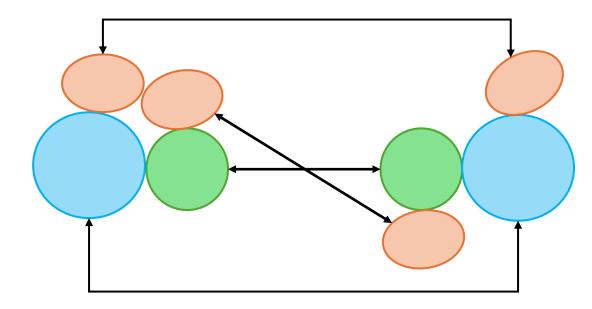
# Analysis and evaluation of CASP16 protein oligomer predictions

Rongqing Yuan, Jing Zhang, Andriy Kryshtafovych, Gabriel Studer, Nick V. Grishin, and Qian Cong

**University of Texas Southwestern Medical Center** 

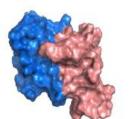
### The classic oligomer scoring routine in CASP



#### Assembly/Global:

IDDT

TM score







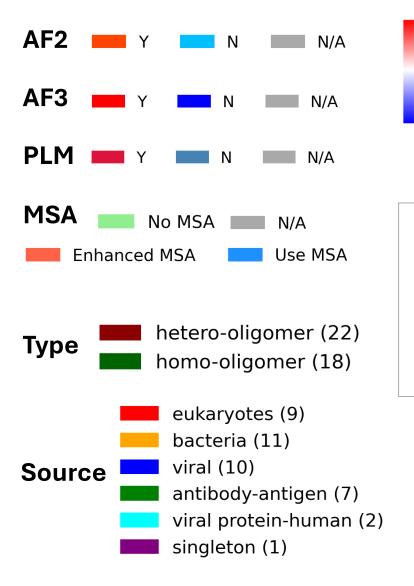
DockQ: interface size-weighted average

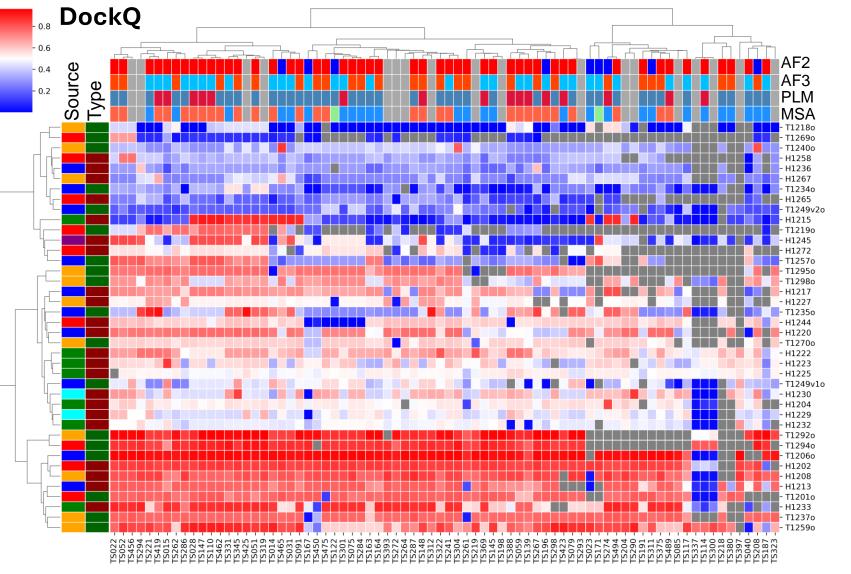
$$IPS(M,T) = J_C(M,T) = \frac{|M_{i-Res} \cap T_{i-Res}|}{|M_{i-Res} \cup T_{i-Res}|}$$

$$ICS(M,T) = 2 \cdot \frac{P(M_{cnt}, T_{cnt}) \times R(M_{cnt}, T_{cnt})}{P(M_{cnt}, T_{cnt}) + R(M_{cnt}, T_{cnt})}$$

$$QS\text{-best}(M,T) = \frac{|M_{\text{cnt}} \cap T_{\text{cnt}}|}{\max(M_{\text{cnt}},T_{\text{cnt}})}$$

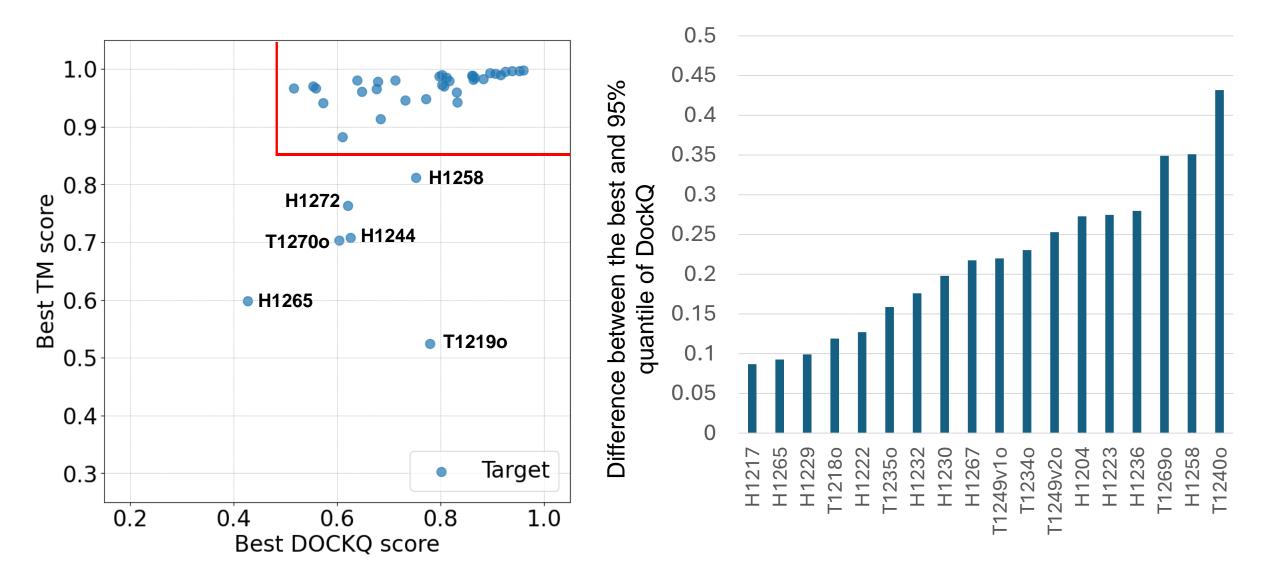
#### An overview of oligomer targets and groups in CASP16



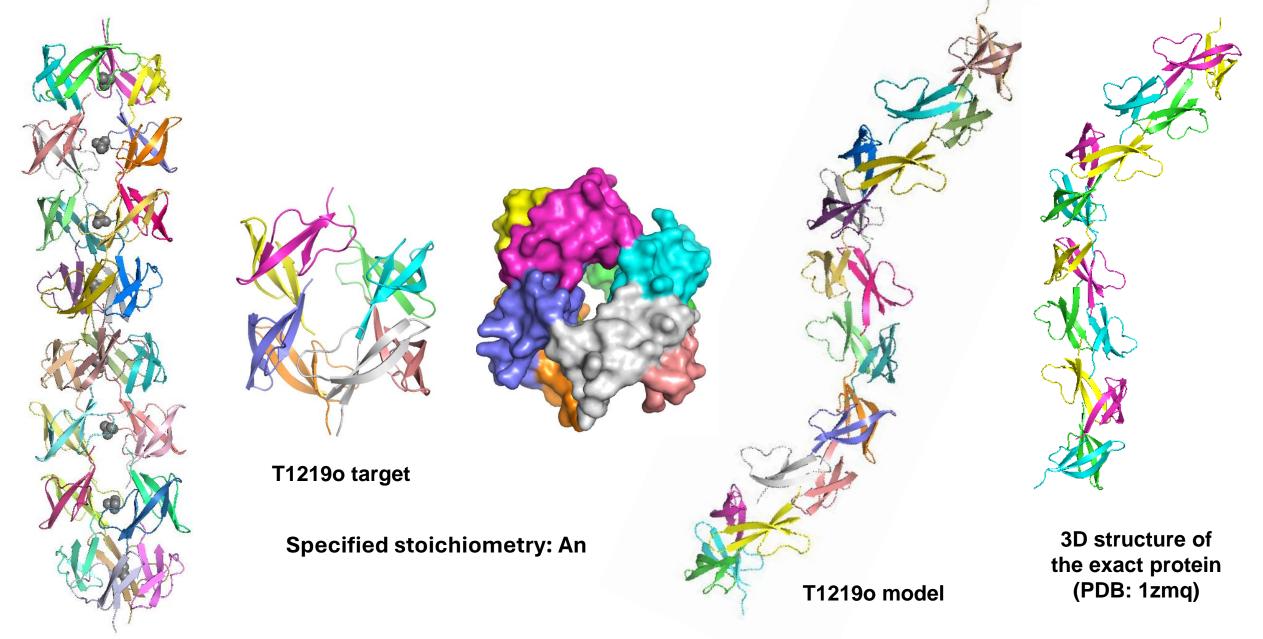


# What targets are still challenging for the community?

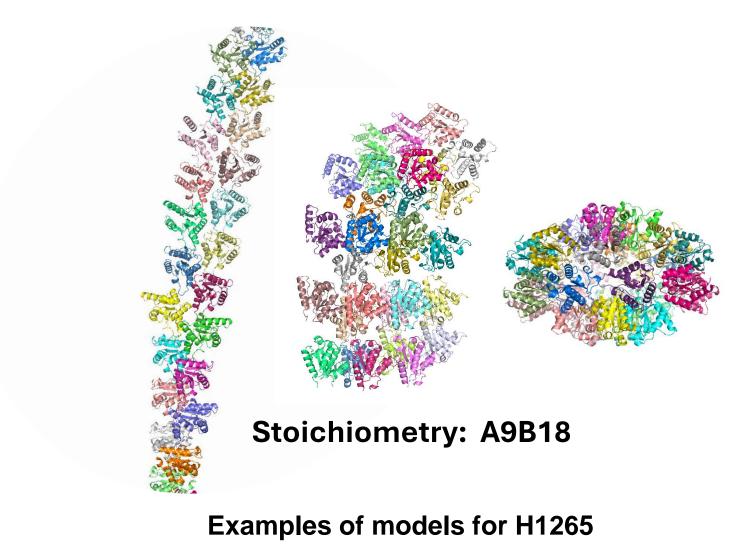
#### What targets are still challenging for the community?

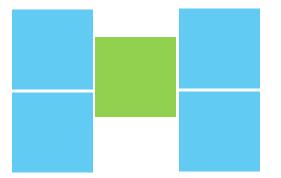


#### Challenge 1: filament maintained by weak interactions



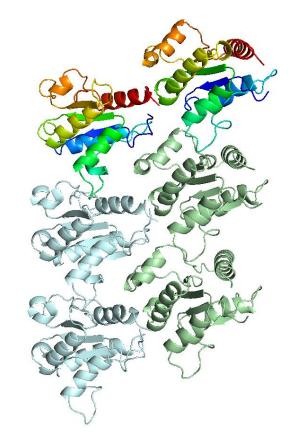
#### Challenge 2: multiple possible interfaces between proteins and unusual shape of protein complex





H1265 target

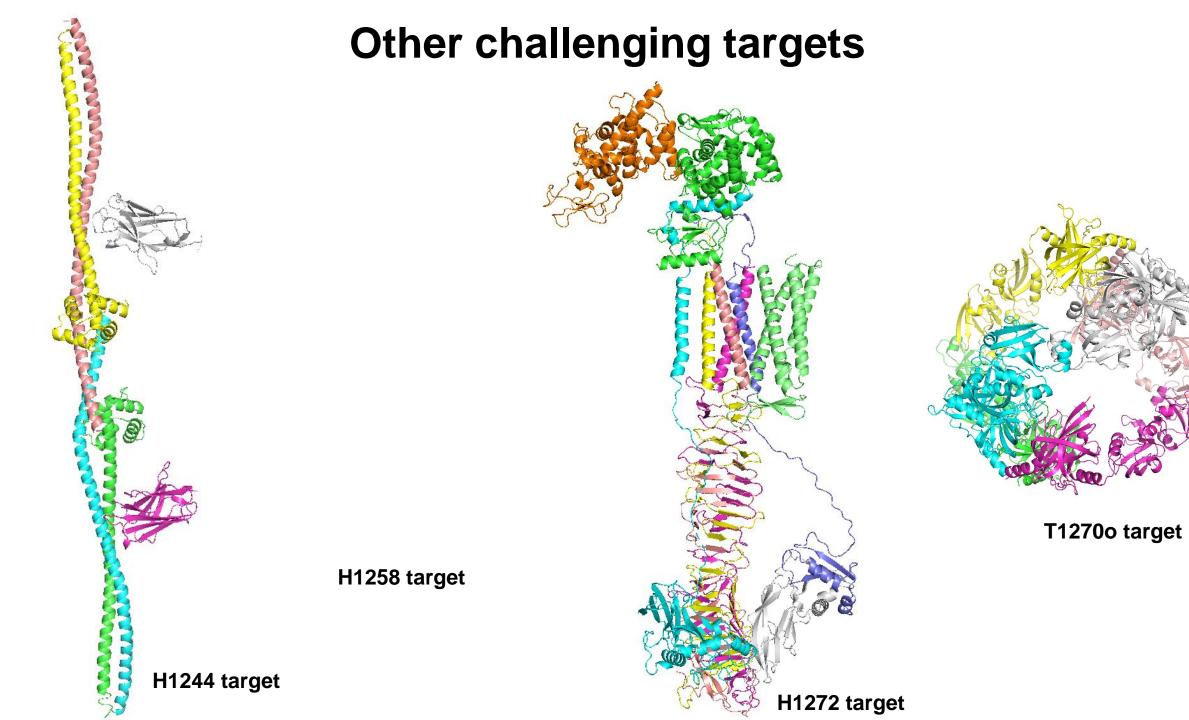
#### Challenge 2: multiple possible interfaces between proteins and unusual shape of protein complex



The shape is correct, but the interfaces are incorrect

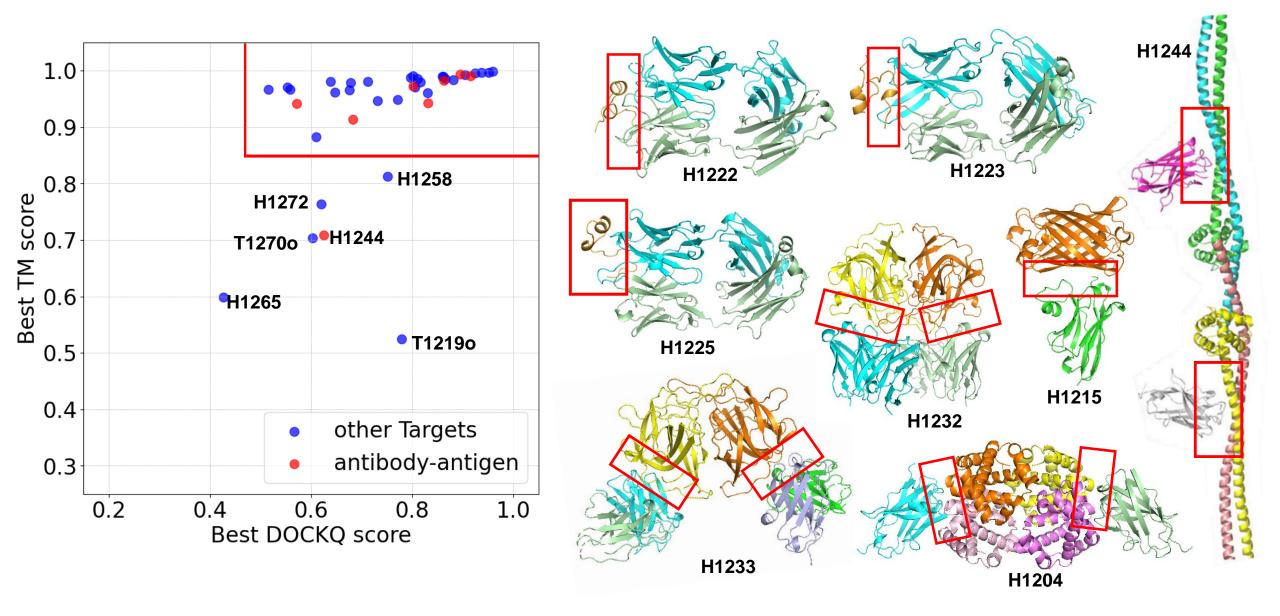
H1265 target

A winning model from the Kihara group

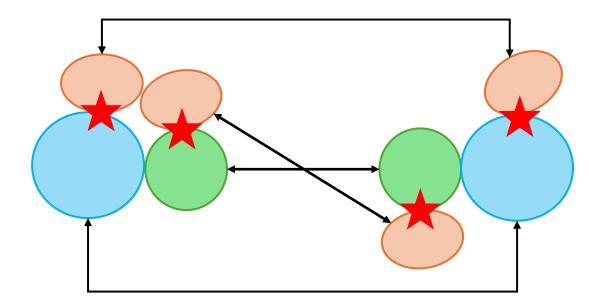


# Are we making progress?

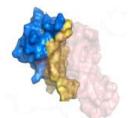
#### What about antibody-antigen interactions ?



#### A specialized scoring routine for antibody-antigen targets





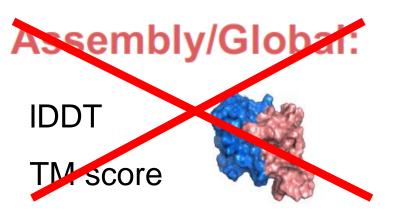


DockQ: interface size-weighted average

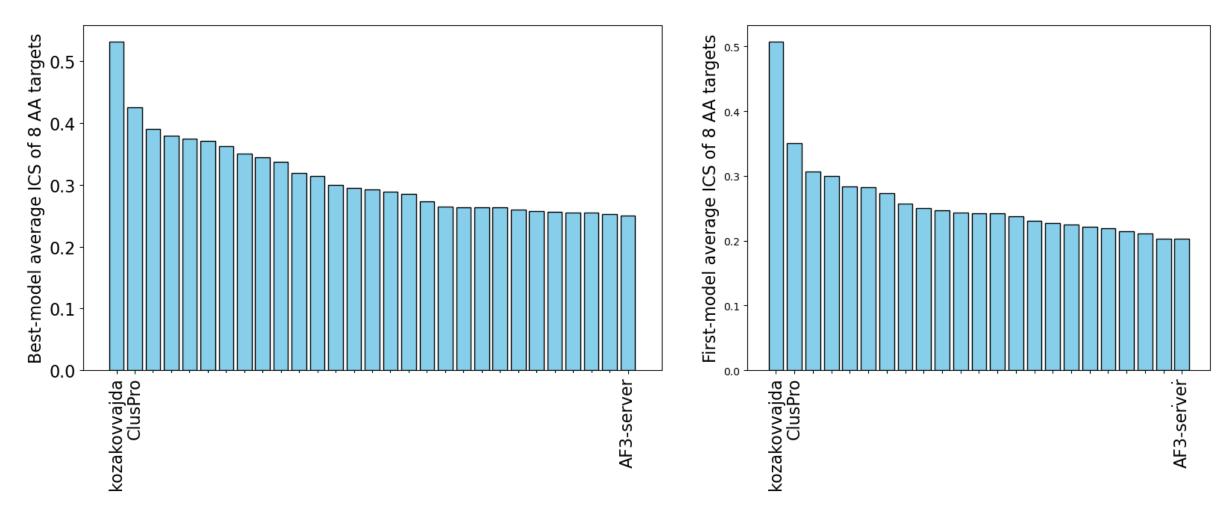
$$IPS(M,T) = J_C(M,T) = \frac{|M_{i-Res} \cap T_{i-Res}|}{|M_{i-Res} \cup T_{i-Res}|}$$

$$ICS(M,T) = 2 \cdot \frac{P(M_{cnt}, T_{cnt}) \times R(M_{cnt}, T_{cnt})}{P(M_{cnt}, T_{cnt}) + R(M_{cnt}, T_{cnt})}$$

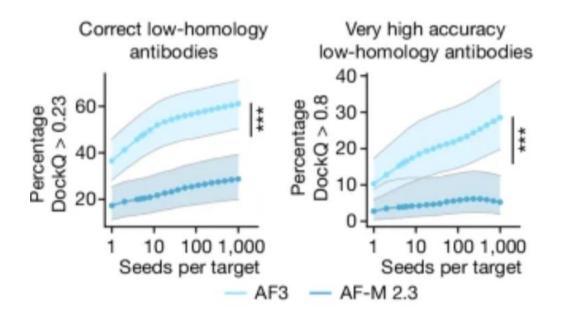
QS-best
$$(M, T) = \frac{|M_{\text{cnt}} \cap T_{\text{cnt}}|}{\max(M_{\text{cnt}}, T_{\text{cnt}})}$$

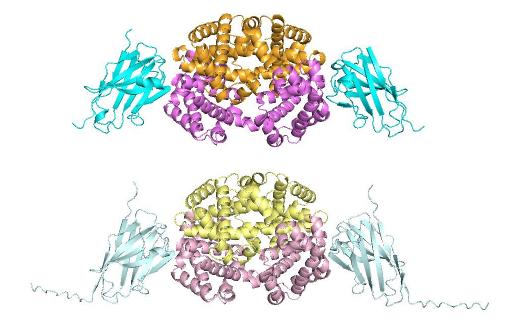


# The Kozakov group is outstanding in antibody-antigen (AA) targets



# Can people remarkably outperform AF3?

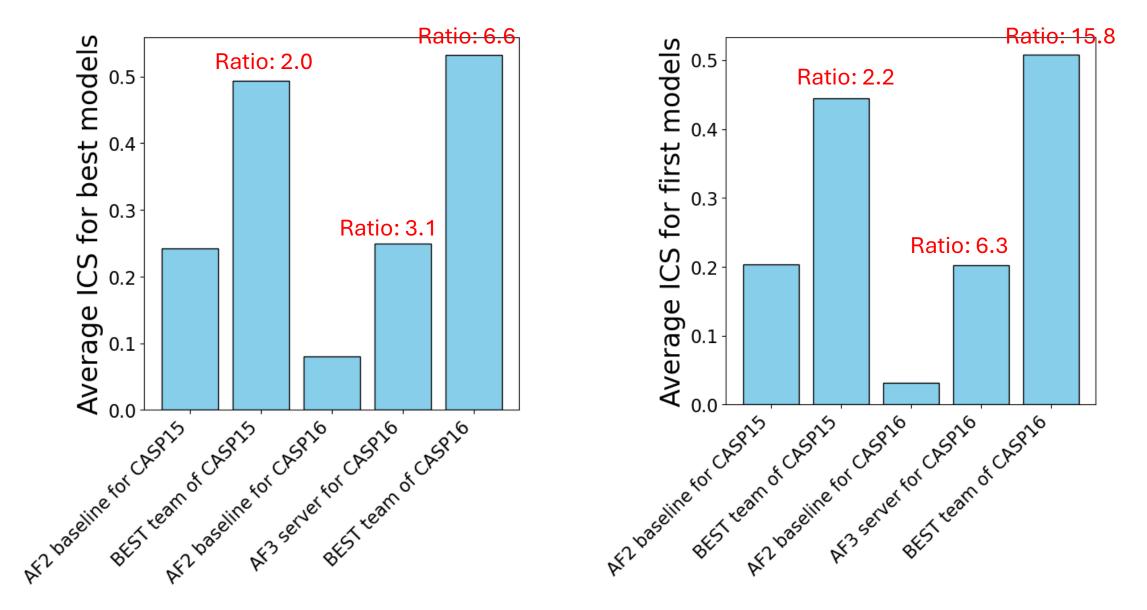




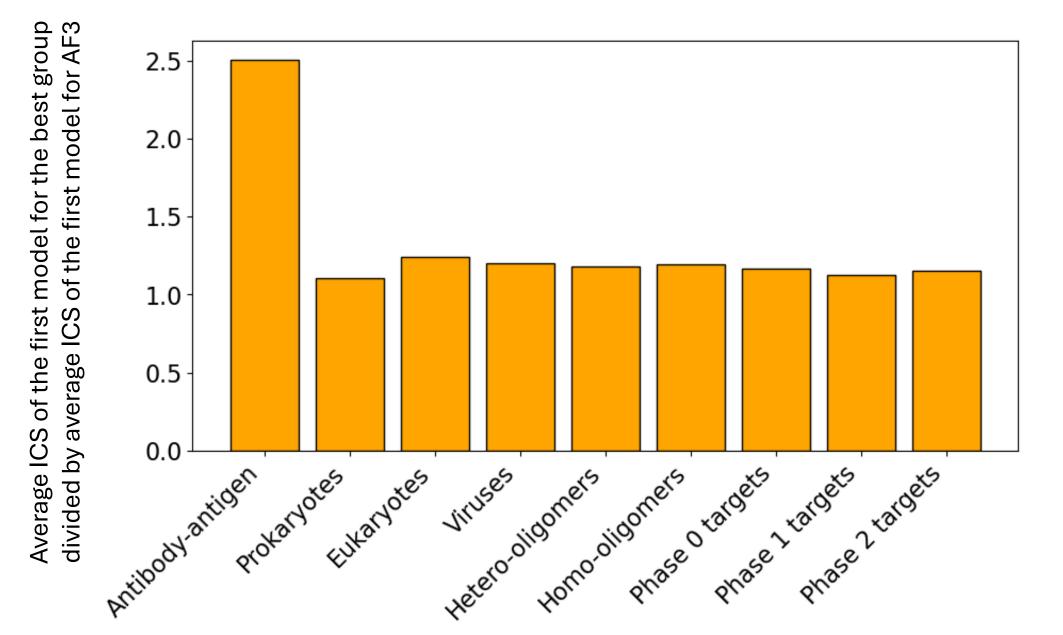
H1204 and a winning model from the Kozakov group

	H1204	H1215	H1222	H1223	H1225	H1232	H1233	H1244	DockQ > 0.23	DockQ > 0.8
Kozakov	0.8755	0.88	0.334	0.7575	0.158	0.0485	0.8995	0.0048	<b>62.5</b> %	37.5%
AF3	0.0192	0.195	0.503	0.1	0.098	0.0245	0.8845	0.0172	25%	<b>12.5%</b>

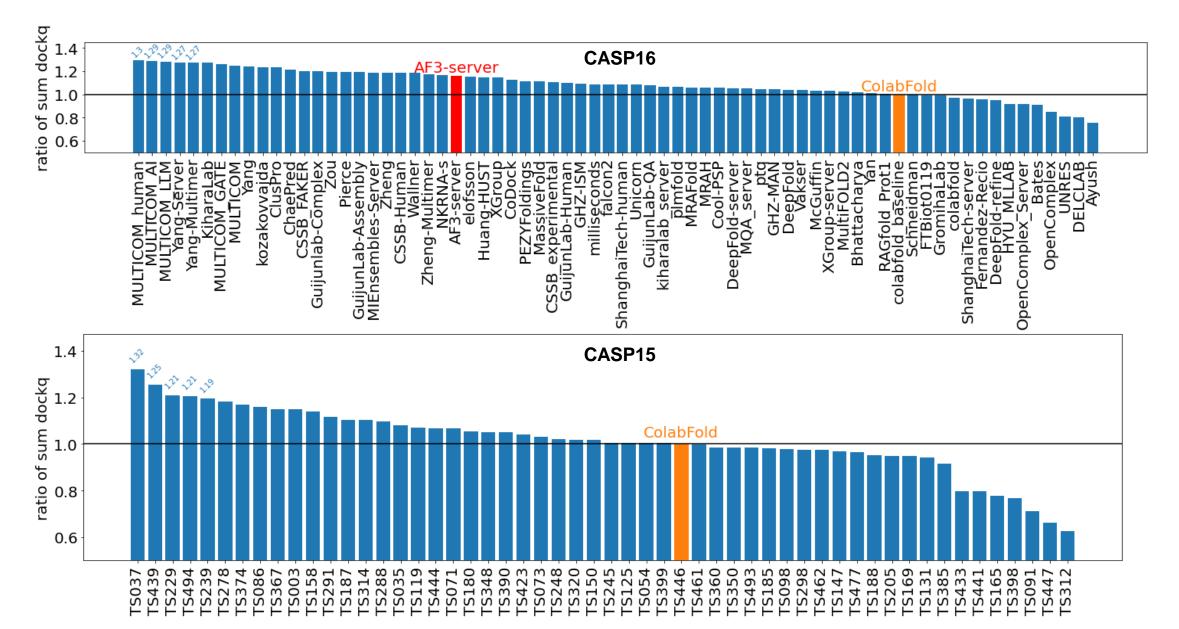
#### Are we making any progress in antibody targets?



#### In what other aspects do people outperform AF3?



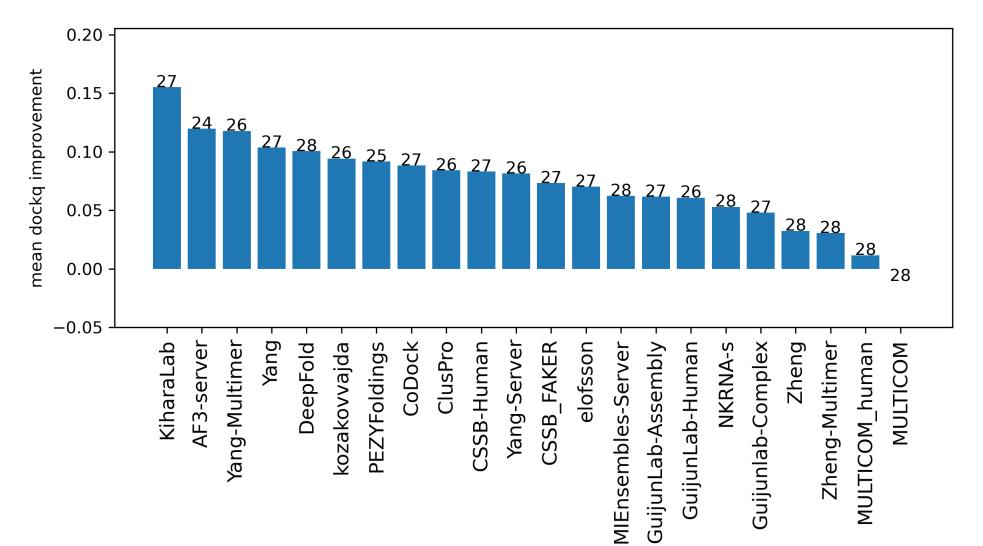
# There is no visible progress in other targets



# Additional experiments of this CASP

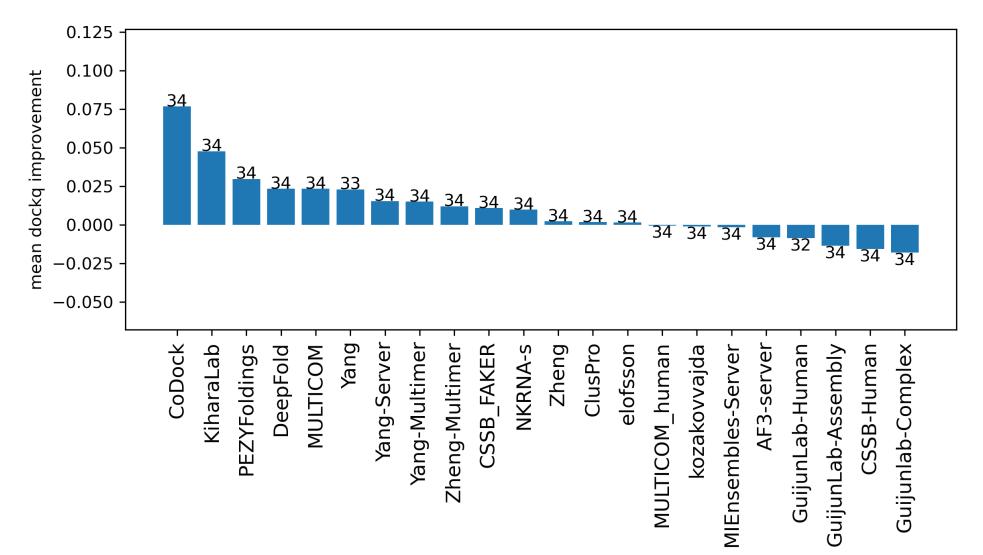
#### T1 vs T0: knowing the stoichiometry help all top 25% groups in predicting oligomer structures

mean dockq improvement from T0 to T1, using first model

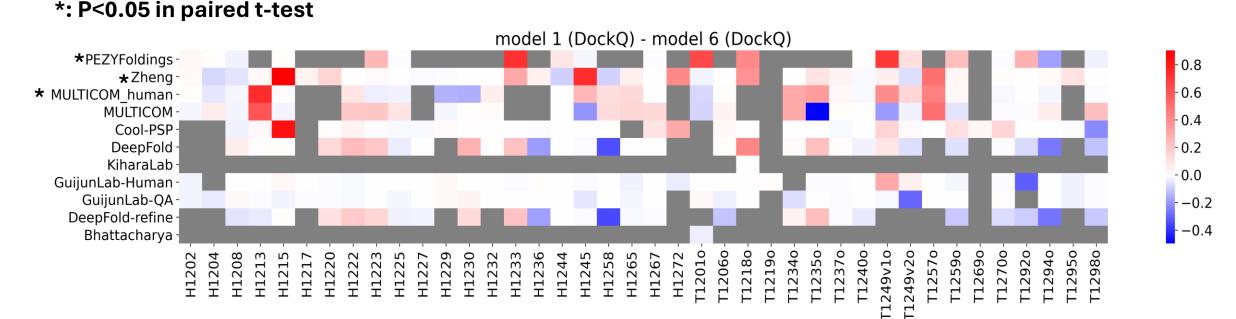


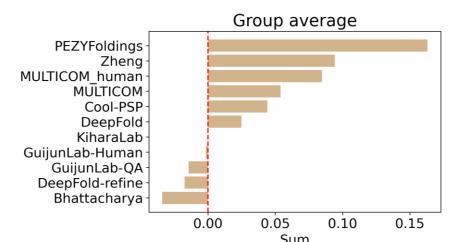
#### T2 vs T1: massive models are helpful for most of top 25% groups in predicting oligomer structures

mean dockq improvement from T1 to T2, using best model

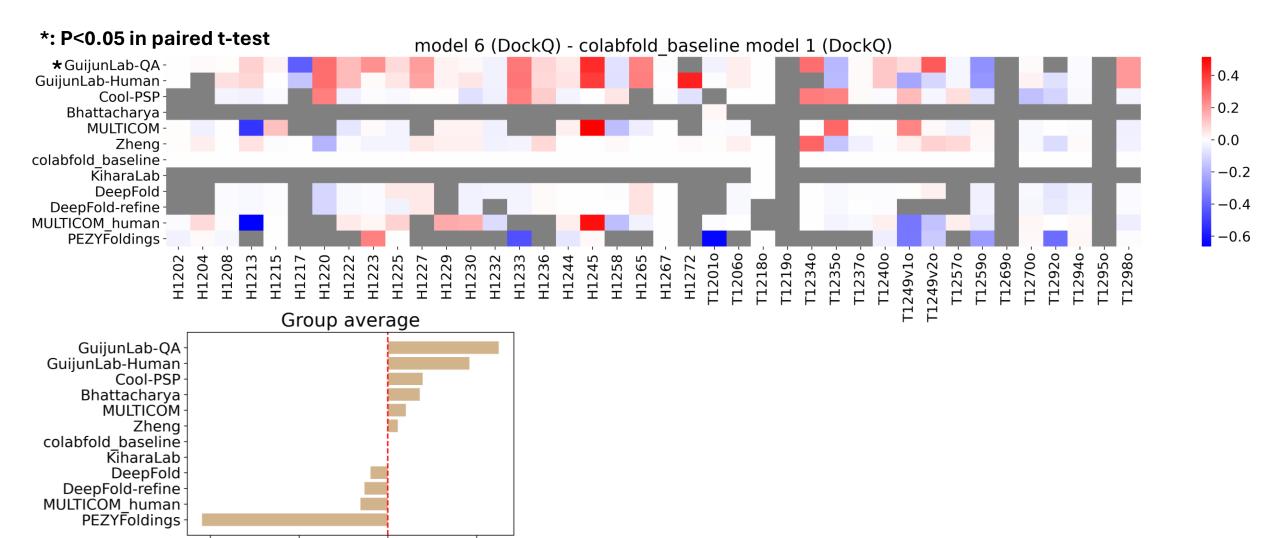


# Comparison of model 1 with model 6 of the same group reveals MSA-building ability





# Comparison of model 6 with ColabFold model 1 reveals modeling ability



0.05

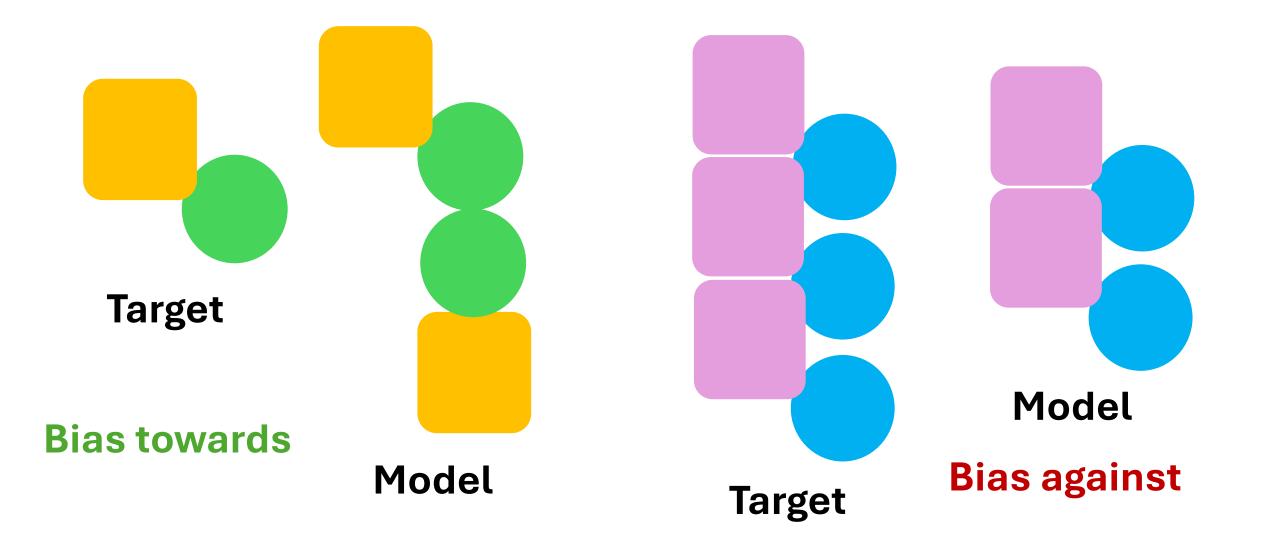
-0.10 -0.05 0.00

#### Phase 0 test people's ability to predict stoichiometry

H0208(A1B1) H0215(A1B1) H0217(A2B2C2D2E2F2) H0220(A1B4) H0222(A1B1C1) H0223(A1B1C1) H0225(A1B1C1) H0227(A1B6) H0229(A1B1) H0230(A1B1) H0232(A2B2) H0233(A2B2C2) H0236(A3B6) H0244(A2B2C2) H0245(A1B1) H0258(A1B2) H0265(A9B18) H0267(A2B2) H0272(A1B1C1D1E1F1G1H1I1) T0206o(A2) T0218o(A2) T0234o(A3) T0235o(A6) T0237o(A4) T0240o(A3) T0257o(A3) T0259o(A3) T0270o(A6) kiharalab\_server -DELCLAB -FTBiot0119 -FrederickFolding MIALAB\_gong Yan pta Ayush ARC GATE Zheng Yang ClusPro GHZ-ISM UNRES -server milliseconds falcon2 MRAFold MRAH APOLLO COAST OmniFold D3D FAKER Σ DeepFold PEZYFoldings elofsson MLLAB ShanghaiTech-server Prot1 XGroup MULTICOM GromihaLab ShanghaiTech-human CoDock Cool-PSP GHZ-MAN mialab\_prediction Bhattacharya MULTICOM\_AI GuijunLab-Human Guijunlab-Complex GuijunLab-Assembly kozakovvajda DeepFold-refine MultiFOLD2 McGuffin smg\_ulava server MULTICOM\_GATE NKRNA-5 Zheng-Multime MULTICOM\_humar experimenta XGroup-serve Yang-Multime KiharaLab GuijunLab-Q/ Yang-Serve Servei **MIEnsembles-Serve** DeepFold-serve Unicorr Schneidma SSB-Huma OpenComplex\_Serve OpenComple CSSB\_FA MULTICOM\_ RAGfold\_ AF3-MQA HYU CSSB

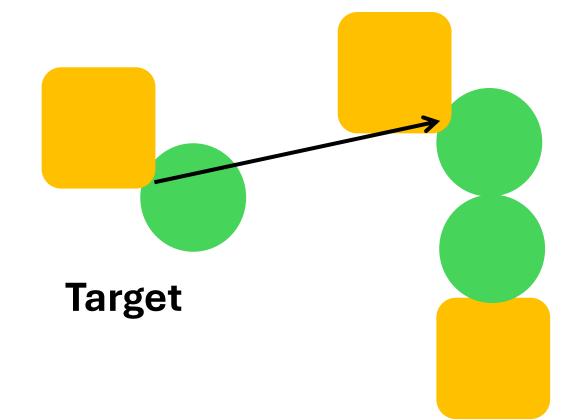
Binary Score Heatmap

# Phase 0, as well as filament targets, challenge our ability to evaluate models



# New evaluation routine for targets of unknown stoichiometry

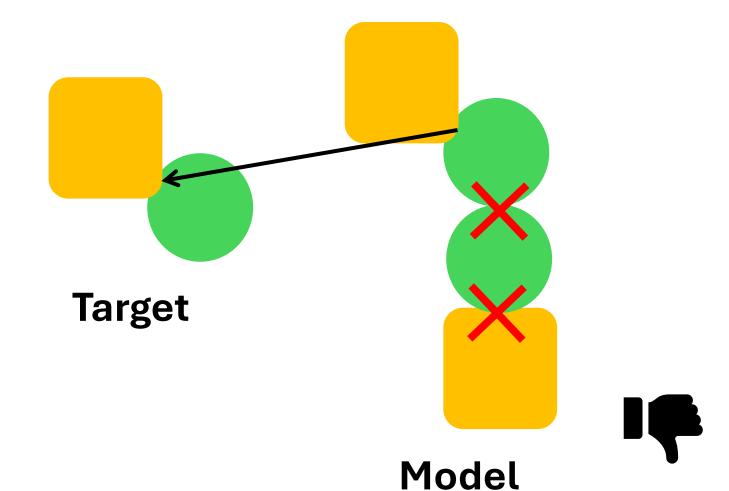
Reciprocal Best Match pair-wise scoring





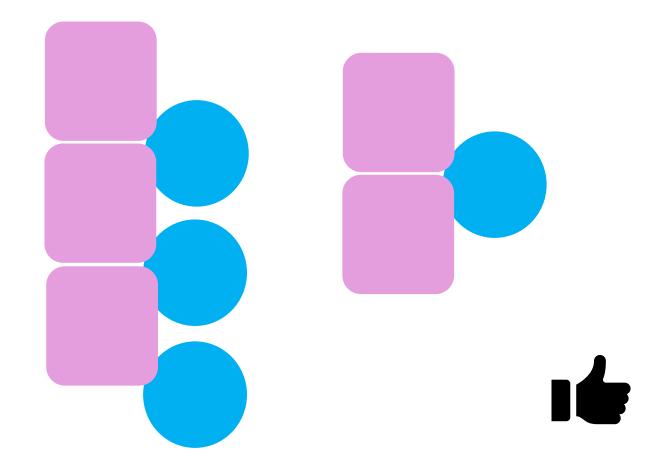
# New evaluation routine for targets of unknown stoichiometry

Reciprocal Best Match pair-wise scoring



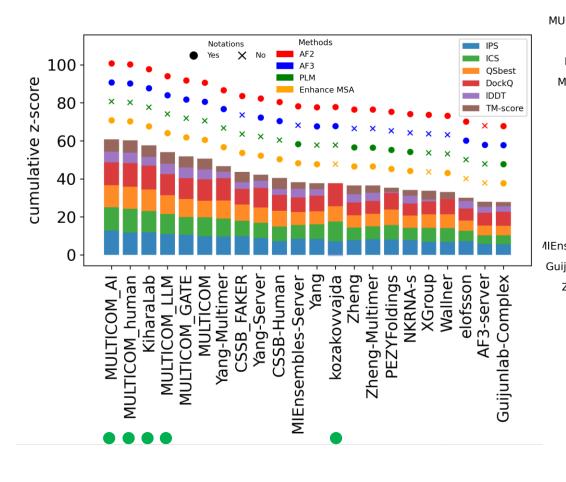
# New evaluation routine for targets of unknown stoichiometry

Reciprocal Best Match pair-wise scoring



# The ranking

### **Ranking on Phase 1 best models**

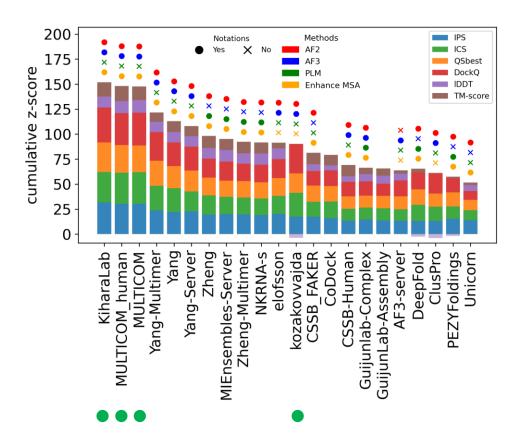


MULTICOM_AI	0.00	0.60	0.63	0.92		0.52	0.99															
JLTICOM_human -	0.40	0.00	0.60	0.87	0.94	0.60	0.99															
KiharaLab -	0.37	0.40	0.00	0.62	0.66	0.56	0.70	0.85	0.89	0.89	0.90					0.93						
MULTICOM_LLM	0.09	0.13	0.38	0.00	0.78	0.36	0.98	0.88	0.90			0.92		0.92		0.94						
MULTICOM_GATE	0.05	0.06	0.34	0.22	0.00	0.35	0.60	0.74	0.86	0.94		0.90		0.89		0.91		0.94				
kozakovvajda -	0.48	0.40	0.44	0.64	0.65	0.00	0.70	0.68	0.75	0.80	0.84	0.74	0.85	0.75	0.90	0.86	0.88	0.80	0.83	0.91	0.86	0.86
MULTICOM -	0.01	0.01	0.30	0.02	0.40	0.30	0.00	0.70	0.80	0.93	0.93	0.86		0.86		0.89		0.92	0.92			
Yang-Multimer	0.00	0.01	0.15	0.12	0.26	0.32	0.30	0.00	0.62	0.88	0.77	0.72	0.90	0.82	0.93	0.87	0.92	0.90	0.89			
CSSB_FAKER	0.02	0.03	0.11	0.10	0.14	0.25	0.20	0.38	0.00	0.54	0.56	0.70	0.72	0.77	0.74	0.65	0.74	0.68	0.82	0.86	0.90	0.92
Yang-Server-	0.00	0.00	0.11	0.01	0.06	0.20	0.07	0.12	0.46	0.00	0.49	0.57	0.73	0.70	0.81	0.74	0.82	0.80	0.78	0.89	0.94	
Yang -	0.00	0.00	0.10	0.00	0.05	0.16	0.07	0.23	0.44	0.51	0.00	0.52	0.74	0.59	0.77	0.74	0.81	0.70	0.74	0.88	0.90	0.94
CSSB-Human -	0.01	0.02	0.05	0.09	0.10	0.26	0.14	0.28	0.30	0.42	0.48	0.00	0.59	0.65	0.64	0.55	0.63	0.79	0.76	0.76	0.84	0.86
sembles-Server-	0.00	0.00	0.05	0.00	0.01	0.15	0.00	0.10	0.28	0.27	0.26	0.41	0.00	0.46	0.69	0.57	0.88	0.62	0.67	0.98	0.81	0.87
ijunlab-Complex -	0.02	0.02	0.02	0.08	0.11	0.25	0.14	0.18	0.23	0.30	0.41	0.35	0.54	0.00	0.57	0.53	0.58	0.49	0.61	0.66	0.57	0.64
Zheng-Multimer	0.00	0.00	0.02	0.00	0.00	0.10	0.00	0.07	0.26	0.19	0.23	0.36	0.31	0.43	0.00	0.52	0.51	0.58	0.65	0.98	0.77	0.83
PEZYFoldings -	0.03	0.02	0.07	0.06	0.09	0.14	0.12	0.13	0.35	0.26	0.26	0.46	0.43	0.47	0.48	0.00	0.44	0.44	0.55	0.54	0.70	0.7
Zheng-	0.00	0.00	0.04	0.00	0.00	0.12	0.00	0.08	0.26	0.18	0.18	0.37	0.12	0.42	0.49	0.56	0.00	0.60	0.60	0.80	0.73	0.80
XGroup -	0.00	0.00	0.00	0.05	0.06	0.20	0.08	0.10	0.32	0.20	0.30	0.21	0.38	0.51	0.42	0.56	0.40	0.00	0.53	0.50	0.74	0.73
Wallner-	0.00	0.00	0.01	0.04	0.05	0.17	0.08	0.11	0.18	0.22	0.26	0.24	0.33	0.39	0.35	0.45	0.40	0.47	0.00	0.45	0.61	0.6
NKRNA-s	0.00	0.00	0.01	0.00	0.00	0.09	0.00	0.04	0.14	0.11	0.12	0.24	0.02	0.34	0.02	0.46	0.20	0.50	0.55	0.00	0.66	0.75
elofsson -	0.00	0.00	0.00	0.00	0.00	0.14	0.01	0.00	0.10	0.06	0.10	0.16	0.19	0.43	0.23	0.30	0.27	0.26	0.39	0.34	0.00	0.65
AF3-server	0.00	0.00	0.00	0.00	0.01	0.14	0.01	0.02	0.08	0.02	0.06	0.14	0.13	0.36	0.17	0.29	0.20	0.27	0.33	0.25	0.35	0.00
·	Ă.	an -	ab -	Σ	TE -	da -	- MC	- Jer	ER -	/er-	Yang -	an -	er -	ex	ler -	- sɓ	- Gu	- dn	her -	-s-	on -	/er-
	MULTICOM_AI	MULTICOM_human	KiharaLab	MULTICOM_LLM	MULTICOM_GATE	kozakovvajda	MULTICOM	Yang-Multimer	CSSB_FAKER	Yang-Server	Yа	CSSB-Human	MIEnsembles-Servei	Guijunlab-Complex	Zheng-Multimer	PEZYFoldings	Zheng	XGroup	Wallner	NKRNA-S	elofsson	AF3-server

- 0.25

- 0

#### Ranking on best models for targets from all phases



KiharaLab	0.00	0.81	0.87	0.96	0.85	0.98	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
MULTICOM_human	0.19	0.00	0.53	0.90	0.68																	
MULTICOM	0.13	0.47	0.00	0.85	0.68																	
Yang-Multimer	0.04	0.10	0.15	0.00	0.54	0.85																
kozakovvajda	0.15	0.32	0.32	0.46	0.00	0.67	0.71		0.86	0.83	0.92	0.92		0.93	0.93	0.94	0.94					
Yang	0.02	0.01	0.02	0.15	0.33	0.00	0.69	0.82	0.92	0.87		0.90	0.94									
Yang-Server	0.01	0.01	0.02	0.01	0.29	0.31	0.00	0.75	0.87	0.86		0.88	0.93									
ClusPro	0.01	0.01	0.03	0.05	0.02	0.18	0.25	0.00	0.53	0.58	0.70	0.60	0.78	0.72	0.71	0.84	0.86	0.86	0.87			
Zheng	0.00	0.00	0.00	0.02	0.14	0.08	0.13	0.47	0.00	0.54	0.69	0.63	0.60	0.81	0.88	0.80	0.82	0.92	0.88	0.86		0.93
elofsson	0.00	0.00	0.01	0.01	0.17	0.13	0.14	0.42	0.46	0.00	0.56	0.67	0.68	0.61	0.58	0.94	0.91	0.78		0.90		
MIEnsembles-Server	0.00	0.00	0.00	0.01	0.08	0.04	0.04	0.30	0.32	0.45	0.00	0.52	0.54	0.79	0.88	0.70	0.80	0.84	0.86	0.85	0.93	0.92
PEZYFoldings	0.00	0.00	0.01	0.01	0.08	0.10	0.12	0.40	0.37	0.33	0.48	0.00	0.73	0.53	0.53	0.57	0.73	0.89	0.72	0.87	0.91	0.90
CSSB_FAKER	0.00	0.00	0.00	0.01	0.05	0.06	0.07	0.22	0.40	0.32	0.47	0.27	0.00	0.52	0.56	0.62	0.71	0.80	0.77		0.84	0.89
Zheng-Multimer	0.00	0.00	0.00	0.01	0.07	0.04	0.02	0.28	0.19	0.39	0.21	0.47	0.48	0.00	0.67	0.66	0.73	0.80	0.81	0.83	0.90	0.90
NKRNA-s	0.00	0.00	0.00	0.00	0.07	0.02	0.02	0.29	0.12	0.42	0.12	0.47	0.44	0.33	0.00	0.68	0.74	0.78	0.81	0.78	0.90	0.89
GuijunLab-Assembly	0.00	0.00	0.00	0.00	0.06	0.04	0.03	0.16	0.20	0.06	0.29	0.43	0.38	0.34	0.32	0.00	0.58	0.59	0.51	0.78	0.85	0.80
Guijunlab-Complex	0.00	0.00	0.00	0.00	0.06	0.01	0.01	0.14	0.18	0.09	0.20	0.27	0.29	0.27	0.26	0.42	0.00	0.47	0.46	0.66	0.79	0.76
CoDock	0.00	0.00	0.00	0.00	0.04	0.00	0.01	0.14	0.08	0.22	0.16	0.11	0.20	0.20	0.22	0.41	0.54	0.00	0.59	0.54	0.85	0.73
AF3-server	0.00	0.00	0.00	0.00	0.03	0.01	0.02	0.13	0.12	0.04	0.14	0.28	0.23	0.20	0.19	0.49	0.55	0.41	0.00	0.51	0.82	0.73
CSSB-Human	0.00	0.00	0.00	0.00	0.01	0.02	0.01	0.04	0.14	0.10	0.15	0.13	0.04	0.17	0.22	0.22	0.34	0.46	0.49	0.00	0.64	0.73
Unicorn	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.03	0.05	0.01	0.07	0.09	0.15	0.10	0.10	0.15	0.21	0.15	0.18	0.36	0.00	0.64
DeepFold	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.01	0.07	0.05	0.08	0.10	0.11	0.10	0.11	0.20	0.24	0.27	0.27	0.27	0.36	0.00
	KiharaLab -	MULTICOM_human -	MULTICOM -	Yang-Multimer -	kozakovvajda -	- Yang	Yang-Server -	- ClusPro	Zheng -	elofsson -	MIEnsembles-Server -	PEZYFoldings -	CSSB_FAKER -	Zheng-Multimer -	NKRNA-S -	GuijunLab-Assembly -	Guijunlab-Complex -	CoDock -	AF3-server -	CSSB-Human -	Unicorn -	DeepFold -

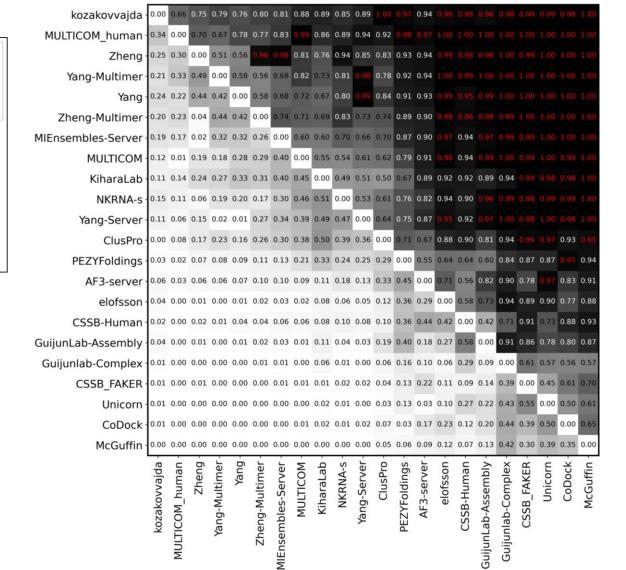
0.75

0.5

0.25

- 0

#### Ranking on first models for targets from all phases

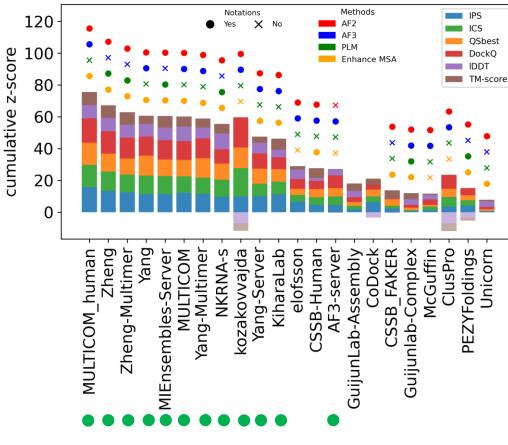


-0.75

0.5

0.25

- 0



#### Additional rankings reveal interest aspects of groups



IPS

ICS

QSbest

DockO

TM-score

IDDT

X X

Zheng

server FAKER

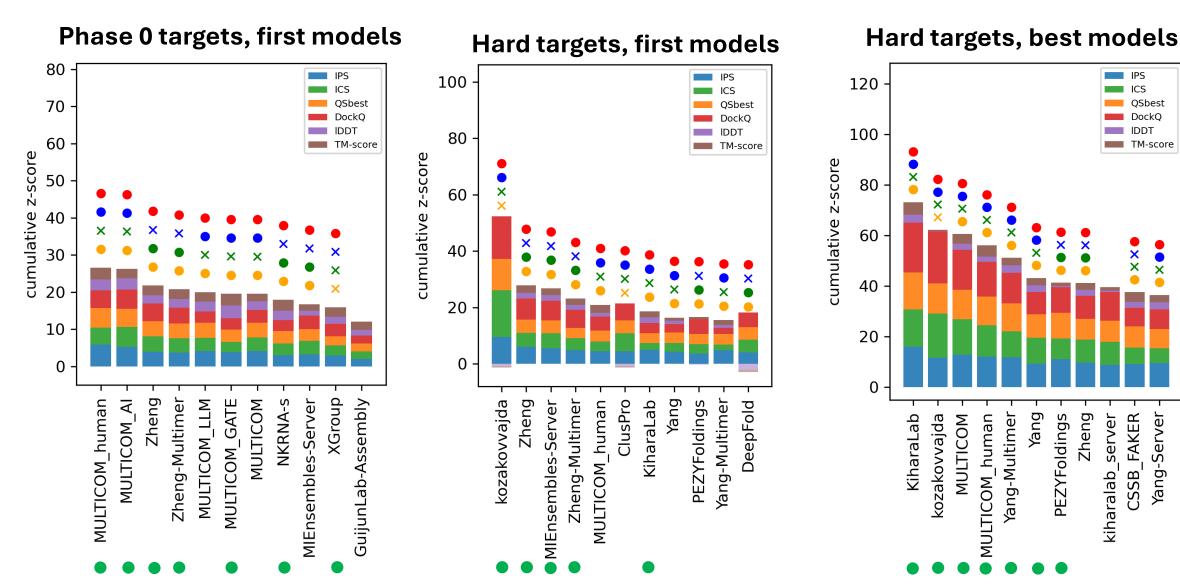
kiharalab

Yang-Server

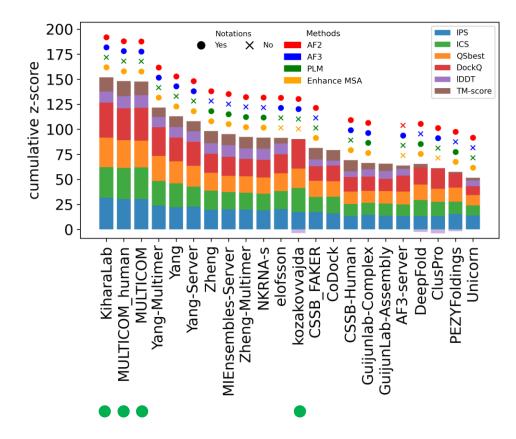
CSSB\_

Yang

PEZYFoldings



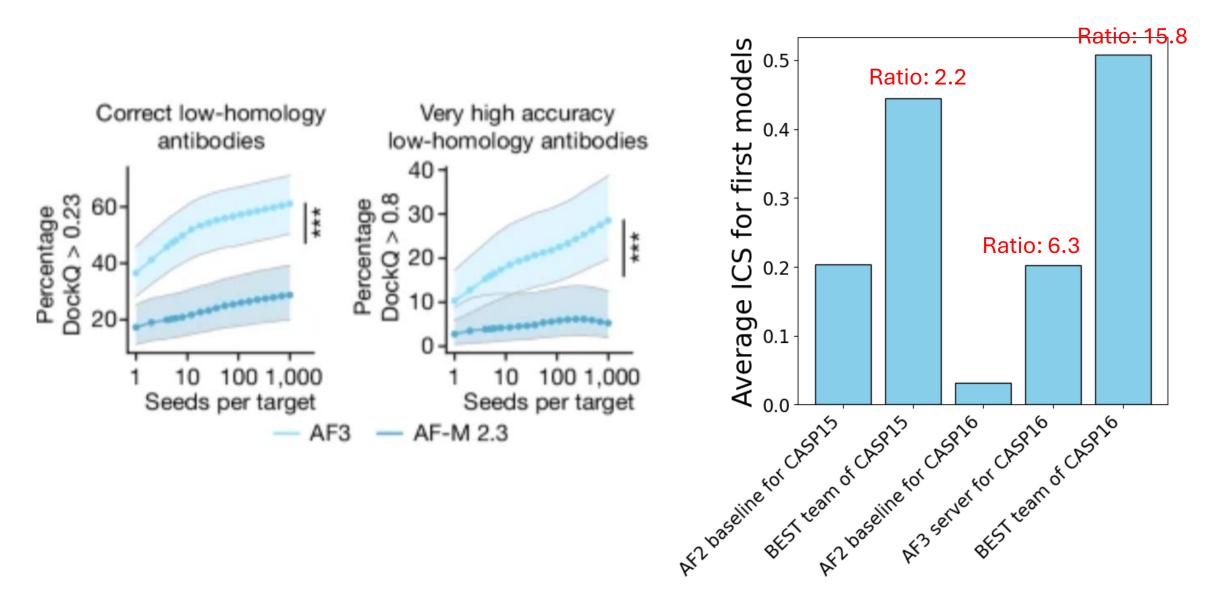
### Final "ranking" (with ties) for oligomers **Best models over all phases**



-										_						_			_			_
KiharaLab -	_				0.85																	
MULTICOM_human	0.19	0.00	0.53	0.90	0.68																	
MULTICOM	0.13	0.47	0.00	0.85	0.68																	
Yang-Multimer	0.04	0.10	0.15	0.00	0.54	0.85																
kozakovvajda -	0.15	0.32	0.32	0.46	0.00	0.67	0.71		0.86	0.83	0.92	0.92		0.93	0.93	0.94	0.94					
Yang -	0.02	0.01	0.02	0.15	0.33	0.00	0.69	0.82	0.92	0.87		0.90	0.94									
Yang-Server-	0.01	0.01	0.02	0.01	0.29	0.31	0.00	0.75	0.87	0.86		0.88	0.93									
ClusPro	0.01	0.01	0.03	0.05	0.02	0.18	0.25	0.00	0.53	0.58	0.70	0.60	0.78	0.72	0.71	0.84	0.86	0.86	0.87			
Zheng	0.00	0.00	0.00	0.02	0.14	0.08	0.13	0.47	0.00	0.54	0.69	0.63	0.60	0.81	0.88	0.80	0.82	0.92	0.88	0.86		0.93
elofsson -	0.00	0.00	0.01	0.01	0.17	0.13	0.14	0.42	0.46	0.00	0.56	0.67	0.68	0.61	0.58	0.94	0.91	0.78		0.90		
MIEnsembles-Server	0.00	0.00	0.00	0.01	0.08	0.04	0.04	0.30	0.32	0.45	0.00	0.52	0.54	0.79	0.88	0.70	0.80	0.84	0.86	0.85	0.93	0.92
PEZYFoldings -	0.00	0.00	0.01	0.01	0.08	0.10	0.12	0.40	0.37	0.33	0.48	0.00	0.73	0.53	0.53	0.57	0.73	0.89	0.72	0.87	0.91	0.90
CSSB_FAKER	0.00	0.00	0.00	0.01	0.05	0.06	0.07	0.22	0.40	0.32	0.47	0.27	0.00	0.52	0.56	0.62	0.71	0.80	0.77		0.84	0.89
Zheng-Multimer -	0.00	0.00	0.00	0.01	0.07	0.04	0.02	0.28	0.19	0.39	0.21	0.47	0.48	0.00	0.67	0.66	0.73	0.80	0.81	0.83	0.90	0.90
NKRNA-s -	0.00	0.00	0.00	0.00	0.07	0.02	0.02	0.29	0.12	0.42	0.12	0.47	0.44	0.33	0.00	0.68	0.74	0.78	0.81	0.78	0.90	0.89
GuijunLab-Assembly -	0.00	0.00	0.00	0.00	0.06	0.04	0.03	0.16	0.20	0.06	0.29	0.43	0.38	0.34	0.32	0.00	0.58	0.59	0.51	0.78	0.85	0.80
Guijunlab-Complex -	0.00	0.00	0.00	0.00	0.06	0.01	0.01	0.14	0.18	0.09	0.20	0.27	0.29	0.27	0.26	0.42	0.00	0.47	0.46	0.66	0.79	0.76
CoDock	0.00	0.00	0.00	0.00	0.04	0.00	0.01	0.14	0.08	0.22	0.16	0.11	0.20	0.20	0.22	0.41	0.54	0.00	0.59	0.54	0.85	0.73
AF3-server-	0.00	0.00	0.00	0.00	0.03	0.01	0.02	0.13	0.12	0.04	0.14	0.28	0.23	0.20	0.19	0.49	0.55	0.41	0.00	0.51	0.82	0.73
CSSB-Human -	0.00	0.00	0.00	0.00	0.01	0.02	0.01	0.04	0.14	0.10	0.15	0.13	0.04	0.17	0.22	0.22	0.34	0.46	0.49	0.00	0.64	0.73
Unicorn -	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.03	0.05	0.01	0.07	0.09	0.15	0.10	0.10	0.15	0.21	0.15	0.18	0.36	0.00	0.64
DeepFold	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.01	0.07	0.05	0.08	0.10	0.11	0.10	0.11	0.20	0.24	0.27	0.27	0.27	0.36	0.00
L	- de	- ue	Σ	er-	la -	- Gu	er-	ro -	- ɓเ	- uc	er -	gs -		er -		- Y	- Xə	÷	er-	- ue	Ė	þ
	KiharaLab	MULTICOM_human	MULTICOM	Yang-Multimer	kozakovvajda	Yang	Yang-Server	ClusPro	Zheng	elofsson	MIEnsembles-Server	PEZYFoldings	CSSB_FAKER	Zheng-Multimer	NKRNA-S	GuijunLab-Assembly	Guijunlab-Complex	CoDock	AF3-server	CSSB-Human	Unicorn	DeepFold
	Kïł	μÖ	MU	N-bu	ozakı		Yang			Ψ	aldr	EZYF	CSSB	Ng-N	Z	sA-de	ab-C		AF	SSB		ŏ
		ורדונ		К	X						sen	Ъ	Ŭ	Zhe		nLâ	ijunl			0		
		ML									IIEn					uijı	Gui					

- 0.25

### Huge progress in antibody-antigen interactions



# Final thoughts on oligomer prediction

0. Exciting progress in antigen-antibody interactions. We may want more antibody targets in the future to more robustly evaluate the progress.

1. Protein complex modeling is not "solved": each group gets a subset correctly.

2. We think Phase 0 should be the future of oligomer prediction.

3. Evaluating targets with unknown stoichiometry needs better tools and we provided a start.

4. Weak interactions and multiple alternative interfaces are hard to predict.

5. Unseen flexibility in experimental structures could be an issue.