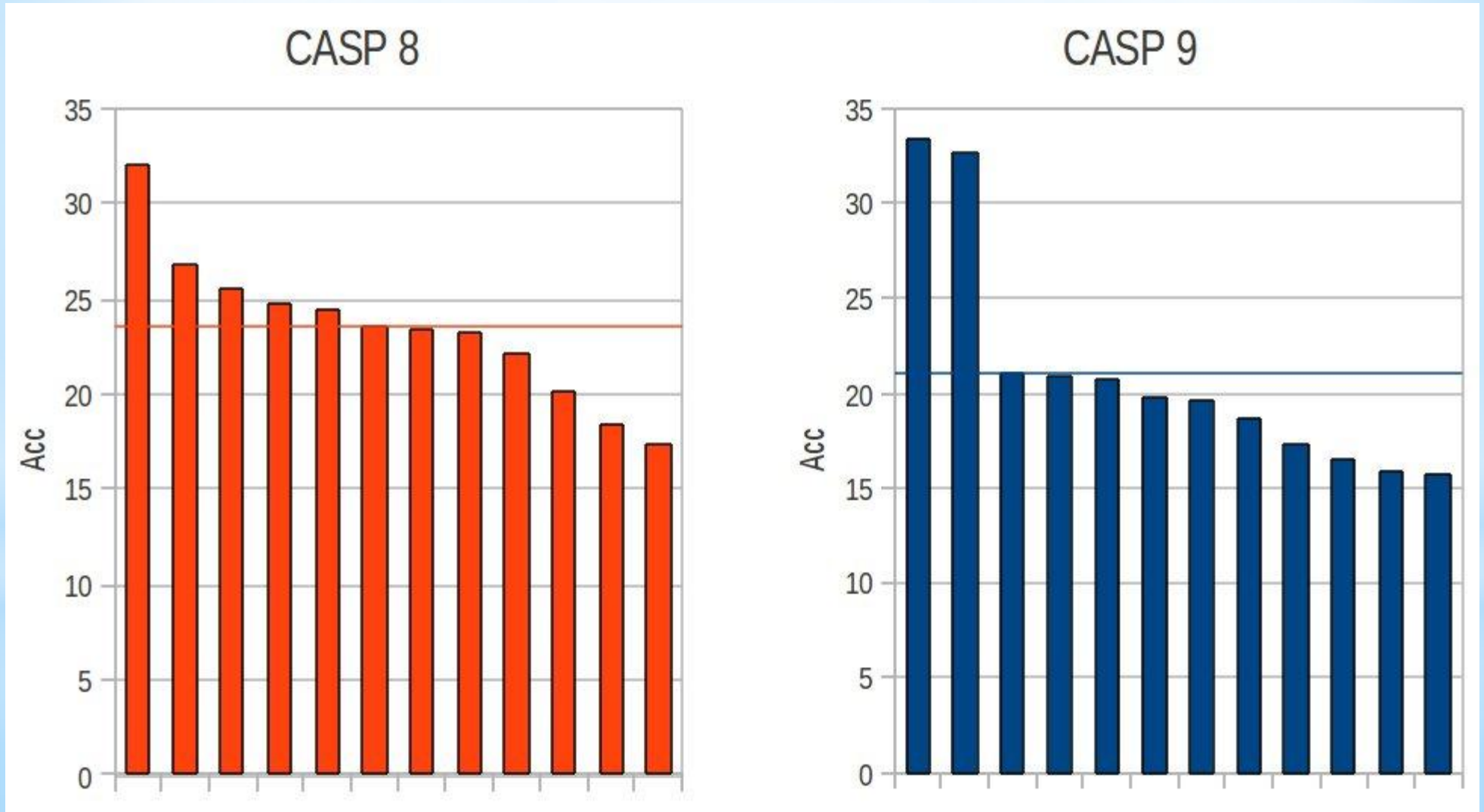
	Groups	L/5	
		Nb Targets	Z, total
SMEG-CCP	391	27	2.3912888889
MULTICOM	490	27	2.3884444444
MULTICOM-CLUSTER	2	25	1.2584
Infobiotics	51	28	1.0733464286
ProC_S3	138	24	1.0118708333
Distill	214	28	0.8808714286
SAM-T08-server	103	28	0.8399928571
MULTICOM-CONSTRUCT	80	26	0.7760653846
ProC_S1	375	25	0.740252
SAM-T06-server	244	25	0.678252
MULTICOM-REFINE	119	26	0.6738884615
PSICON	422	28	0.6277285714



SMEG-CCP abstract:

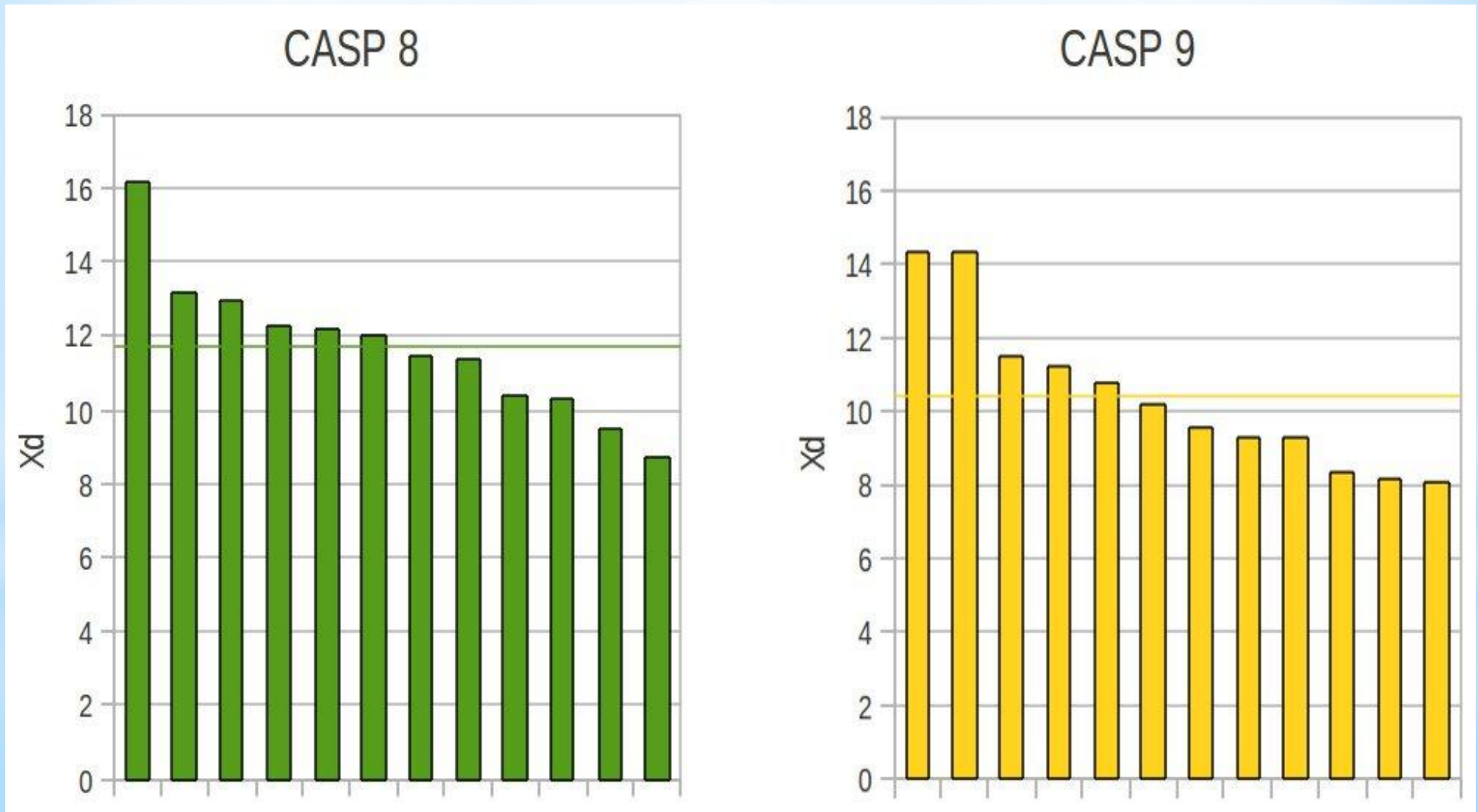
“For each target, the server predictions marked as ‘model 1’ were converted to contact maps using the CASP contact definition...”

Comparison with CASP8 (12 best groups, Acc)





Comparison with CASP8 (12 best groups, Xd)





Head to head comparison of groups

% of common targets for which group I beats group II (Acc, L/5)

		group II											
		391	490	2	51	138	214	103	80	375	244	119	422
group I	391	NA	19.2%	68.0%	70.4%	65.2%	74.1%	51.9%	80.0%	66.7%	70.8%	80.0%	74.1%
	490	26.9%	NA	64.0%	63.0%	62.5%	77.8%	63.0%	80.8%	64.0%	68.0%	80.8%	70.4%
	2	24.0%	32.0%	NA	48.0%	47.8%	48.0%	44.0%	72.0%	33.3%	45.8%	64.0%	48.0%
	51	22.2%	25.9%	40.0%	NA	45.8%	57.1%	50.0%	65.4%	52.0%	56.0%	65.4%	60.7%
	138	30.4%	29.2%	39.1%	33.3%	NA	58.3%	58.3%	66.7%	62.5%	45.8%	58.3%	62.5%
	214	14.8%	18.5%	48.0%	28.6%	33.3%	NA	39.3%	57.7%	48.0%	40.0%	57.7%	46.4%
	103	33.3%	29.6%	40.0%	35.7%	41.7%	57.1%	NA	57.7%	40.0%	44.0%	57.7%	50.0%
	80	12.0%	11.5%	24.0%	26.9%	29.2%	30.8%	38.5%	NA	48.0%	28.0%	26.9%	42.3%
	375	16.7%	20.0%	54.2%	36.0%	37.5%	52.0%	32.0%	52.0%	NA	36.0%	52.0%	48.0%
	244	16.7%	20.0%	45.8%	36.0%	41.7%	48.0%	36.0%	64.0%	40.0%	NA	64.0%	60.0%
	119	12.0%	15.4%	28.0%	19.2%	33.3%	34.6%	34.6%	46.2%	44.0%	24.0%	NA	50.0%
	422	22.2%	22.2%	28.0%	17.9%	25.0%	46.4%	25.0%	42.3%	32.0%	32.0%	46.2%	NA