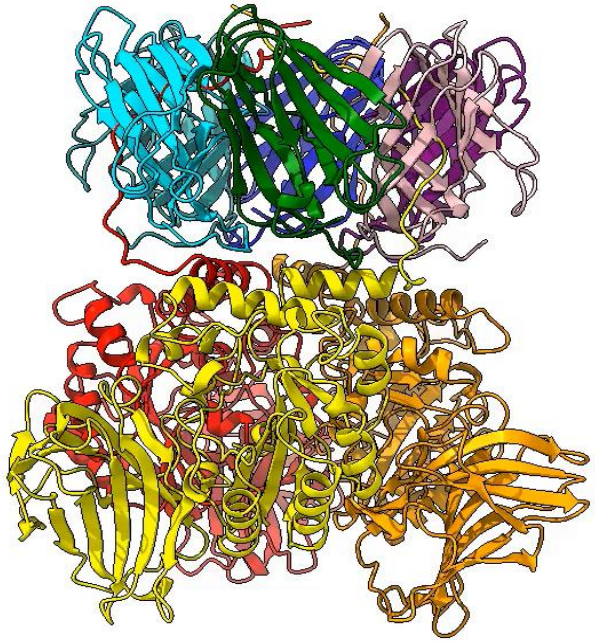
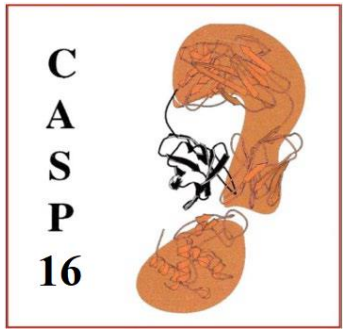


Improving AlphaFold2/3-based Protein Complex Structure Prediction with MULTICOM4



Jianlin Cheng

University of Missouri, Columbia, USA

Initial MULTICOM4 System Based on AlphaFold2

Sequence Structure Databases

MSA Concatenation & Different Parameters

Model Generation (25 – 1000 per MSA)

Model Ranking

Predictors

Each Subunit of a Complex

Uniref BFD, Magnify, UniProt, JGI, TaraDB, MetaSourceDB

AF2 Default DB (updated)

AF2 Prediction & FoldSeek

AlphaFoldDB & PDB

Interaction (homomer: 11; heteromer:17)

AF2 Default (14)

DeepMSA (homomer: 8 heteromer: 16)

AFsample (6)

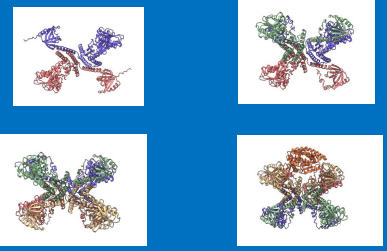
FoldSeek (homomer:16 ; heteromer: 14)

~50 - 70 MSAs + One ColabFold MSA

AlphaFold2-Multimer

32 A100 GPUs
8 H100 GPUs

~1000 – 70,000 AF2 models



1. AF2 Confidence Score
2. Gate: Graph Transformer
3. GCPNet-EMA
4. ENQA
5. Pairwise Similarity

MULITCOM_human (345)

MULTICOM (051)

MULITCOM_GATE (425)

MULITCOM_AI (331)

MULTICOM_LLM (319)

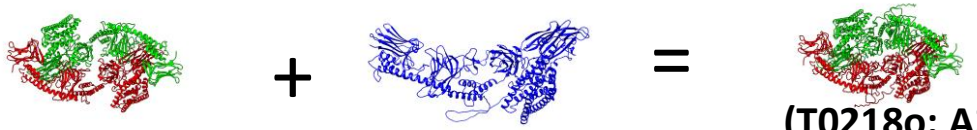
Diversity Rule

Exception Handling

Divide and Conquer

(H0227: A1B6)

Template-Based Modeling (Modeller + AlphaFold)



(T0218o: A2)

Updated MULTICOM4 System Based on AlphaFold2 & 3

Sequence Structure Databases

MSA Concatenation & Different Parameters

Model Generation (25 – 1000 per MSA)

Model Ranking

Predictors

Each Subunit of a Complex

Uniref BFD, Magnify, UniProt, JGI, TaraDB, MetaSourceDB

AF2 Default DB (updated)

AF2 Prediction & FoldSeek

AlphaFoldDB & PDB

Interaction (homomer: 11; heteromer:17)

AF2 Default (14)

DeepMSA (homomer: 8 heteromer: 16)

AFsample (6)

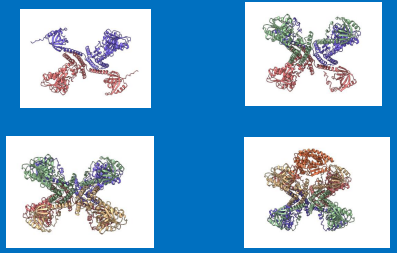
FoldSeek (homomer:16 ; heteromer: 14)

~50 - 70 MSAs + One ColabFold MSA

AlphaFold2-Multimer

32 A100 GPUs
8 H100 GPUs

~1000 – 70,000 AF2 models



(confidence score)

1. AF2 Confidence Score
2. Gate: Graph Transformer
3. GCPNet-EMA
4. ENQA
5. Pairwise Similarity

MULITCOM_human (345)

MULTICOM (051)

MULITCOM_GATE (425)

MULITCOM_AI (331)

MULTICOM_LLM (319)

Diversity Rule

Whole Complex

AlphaFold 3

~100 – 7000 AF3 models

AF3 Ranking Score

Exception Handling

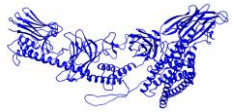
Divide and Conquer

A + B = + =

(H0227: A1B6)

Template-Based Modeling (Modeller + AlphaFold)

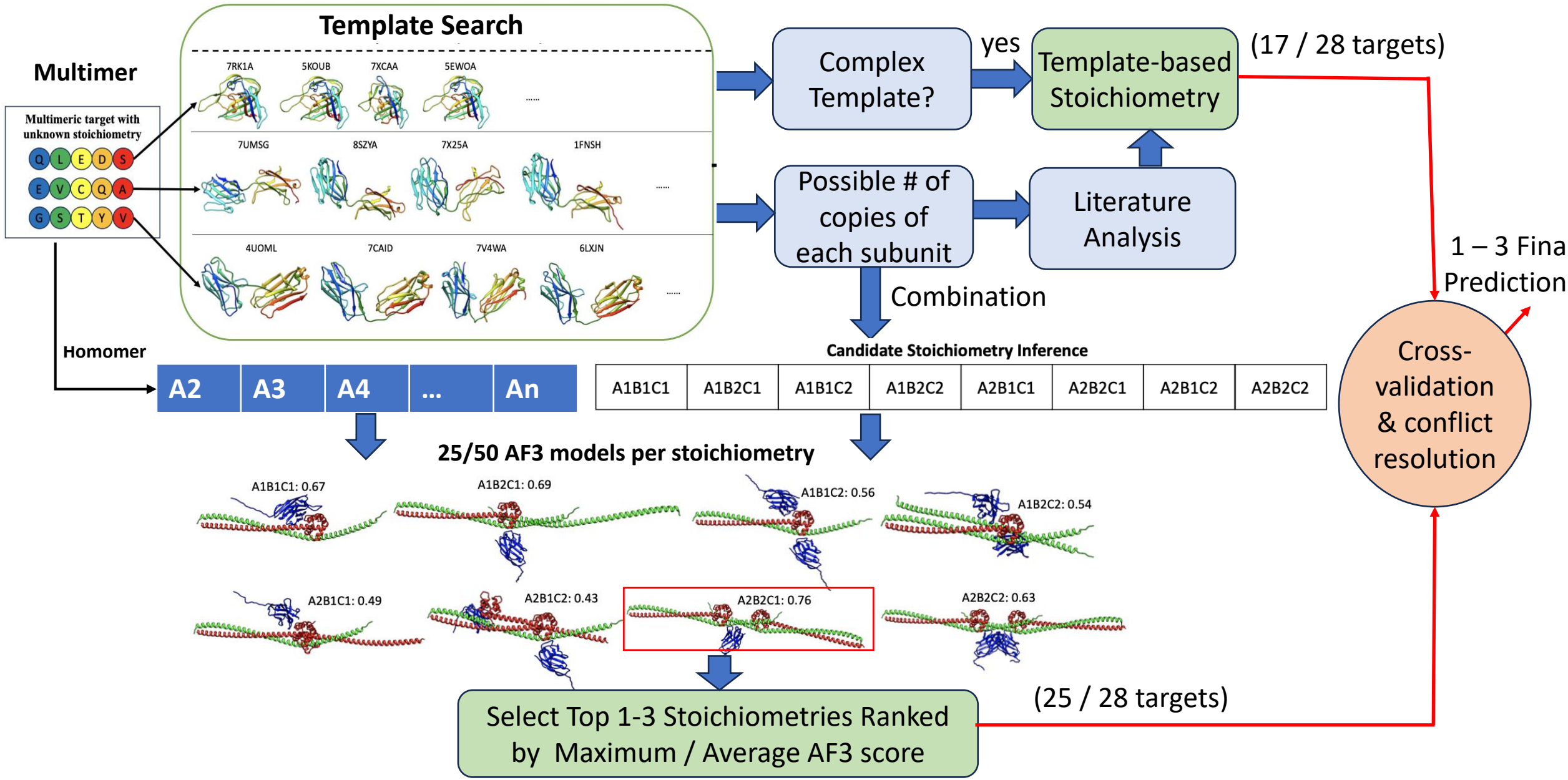
+



=

(T0218o: A2)

Stoichiometry Prediction (Phase 0)



Stoichiometry Predictions of MULTICOM

(28 targets: 19 heteromers 9 homomers)

18 Easy / Medium

Accuracy of Top 1 = 71%
Best of Top 3 = 93%

10 Hard

Note: MULTICOM_human is similar to MULTICOM

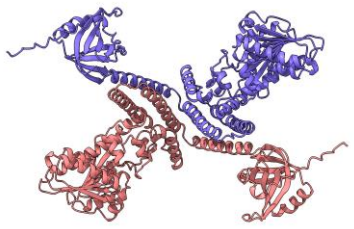
Target (28)	Difficulty (% CASP groups being correct)	True stoichiometry	AF3 maximum ranking score (AF-max). Red – corr.	AF3 average ranking score (AF-avg). Red - correct	Template-based prediction (TB). Red - correct	Final prediction (Top 1): Red - correct	Decision choice	Best of Top 3
H0225	90%	A1B1C1	A1B1C1	A1B1C1	A1B1C1	A1B1C1	AF-max + AF-avg + TB	Yes
H0222	88%	A1B1C1	A1B1C1	A1B1C1	A1B1C1	A1B1C1	AF-max + AF-avg + TB	Yes
H0223	87%	A1B1C1	A1B1C1	A1B1C1	A1B1C1	A1B1C1	AF-max + AF-avg + TB	Yes
H0245	85%	A1B1	A1B1	A1B1	N/A	A1B1	AF-max + AF-avg	Yes
H0215	75%	A1B1	A1B1	A2B2	A1B1	A1B1	AF-max + TB	Yes
T0206o	72%	A2	A2 (AFM)	A2 (AFM)	A2	A2	AF-max + AF-avg + TB	Yes
T0257o	72%	A3	A3	A2	A3	A3	AF-max + TB	Yes
T0259o	71%	A3	A4	A3	N/A	A3	AF-avg	Yes
H0208	66%	A1B1	A1B1 (AFM)	A1B1 (AFM)	A2B2	A2B2	TB	Yes
T0240o	63%	A3	A3	A2	N/A	A3	AF-max	Yes
H0232	57%	A2B2	A2B2	A2B2	A2B2	A2B2	AF-max + AF-avg + TB	Yes
H0272	57%	A1B1C1D1E1F1G1H1I1	N/A	N/A	A1B1C1D1E1F1G1H1I1	A1B1C1D1E1F1G1H1I1	TB	Yes
H0220	53%	A1B4	A1B1	A1B1	A1B4	A1B4	TB	Yes
T0237o	41%	A4	A4	A4	N/A	A4	AF-max + AF-avg	Yes
H0233	38%	A2B2C2	A1B1C1	A1B1C1	A2B2C2	A2B2C2	TB	Yes
H0227	37%	A1B6	N/A	N/A	A1B6	A1B6	TB	Yes
T0235o	37%	A6	A6	A6	N/A	A6	AF-max + AF-avg	Yes
T0234o	34%	A3	A3	A3	N/A	A3	AF-max + AF-avg	Yes
H0229	31%	A1B1	A1B1	A2B2	N/A	A2B2	AF-avg	Yes
H0230	31%	A1B1	A2B2	A2B2	N/A	A2B2	AF-max + AF-avg	Yes
H0217	28%	A2B2C2D2E2F2	N/A	N/A	A2B2C2D2E2F?(1/2)	A2B2C2D2E2F1	TB	No
H0258	25%	A1B2	A1B1	A1B2	A1B2	A1B2	AF-avg + TB	Yes
T0218o	25%	A2	A3	A3	A2	A2	TB	Yes
H0236	24%	A3B6	A3B3	A3B3	N/A	A3B3	AF-max + AF-avg	Yes
T0270o	22%	A6	A3	A3	A6	A6	TB	Yes
H0244	9%	A2B2C2	A2B2C1	A2B2C1	N/A	A2B2C1	AF-max + AF-avg	Yes
H0267	7%	A2B2	A1B1	A1B1	A1B1	A1B1	AF-max + AF-avg + TB	Yes
H0265	0%	A9B18	A1B1	A1B1	N/A	A1B1	AF-max + AF-avg	No
Accuracy			14 / 25 = 56% (50%)	12 / 25 = 48% (43%)	14 / 17 = 82% (50%)	20 / 28 = 71%	71%	93%

Homomultimer (Top-1 Accuracy = 100%)

T0237o, A4, No template

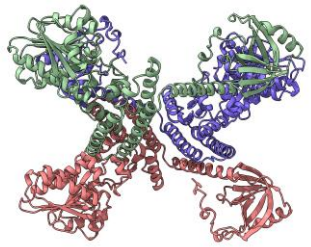
A2

AF3 ranking score: 0.69



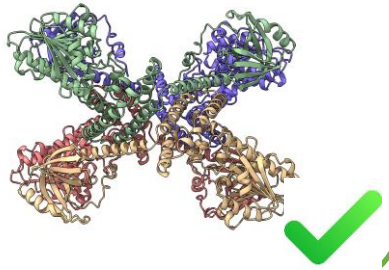
A3

AF3 ranking score: 0.88



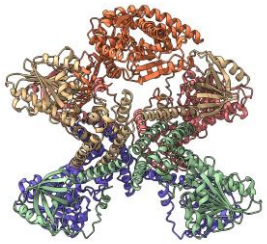
A4

AF3 ranking score: 0.91



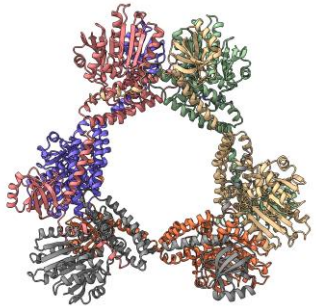
A5

AF3 ranking score: 0.61



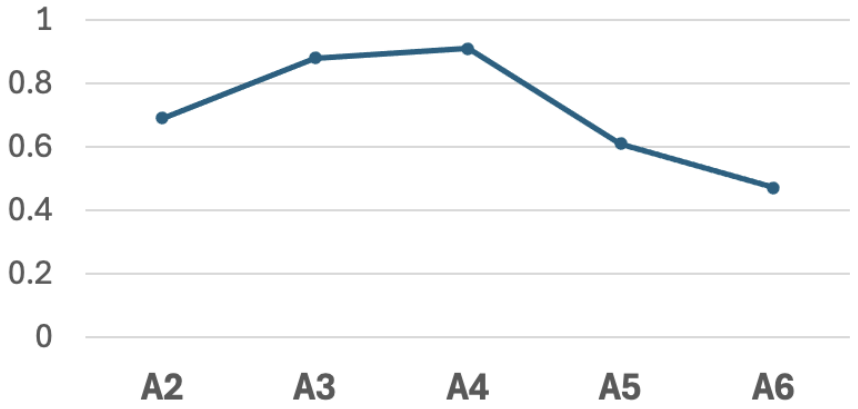
A6

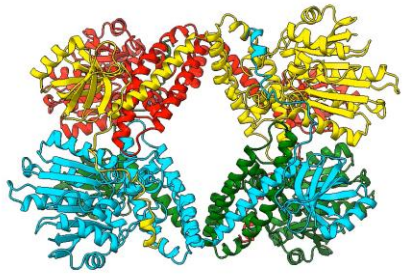
AF3 ranking score: 0.47



41% correctness: 41%
CASP16 groups are correct

AF3 Max Ranking Score





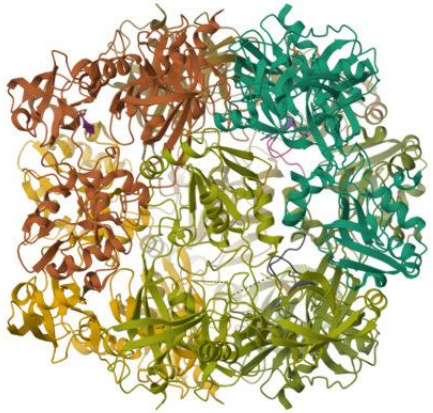
Predicted A4 Model

T0270o, A6

Use templates to correct AF3 predictions

Target	True Stoichiometry	Candidates	Stoichiometry selected by max ranking score	Final prediction of MULTICOM
T0270o	A6	A2,A3,A4,A5,A6,A9	A3	A6

Template



PDB code: 4FLN

Stoichiometry: A6B4C2

22% correctness

Why does it work well for homo-multimers?

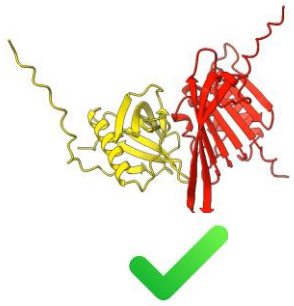
- Fewer choices
- more symmetry
- more template information

Hetero-Multimers: Top-1 Accuracy = 58%; Best-of-top-3 Accuracy = 89.5%

H0245, A1B1, No Template

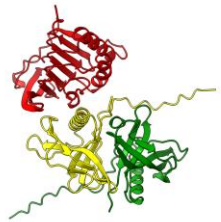
A1B1

AF3 ranking score: 0.35



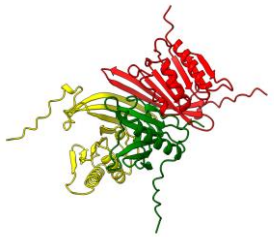
A2B1

AF3 ranking score: 0.24



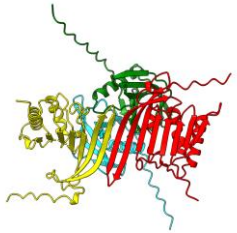
A2B1

AF3 ranking score: 0.34



A2B2

AF3 ranking score: 0.23



Why is it harder to predict stoichiometry of hetero-multimers?

- **More combinations**
- **Less symmetry**
- **Fewer templates**
- **Ambiguity in compatible stoichiometries (**A1B1 and A2B2)

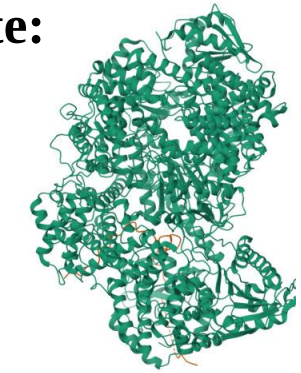
H0220, A1B4

Using templates to correct AF3 predictions

Target	True Stoichiometry	Candidates	Stoichiometry selected by max ranking score	Prediction of MULTICOM
H0220	A1B4	A1B1,A1B2,A1B3,A1B4,A1B5	A1B1	**A1B4**

Subunit: A B

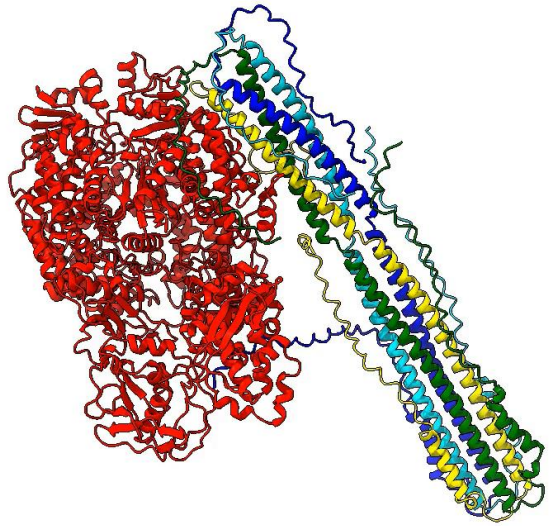
Template:



PDB code: 6UEB 8B8B

A1B1 B4

53% correctness



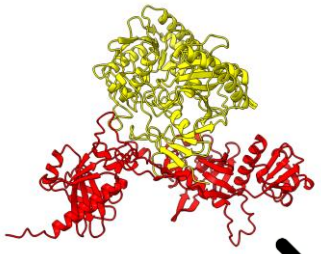
Predicted A1B4 Model

H0267, A2B2, Under-prediction (Tetramer to Dimer Failure)

Target	True Stoichiometry	Candidates	Stoichiometry selected by max ranking score	Stoichiometry by template	Prediction of MULTICOM
H0267	A2B2	A1B1,A1B2,A2B1,A2B2,A4B4	A1B1	A1B1	A1B1

A1B1

AF3 ranking score: 0.88

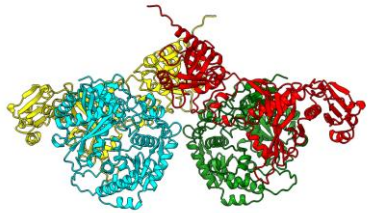


Predicted ✗

Dimer

A2B2

AF3 ranking score: 0.65

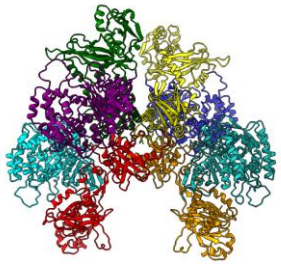


True

Tetramer

A4B4

AF3 ranking score: 0.71

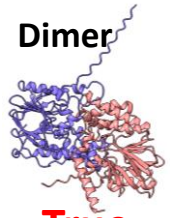


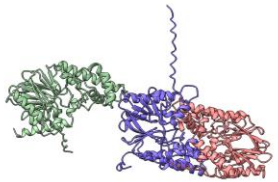
Octamer

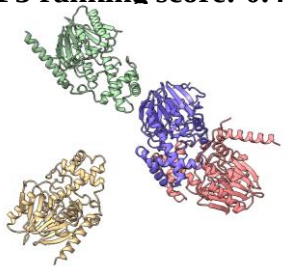
Dimer VS tetramer (compatible, confused)

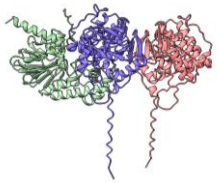
7% correctness

H0208, A1B1, Over-prediction (Dimer to Tetramer Failure)

A1B1
AF3 ranking score: 0.96
Dimer

True

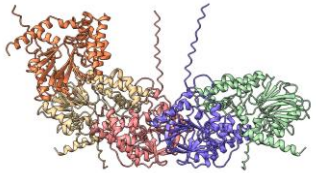
A1B2
AF3 ranking score: 0.75


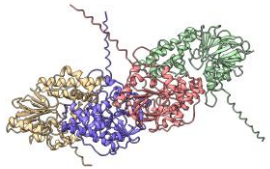
A1B3
AF3 ranking score: 0.46


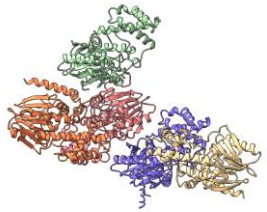
A2B1
AF3 ranking score: 0.88


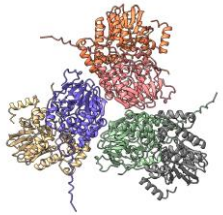
A2B2
AF3 ranking score: 0.89
Tetramer

Predicted

A2B3
AF3 ranking score: 0.61


A3B1
AF3 ranking score: 0.78


A3B2
AF3 ranking score: 0.54


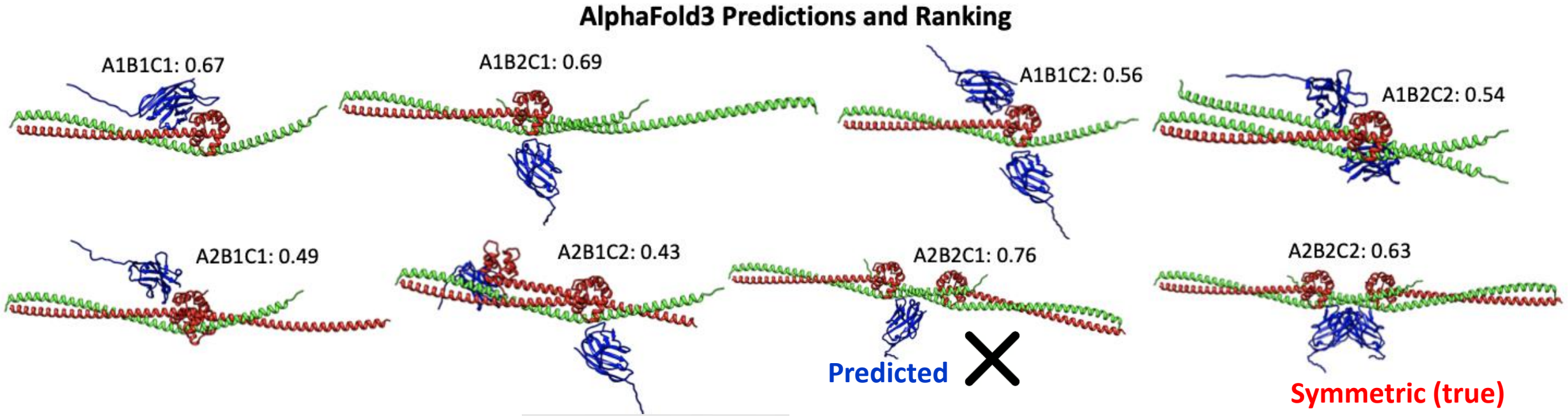
A3B3
AF3 ranking score: 0.48


Dimer VS tetramer (compatible, confused)

Incorrectly interpret templates

66% correctness

H0244, A2B2C2, Not Consider Symmetry Enough



9% correctness

H0265, A9B18, Filament, Too Big, Failed to Propose Stoichiometry

A1B1

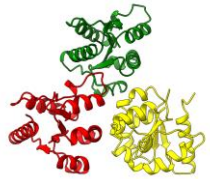
AF3 ranking score: 0.74

A2B1

AF3 ranking score: 0.44

A2B1

AF3 ranking score: 0.46

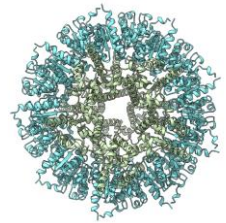


A2B2

AF3 ranking score: 0.5

A9B18

AF3 ranking score: 0.26



AF3 Model

MULTICOM Predicted (worse)

MULTICOM_human Predicted (better)

0% correctness

The only one all CASP16 predictors failed!

True Structure

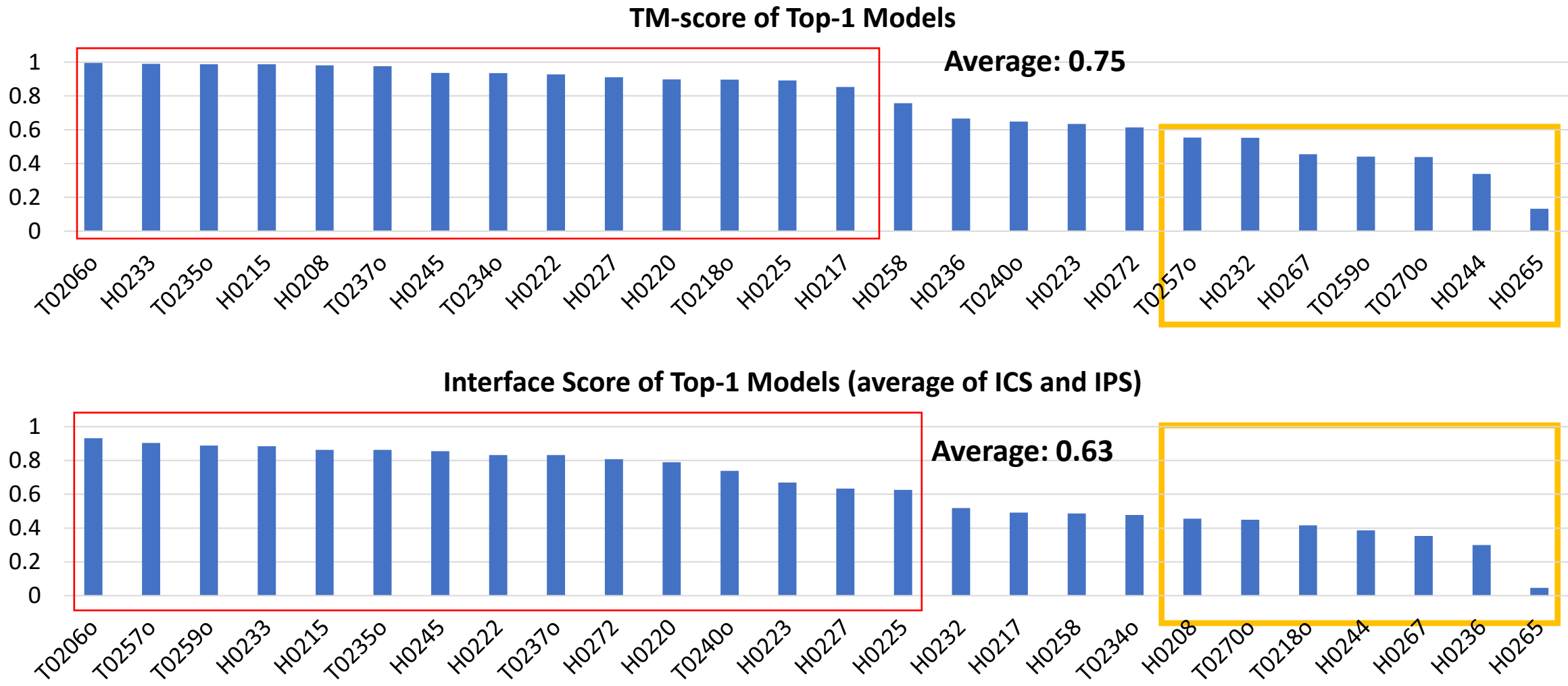
What Went Right?

- Templates for proposing candidate stoichiometries
- AF3 ranking score for selecting stoichiometries (partially successful)
- Combined template-based and AF3-based prediction

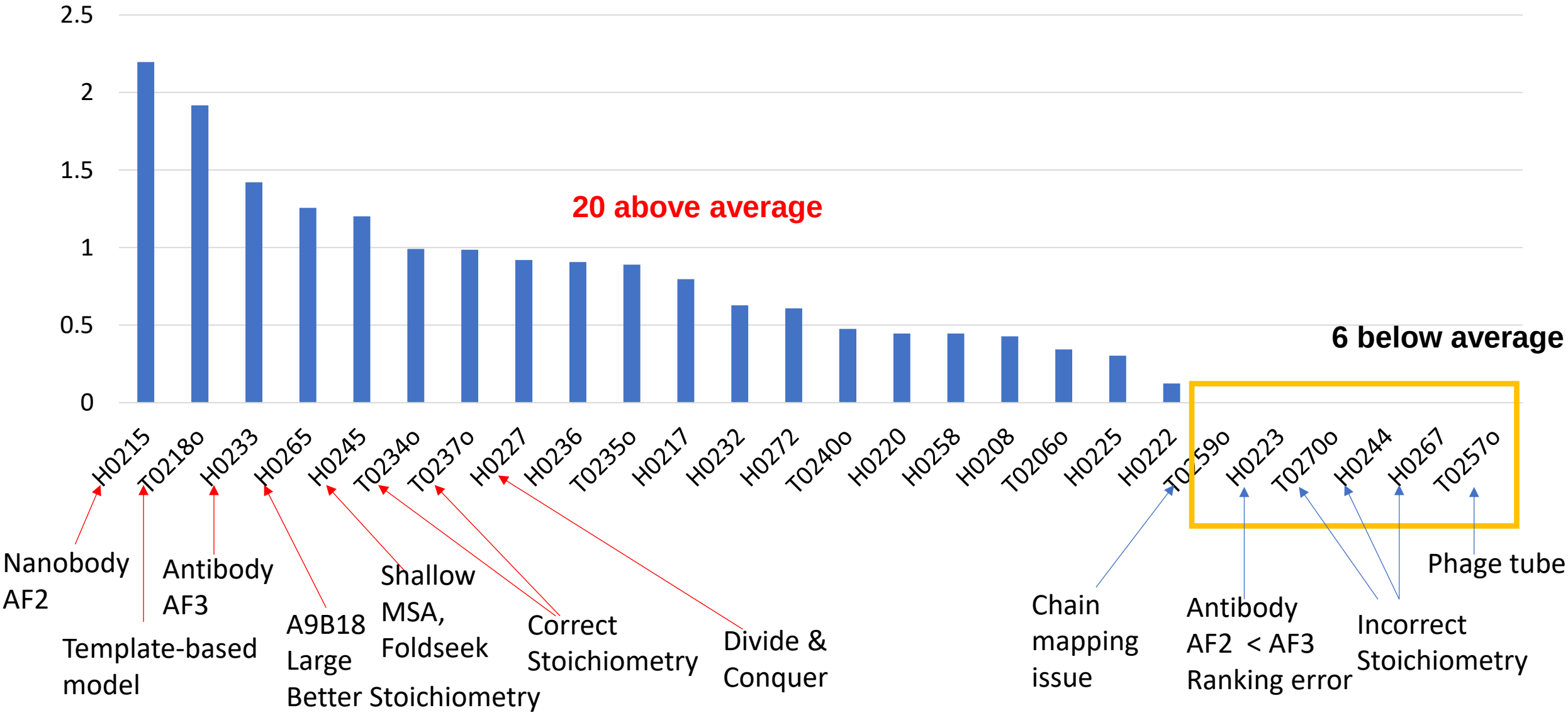
What Went Wrong?

- Failed to resolve the ambiguity in some compatible stoichiometries (e.g., dimer VS tetramer)
- Did not consider symmetry enough
- Constrained by protein size limit set by AlphaFold

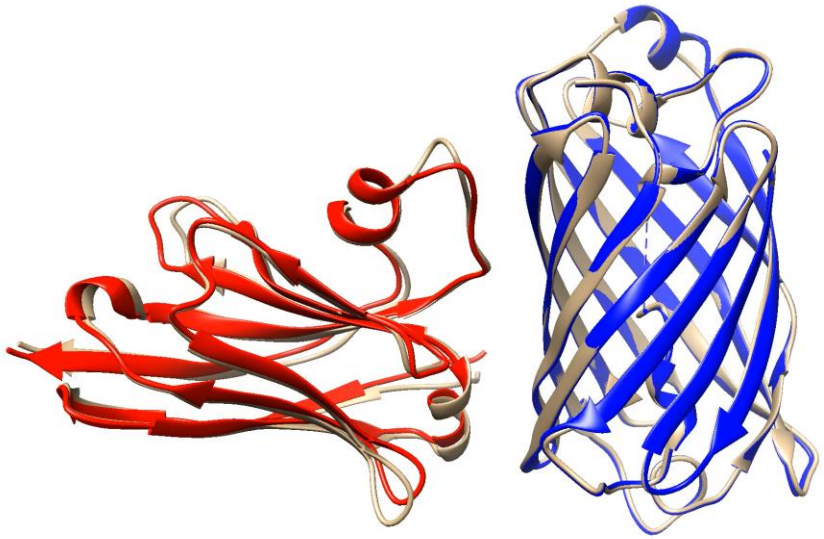
TM-Scores and Interface Scores of MULTICOM_human in Phase 0 (26 Targets)



Z-scores of TM-scores of Top-1 Models of MULTICOM_human on 26 Targets



H0215, A1B1, Nanobody Complex, Hard



AF2 model
TM-score: 0.99
ICS: 0.87
IPS: 0.86
Z-score: 2.2

AF2 better than (>) AF3

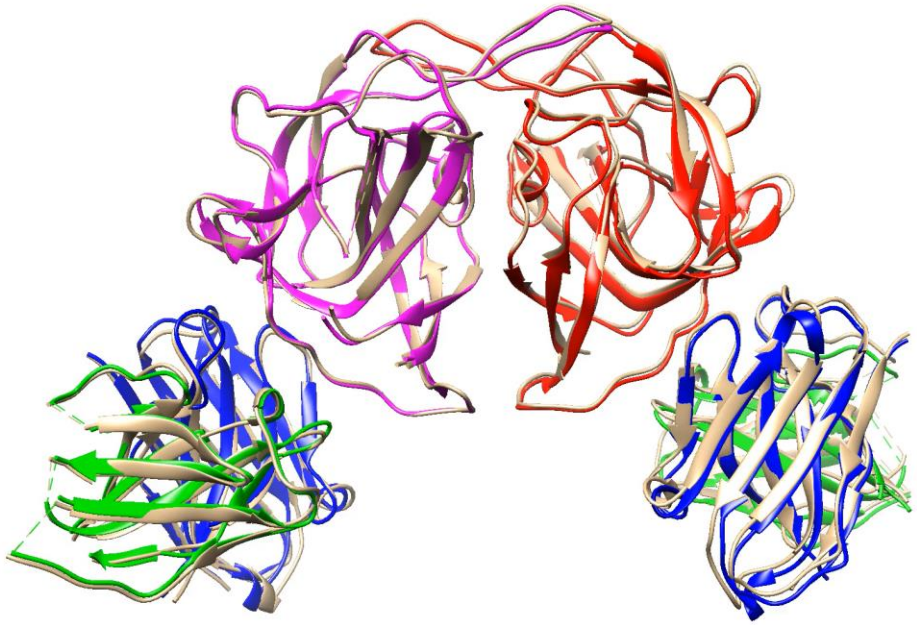
Red & blue: Chain A & B of MULTICOM_human model 1

Brown: true structure

Parameters: 21 cycles, no template, or FoldSeek MSA, 1000 models

H0233, A2B2C2, Antibody, Hard

Fab 3H4 complex, virus capsid protein



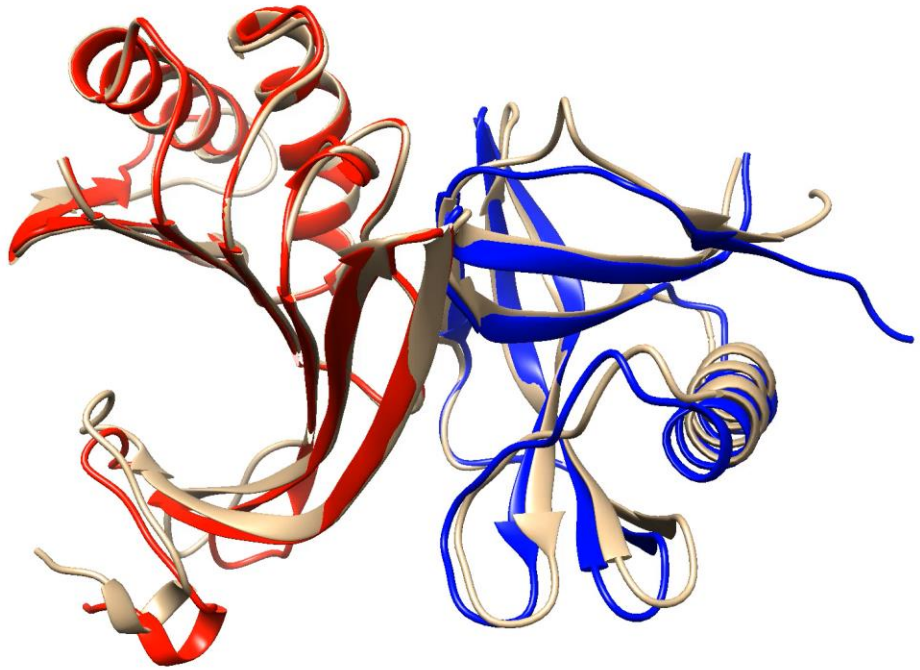
AF3 model
TM-score: 0.99
ICS: 0.87
IPS: 0.90
Z-score: 1.4

Selected from 500 AF3 models
AF3 better than (>) AF2

A2 (Red & magenta) B2 (Blue) C2 (Green) of MULTICOM_human model 1

Brown: true structure

H0245, A1B1, FUNComplex, Shallow MSA, Hard



AF2 model
TM-score: 0.94
ICS: 0.86
IPS: 0.85
Z-score: 1.20

FoldSeek MSA
Selected from 1000 AF2 models

Red & blue: Chains A and B of MULTICOM_human model 1

Brown: true structure

T0234o, A3, Better Stoichiometry Prediction

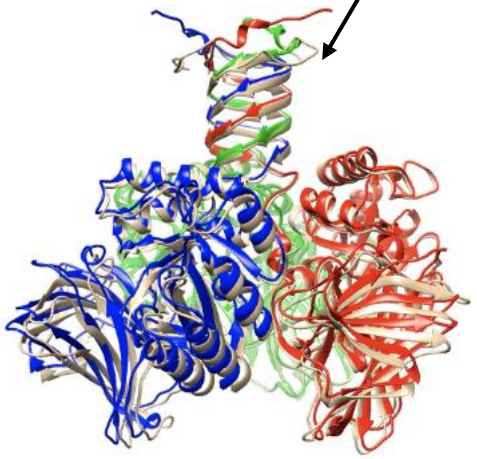
(only 34% groups are correct in stoichiometry)

Model 1
AF3 model
TM-score: 0.93
ICS: 0.26
IPS: 0.7
Z-score: 0.99

Red & blue & green: Chains A, B, C of MULTICOM_human Model 1

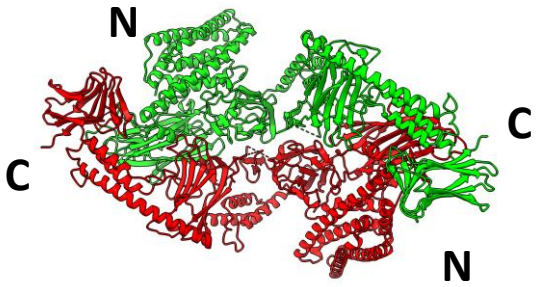
Brown: true structure

Beta tube predicted correctly!

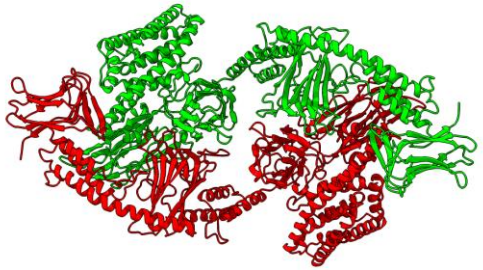


Model 3
AF2 + DeepMSA MSA
TM-score: 0.97
Best among all CASP16 models
Selected by average of GATE and AF confidence score

T0218o, A2, Template-based + AF3 Model

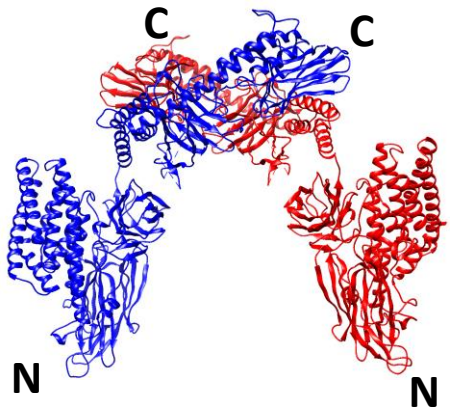


True Structure



Template 4W8J

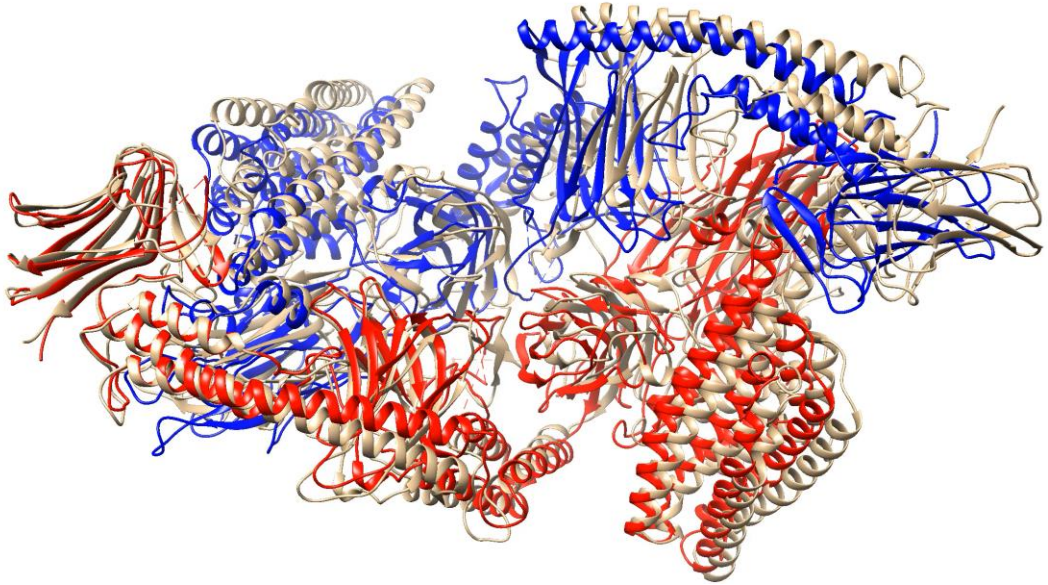
C-C Interaction (incorrect)



A top ranked AF3 model

Red & Blue: Chains A and B of MULTICOM_human model

Brown: true structure



Template-based + AF3 model

TM-score: 0.90

ICS: 0.40

IPS: 0.43

Z-score: 1.92

H0227, A1B6, 5689 Residues, Divide and Conquer

+ =

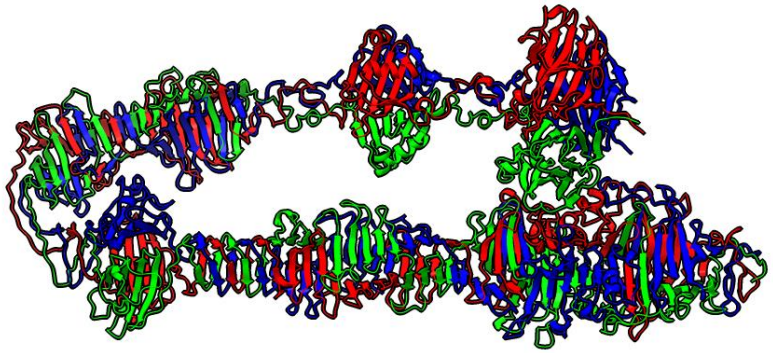
AF3-predicted A1B6 model
Subunit B: Residue 1-745

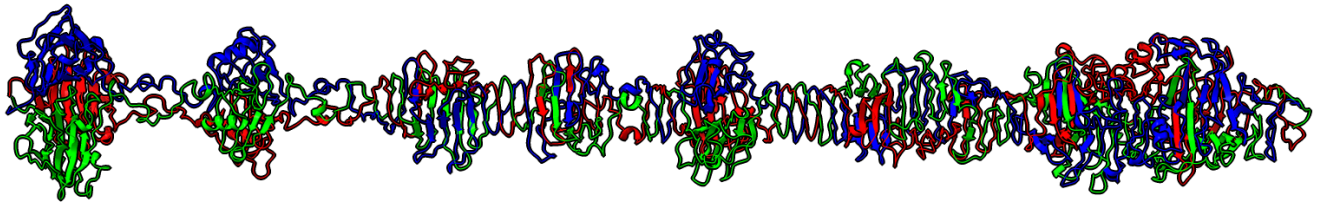
AF3-predicted B6 model
Subunit B: Residue 390 - 877

Combined A1B6 model
TM-score: 0.91
ICS: 0.59
IPS: 0.68
Z-score: 0.92

T0257o, A3, Tube vs Globular, Failure

Enterobacteria phage T5





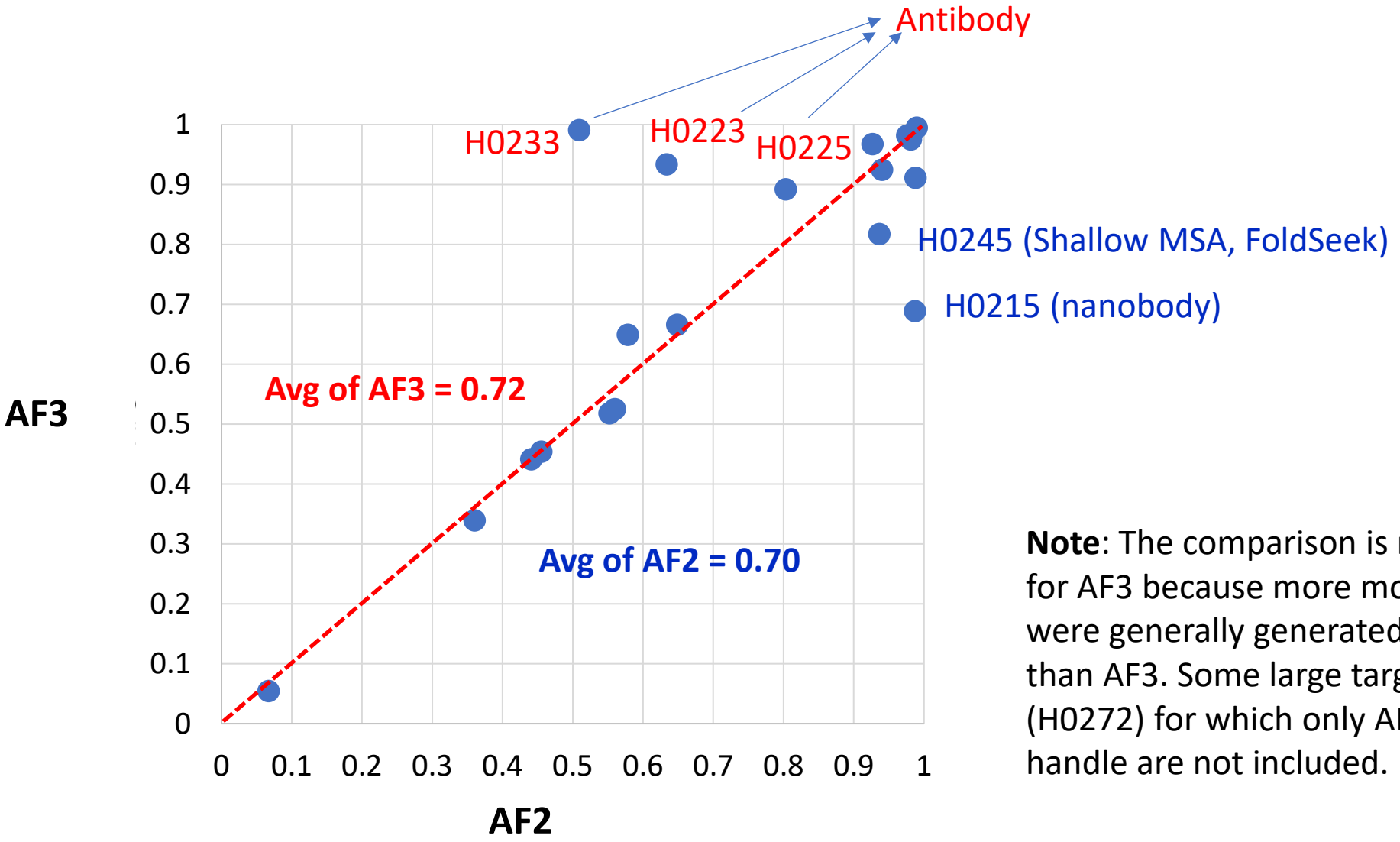
MULTICOM_human model (AF3)
TM-score: 0.55
ICS: 0.87
IPS: 0.94
Z-score: 0

True structure

More Globular, Hallucination of AF3

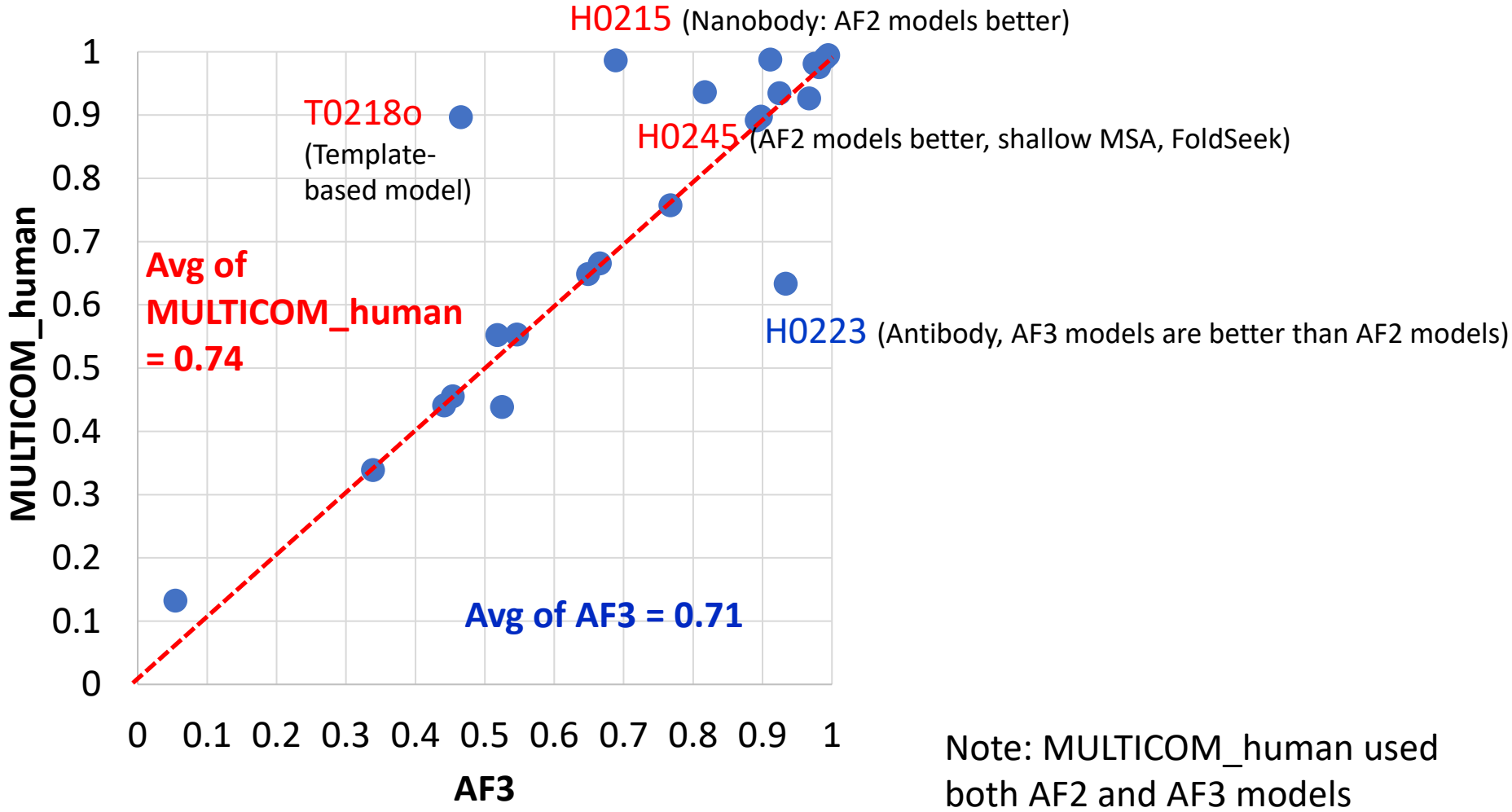
Straight Tube

AF3 Outperformed AF2 on 19 Common Targets in Phase 0 (top-1 model)



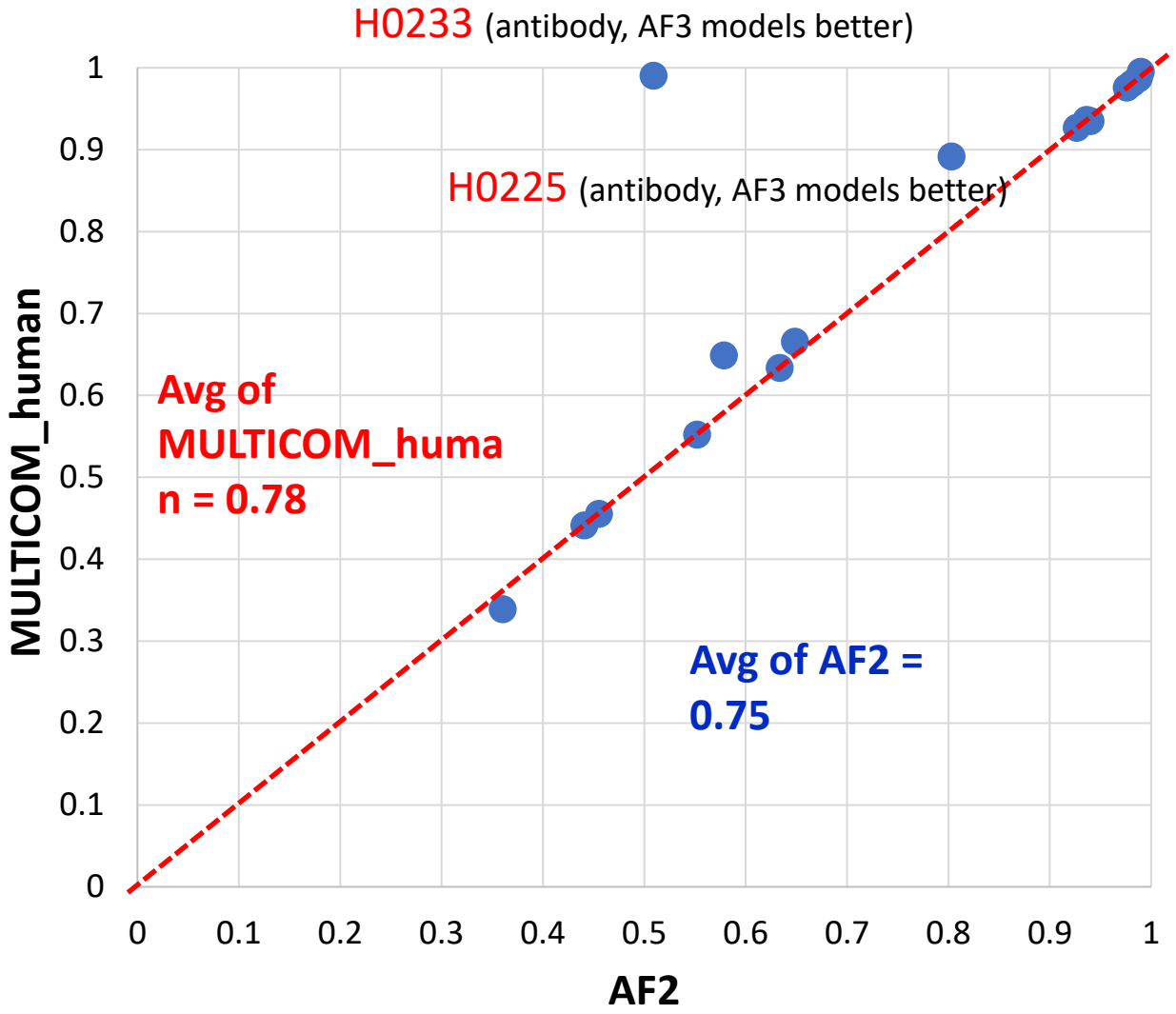
Note: The comparison is not fair for AF3 because more models were generally generated by AF2 than AF3. Some large targets (H0272) for which only AF3 can handle are not included.

MULTICOM_human Outperformed AF3 on 23 Targets in Phase 0 (top-1 model)

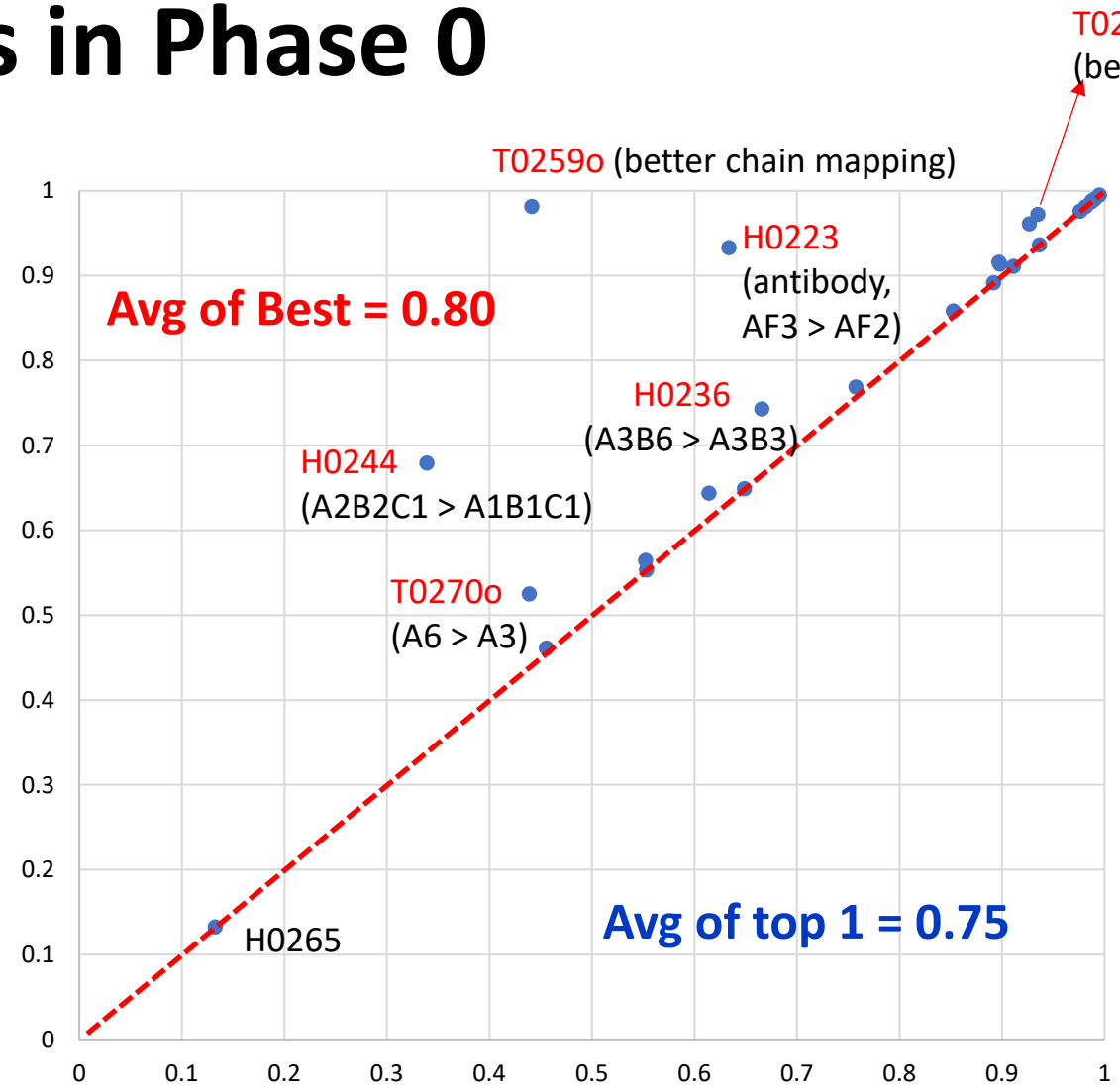


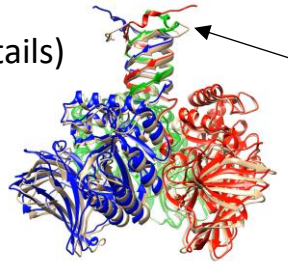
Note: MULTICOM_human used both AF2 and AF3 models

MULTICOM_human Outperformed AF2 on 19 Targets in Phase 0 (top-1 model)



Top 1 vs Best of 5 Models of MULTICOM_human on 26 Targets in Phase 0





Beta tube predicted correctly!

Challenges

- **Select correct stoichiometry as no. 1**
- **Select correct/best model as no. 1**

What Went Right?

- **Used both AlphaFold3 and AlphaFold2** to generate models (antibody, nanobody, large complexes)
- **Used different MSAs and parameters** to generate thousands of models (e.g., FoldSeek MSAs & AFSample parameters)
- **Used multiple model ranking metrics** to select five models and consider alternative and diverse models when uncertain
- **Dealt with the failure and limitation of AlphaFold** (divide & conquer and templates)

What Went Wrong?

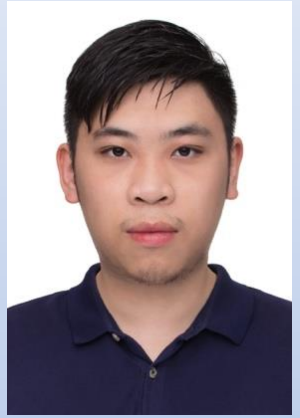
- **Failed to select correct/best (top-1) models for some (hard) targets** when multiple conflicting conformations existed (antibody/nanobody, AF3 vs AF2 models; H0233 and T0234o)

- **Failed to generate good/correct models for several non-globular protein structures** (e.g., T0257o, phage tube; T1240, A3) and **special large targets such as filament** (e.g., H0265/H1265, A9B18; T1295, A8)

Conclusion

- **MULTICOM4 made exciting progress in stoichiometry prediction** by combining template information and AF3 model ranking scores.
- **MULTICOM4 generated correct models for all but several complex targets** using AF3 and AF2 with diverse inputs and parameter settings.
- **Model ranking has been improved but still cannot consistently select best/correct models as top 1** from conflicting conformations.
- **Handling failure and limitation of AlphaFold is useful**.

Acknowledgements



Jian Liu

Pawan Neupane

NATIONAL INSTITUTES OF HEALTH

NSF