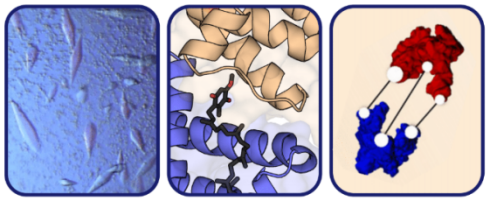
Assessment of FM and FM/TBM modeling in CASP13

Luciano A. Abriata and Matteo Dal Peraro

Laboratory for Biomolecular Modeling – LBM
Institute of Bioengineering, School of Life Sciences
Ecole Polytechnique Fédérale de Lausanne - EPFL



LBM

EPFL
ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

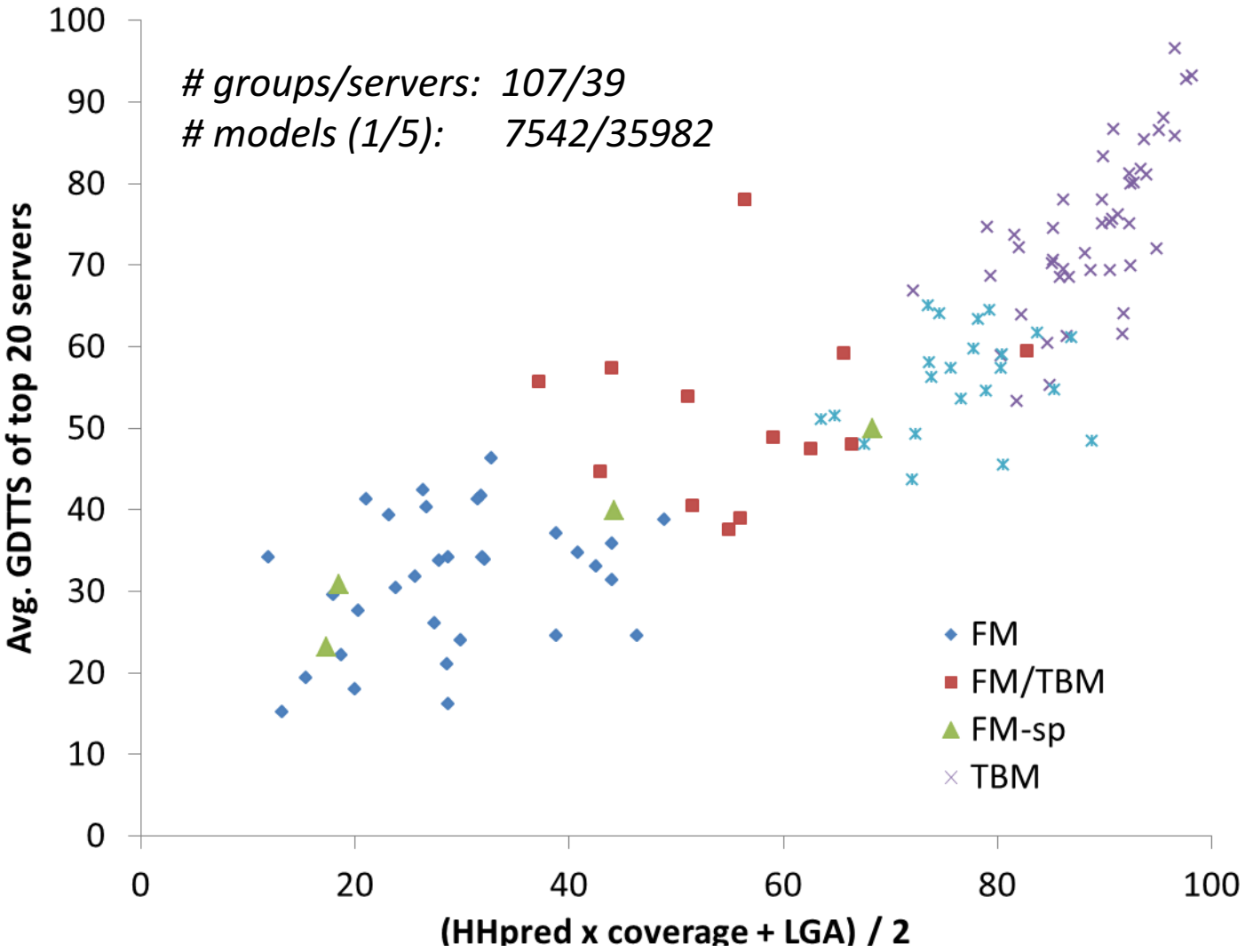
SIB

Acknowledgements

- **CASP13 (and CASP12)** organizers
  - Andriy, John, Torsten, Krzysztof, Maya and **Anna!**
  - *Prediction Center*
  - Previous assessors and all predictors

- Great and constructive experience
- Exciting to witness 2 consecutive huge transformations in protein structure prediction

CASP13 tertiary structure track: 32 FM & 13 FM/TBM (+ 4 FM-special)



Details on classification by Lisa Kinch & Andriy Kryshtafovych

Part 1: EU-specific evaluations

Strategy for EU-specific evaluations

Target EUs, models and tables from the *Prediction Center*

Clustering at 3 Å

~half of the models for initial inspection, represented by top GDTTS models

JS + HTML

Web App for interactive navigation of model clusters: 6 main scores, others available too

Visual Inspection

Designation of best cluster(s) for each EU

Further evaluation of models in best cluster, if worth

Model(s) designated best for each EU

CASP12-like web app: facilitates assessment, and is easily opened to the public

http://lucianoabriata.altervista.org/papersdata/casp12fmassessment/casp12-fm-fmtbm-assessment-3Aclusters.html

Select a target: T0960-D2 (FM) H/S (32.92) Range: 43-126

In these plots correlating different metrics to GDTTS, each dot represents a set of models clustered at 3A RMSD. Click on dots to see a representative model in 3D and compare it to the target.

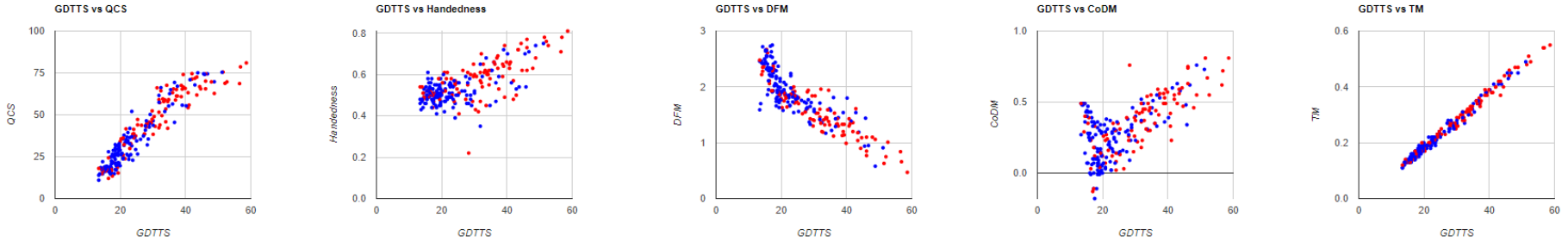


Table summarizing top models

Rank by GDTTS	GDTTS	Metric	Representative Model (hover to see all)
1	58.63	GDTTS	T0960TS322_1-D2 (1) ...
2	56.84	GDTTS	T0960TS457_4-D2 ...
3	56.55	GDTTS	T0960TS055_5-D2 ...
4	52.68	GDTTS	T0960TS366_5-D2 (1) ...
5	52.08	GDTTS	T0960TS086_3-D2 ...
6	51.49	QCS	T0960TS322_3-D2 ...
7	51.19	QCS	T0960TS324_1-D2 (1) ...
19	43.75	QCS	T0960TS145_3-D2 ...
9	48.81	DFM	T0960TS261_2-D2 ...
6	51.49	CoDM	T0960TS322_3-D2 ...
85	28.27	CoDM	T0960TS068_5-D2 (1) ...
13	46.13	CoDM	T0960TS192_2-D2 (1) ...
14	46.13	Handed	T0960TS322_5-D2 ...

(1) Contains at least one model submitted as #1

Pymol script to compare filtered models to target

[Download script]

Run from PyMOL's File menu or by lanching PyMOL with:

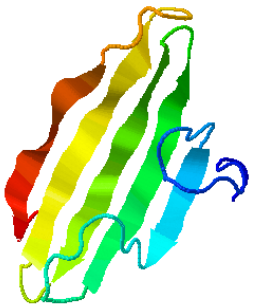
```
pymol nameofscript.pml
```

```
load T0960-D2.pdb
select T0960-D2
spectrum selection=sele

load T0960TS322_1-D2
select T0960TS322_1-D2
spectrum selection=sele
align T0960TS322_1-D2,T0960-D2

load T0960TS457_4-D2
select T0960TS457_4-D2
spectrum selection=sele
align T0960TS457_4-D2,T0960-D2

load T0960TS055_5-D2
select T0960TS055_5-D2
spectrum selection=sele
align T0960TS055_5-D2,T0960-D2

load T0960TS366_5-D2
```

Interactive 3D visualization of **Target** synchronized to model

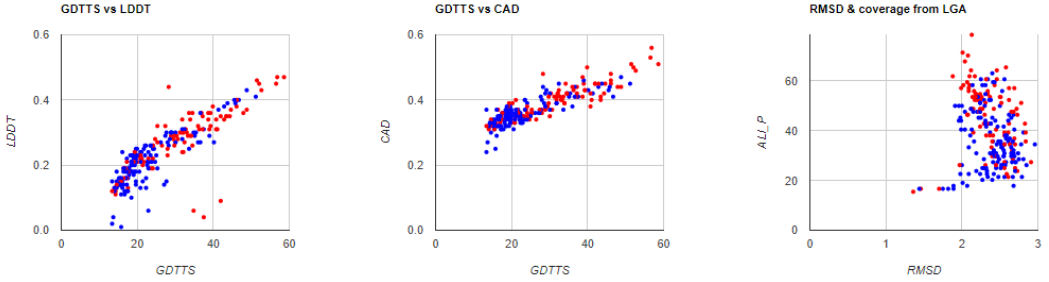
JSmol

Interactive 3D visualization of **Model** synchronized to target



JSmol

Extended Information

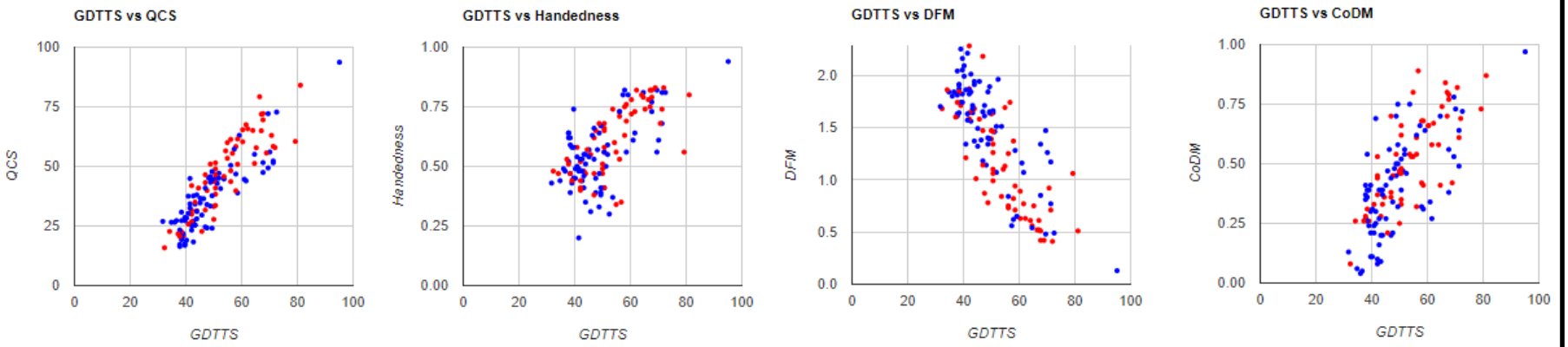


NEW: **more scores**, **show servers in distinct color**, and **built auxiliary web apps also for models clustered at 1 Å and for analysis with no splitting**

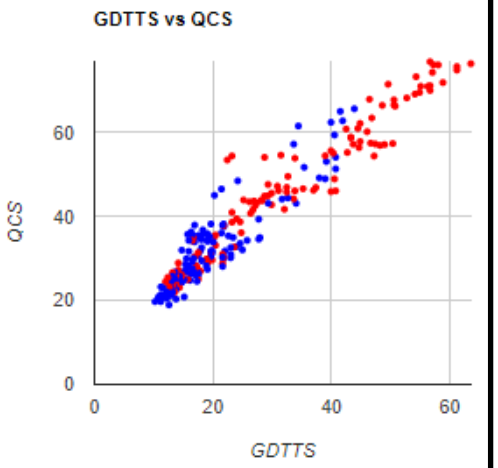
Part 1: EU-specific evaluations

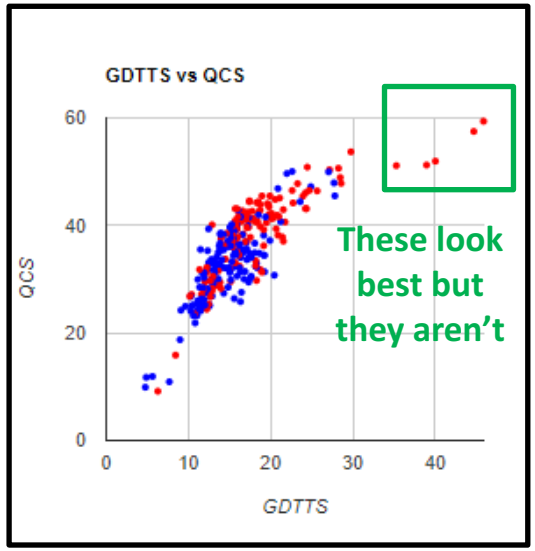
Examples of correlation plots

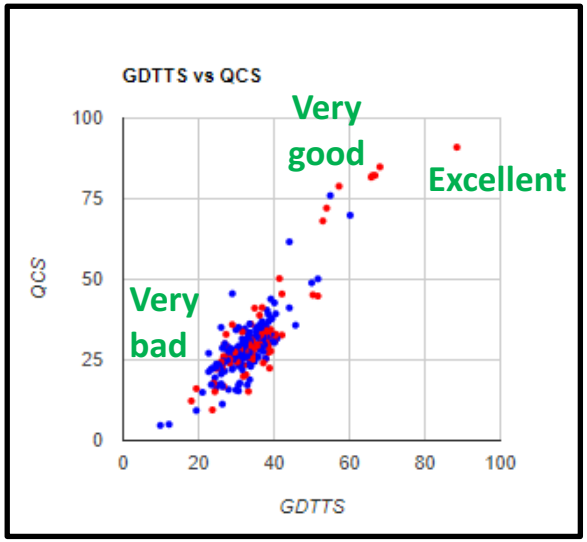
Clear best



Unclear best, visual inspection of multiple models





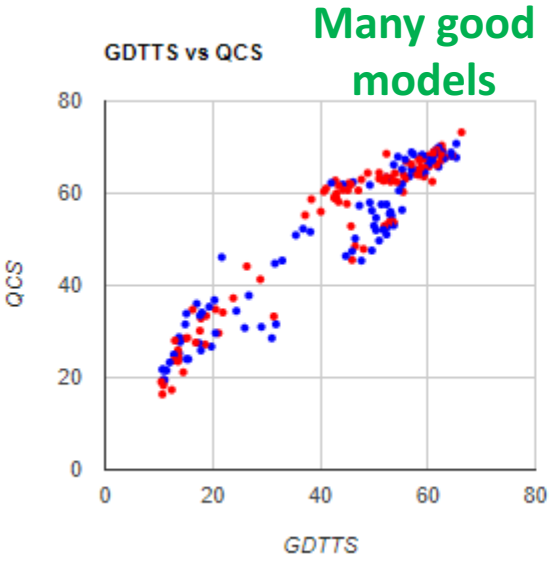


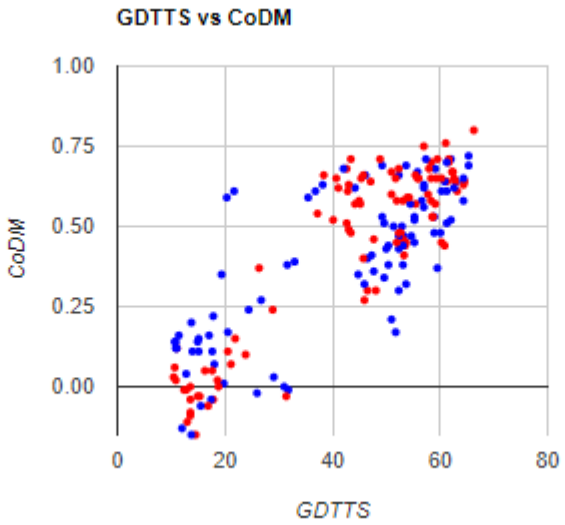
→ GDTTS & QCS turn out to be the two most informative scores, in our experience

* For QCS see Cong et al *Bioinformatics* 2011

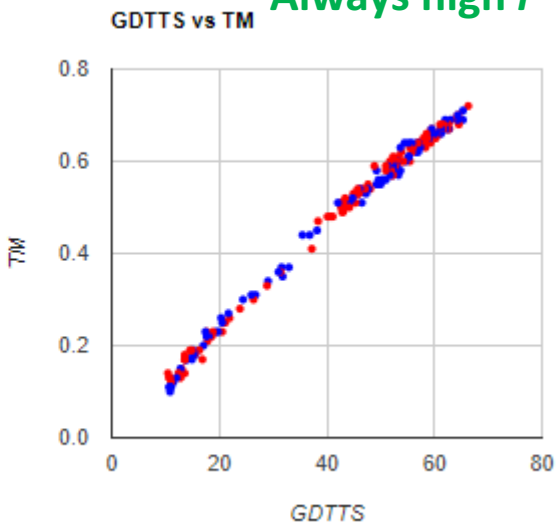
Examples of correlation plots

Many good models

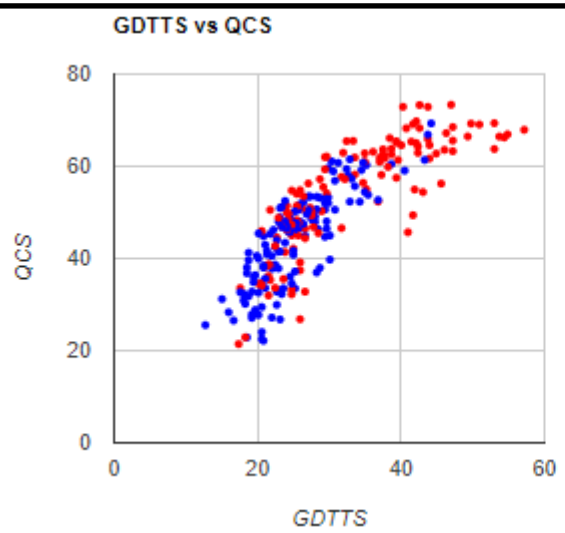


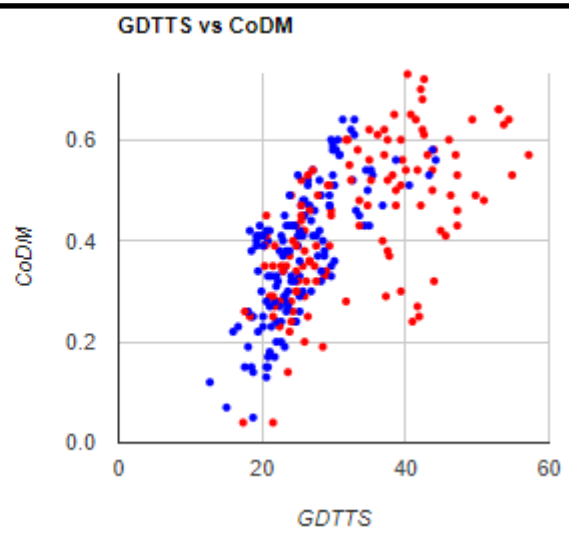


Always high *r*



Different scores propose different best models





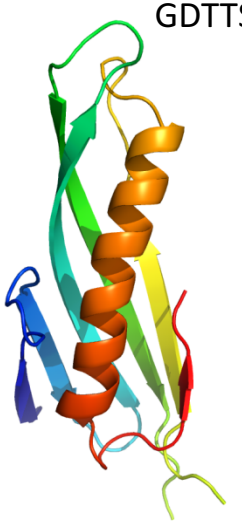
GDTTS & QCS indeed grouped separately in analysis by Olechnovic et al. *Bioinformatics* 2018

Importance of guiding visual assessment by multiple scores

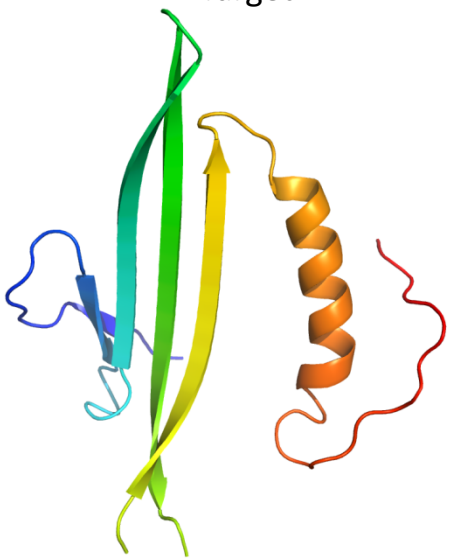
T0991-D1 (FM)

Top GDTTS

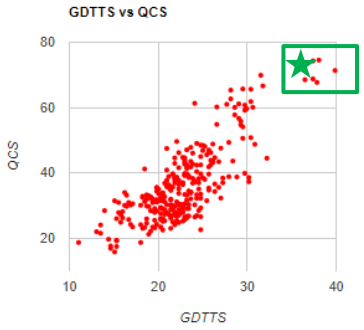
Top QCS, second GDTTS

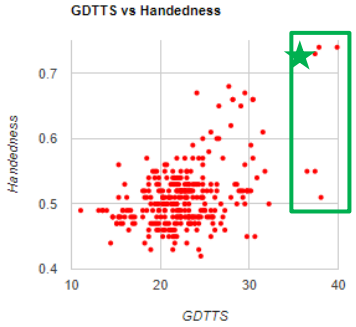


Target



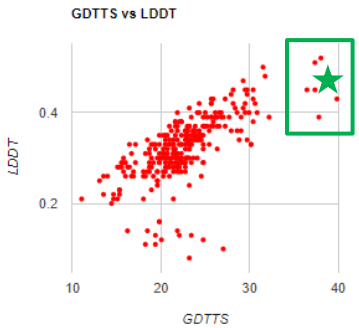
TS366_3 (top by DFM, designated best)

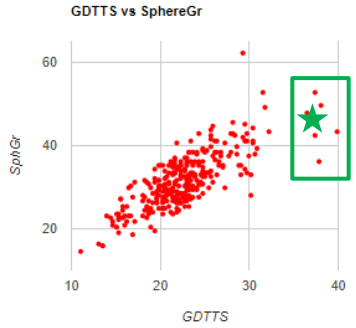


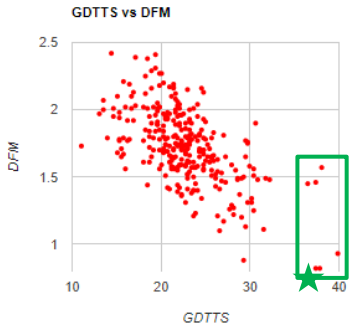


HHscore 13.98
LGA 73.5
Neff/L HHblits 0.01

GDTTS 37.4
QCS 68.8
DFM 0.82





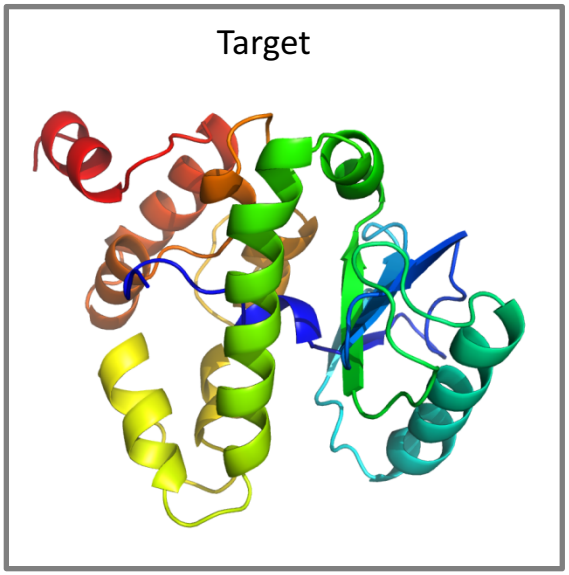


Part 1: EU-specific evaluations

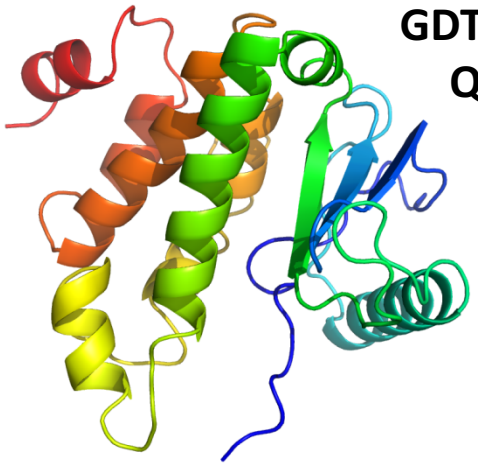
Several very hard targets with folds captured

T0990-D3
(FM)

HHscore 2.76
LGA 23.8
Neff/L HHblits 0.2



TS043_1
(designated best)



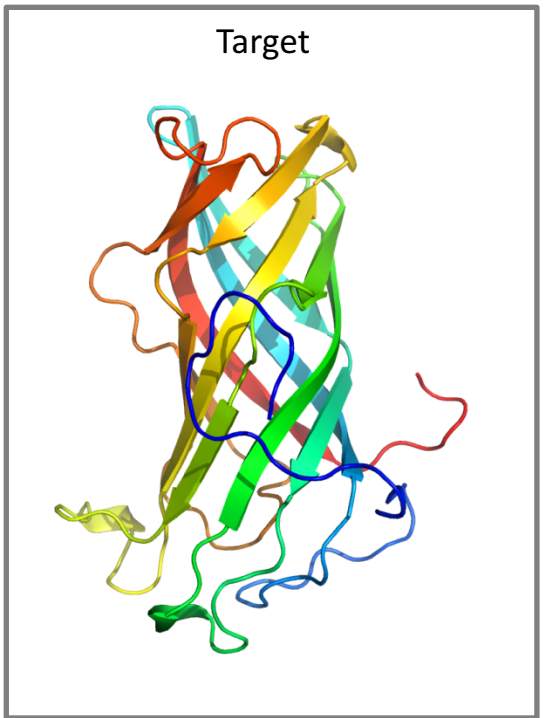
GDTTS 50
QCS 80

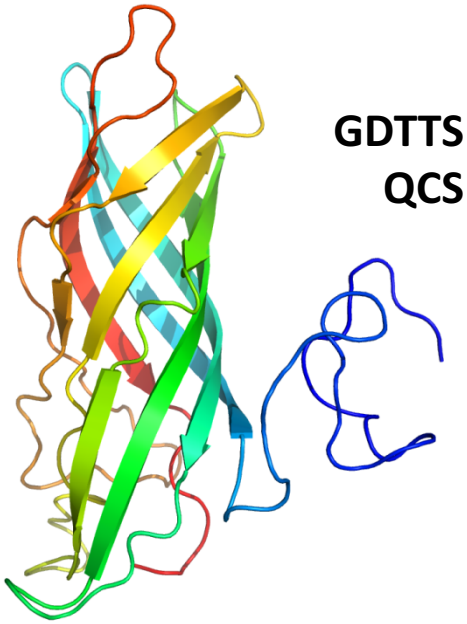
T1010-D1
(FM)

HHscore 0.69

LGA 39.5

Neff/L HHblits 0.07





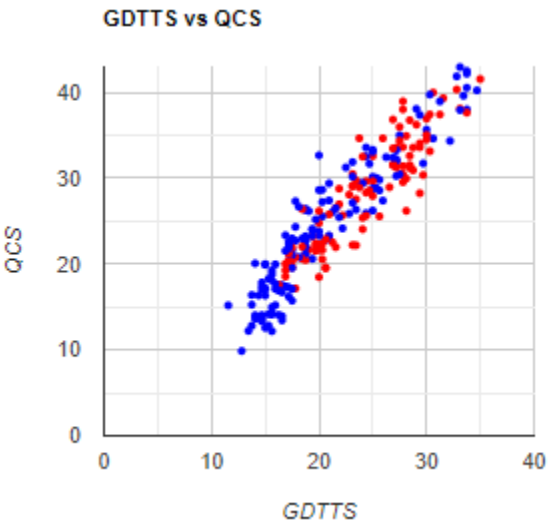
GDTTS 50
QCS 80

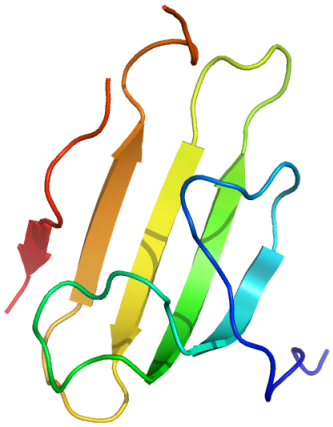
TS117_1

Part 1: EU-specific evaluations

Only two very difficult EUs with no best model

T0981-D2 (FM)

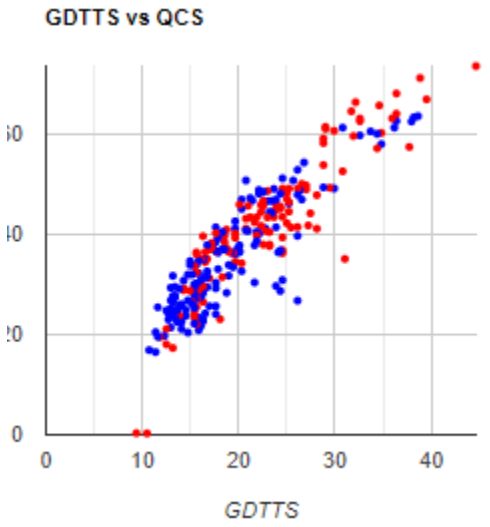




All scores low; here model of highest GDTTS looks reasonable but is missing the last strand which is separated in sequence. And models that are complete are too bad…

HHscore 14 **LGA** 63.8 **Neff/L HHblits** 0.07

T0989-D2 (FM)



Long extended N-terminus and C-terminal beta hairpin, none is well positioned; but the central beta sheet is quite good in some models.

HHscore 4.9 **LGA** 55.1 **Neff/L HHblits** 0.01

Impact of progress in CASP13:

Examples of “FM-special” targets for which full models were very good

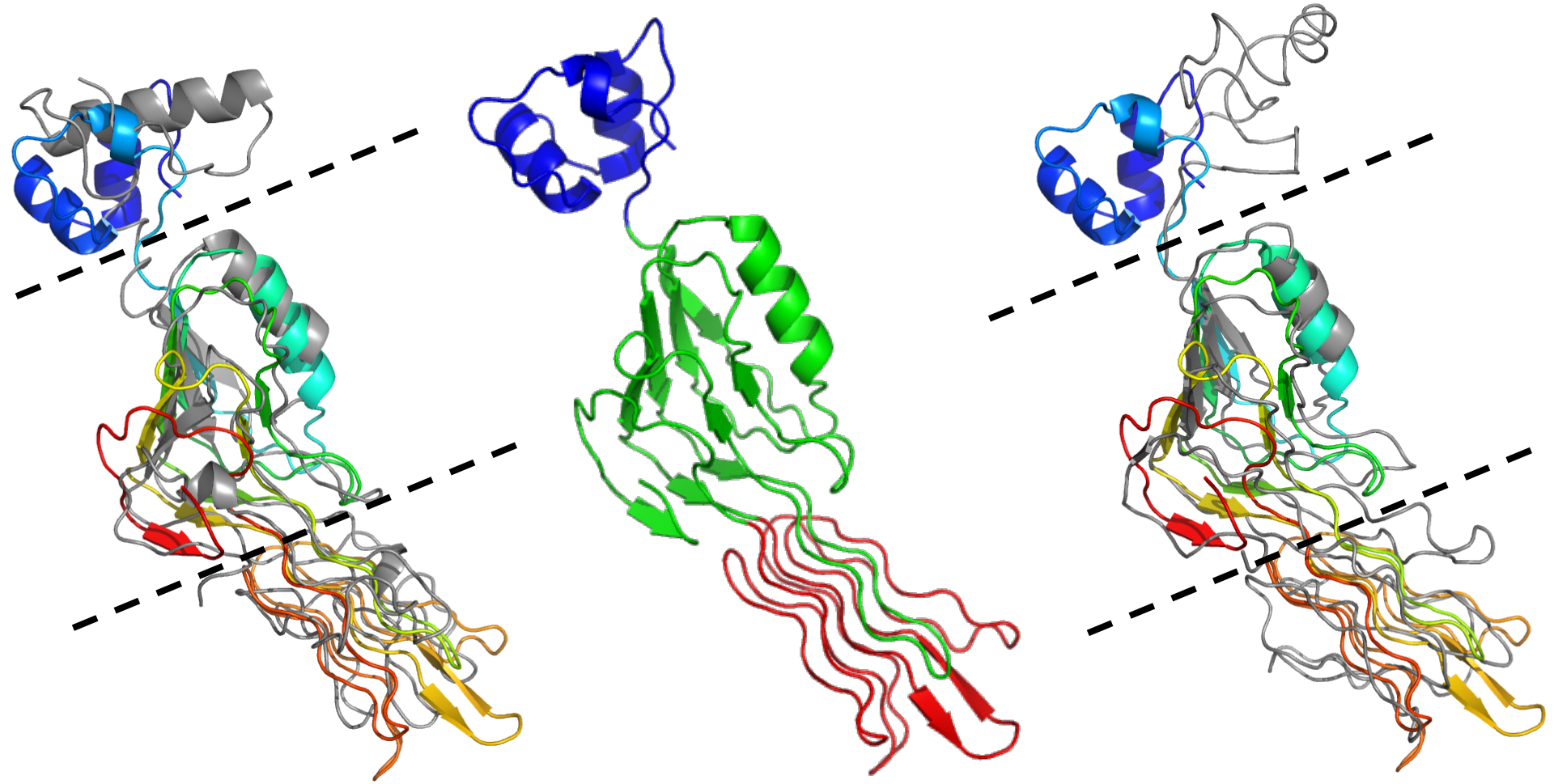
Part 1: EU-specific evaluations

Example: T0953s2 (D1: FM/TBM, D2 & D3: FM)

TS117_4 (Top by TM, 2.53 Å RMSD over 61% of sequence)

Target by EU (D1, D2, D3)

TS224_3 (Top by GDTTS)

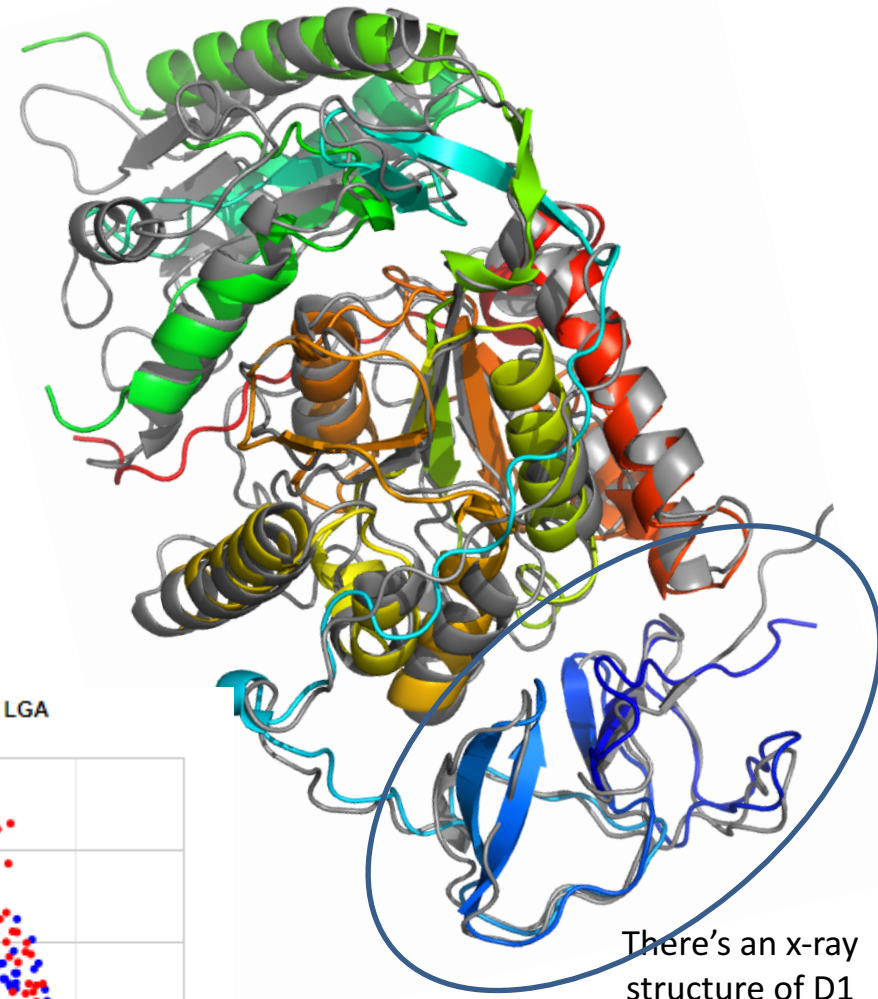


Example: T1000 (D1: TBM not eval., D2:FM)

TS043_1 (Top by GDTTS, scores quite good by all metrics)

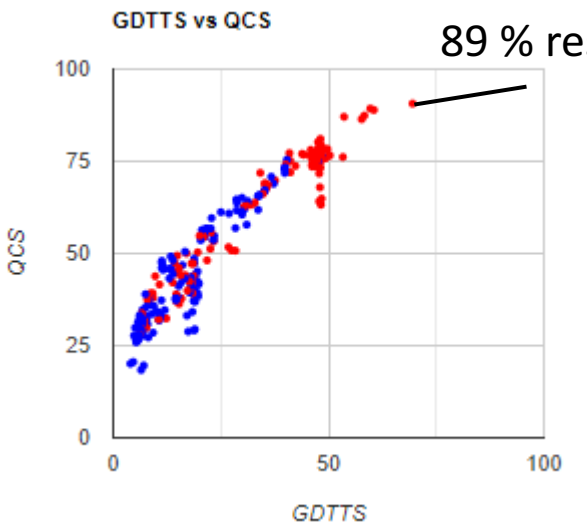
GDTTS 69.5
QCS 90.7

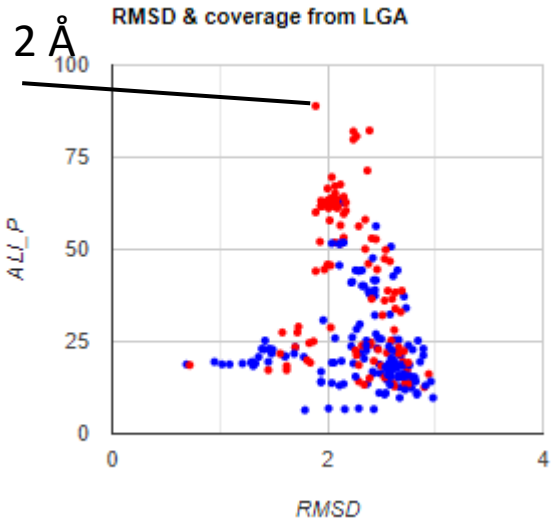
Parts missing in exp target structure



There's an x-ray structure of D1

89 % res < 2 Å

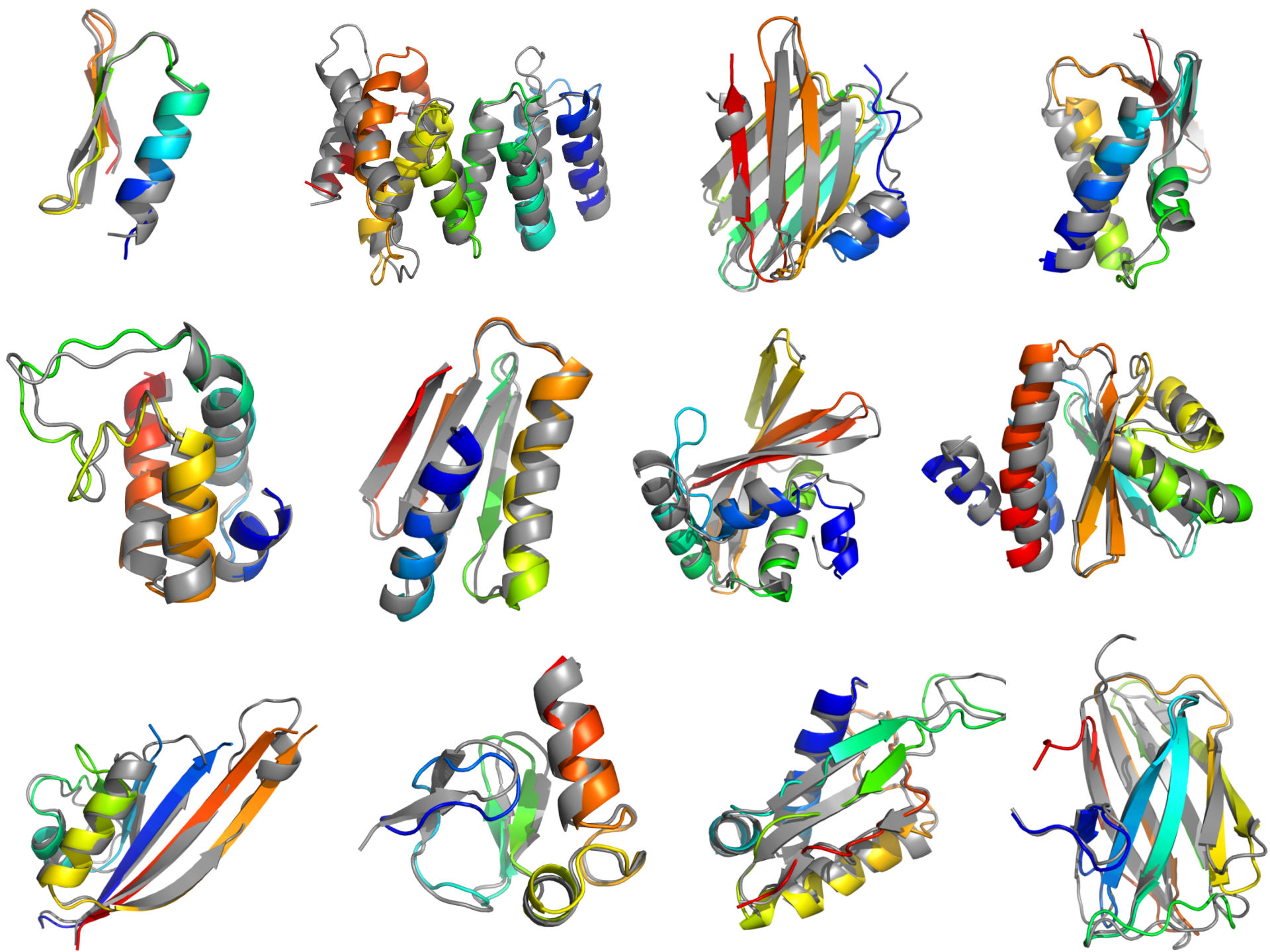




Notable progress in CASP13:

12 hard EUs that reached near atomistic resolution by many groups

Part 1: EU-specific evaluations

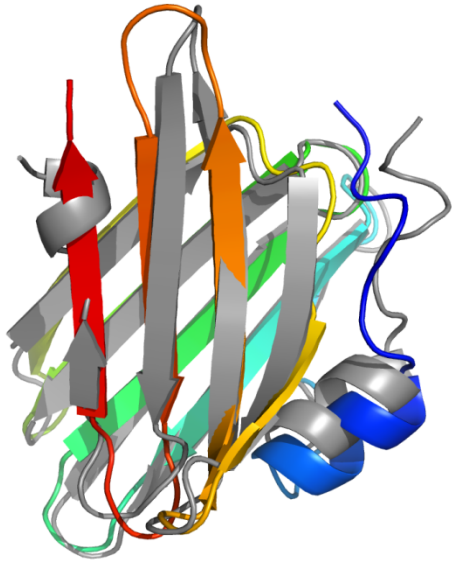


T0968s2-D1 (FM)

TS043_1-D1 (12 models)

2.33 Å over full sequence (115 residues)

HHscore 19
LGA 50
Neff/L HHblits 1.23



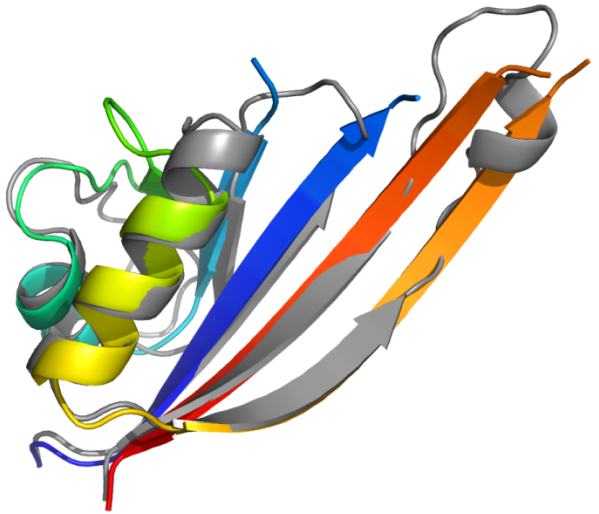
GDTTS 80 & QCS 90

T0970-D1 (FM/TBM)

TS043_2-D1 (5 models plus 4 from TS347)

2.78 Å over 89% of sequence (total 96 residues)

HHscore 17
LGA 67
Neff/L HHblits 1.61



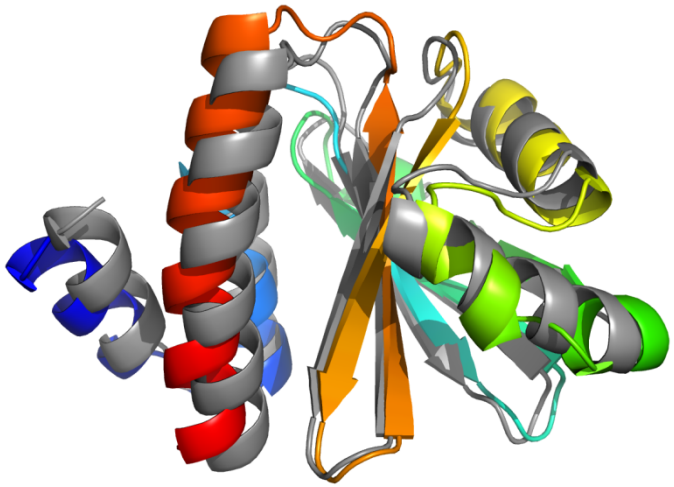
GDTTS 80 & QCS 90

T1001-D1 (FM)

TS222_4-D1 (106 models)

2.32 Å over full sequence (139 residues)

HHscore 11
LGA 55
Neff/L HHblits 0.04



GDTTS 74 & QCS 93

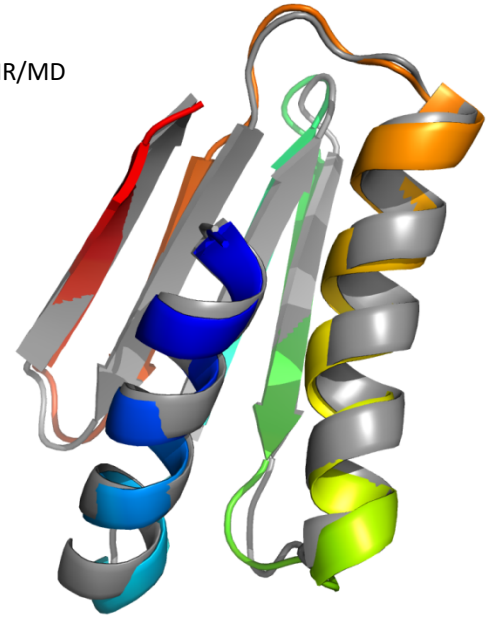
T1008-D1 (FM/TBM)

NMR/MD

TS281_1-D1 (126 models)

1.14 Å over full sequence (77 residues)

HHscore 61
LGA 74
Neff/L HHblits 0.01
GDTTS 91
QCS 95



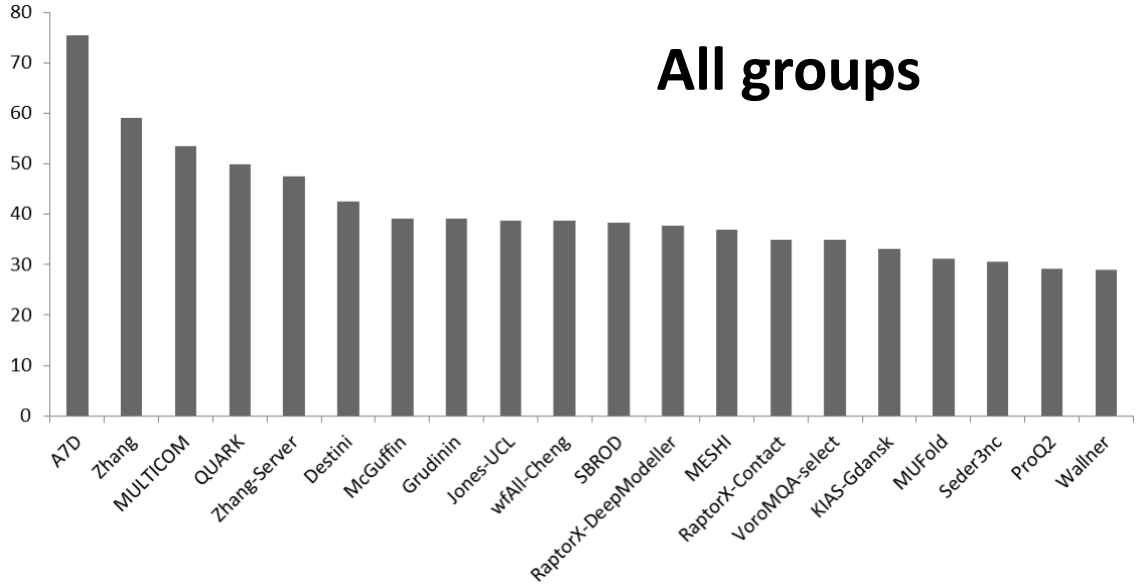
Part 1: EU-specific evaluations

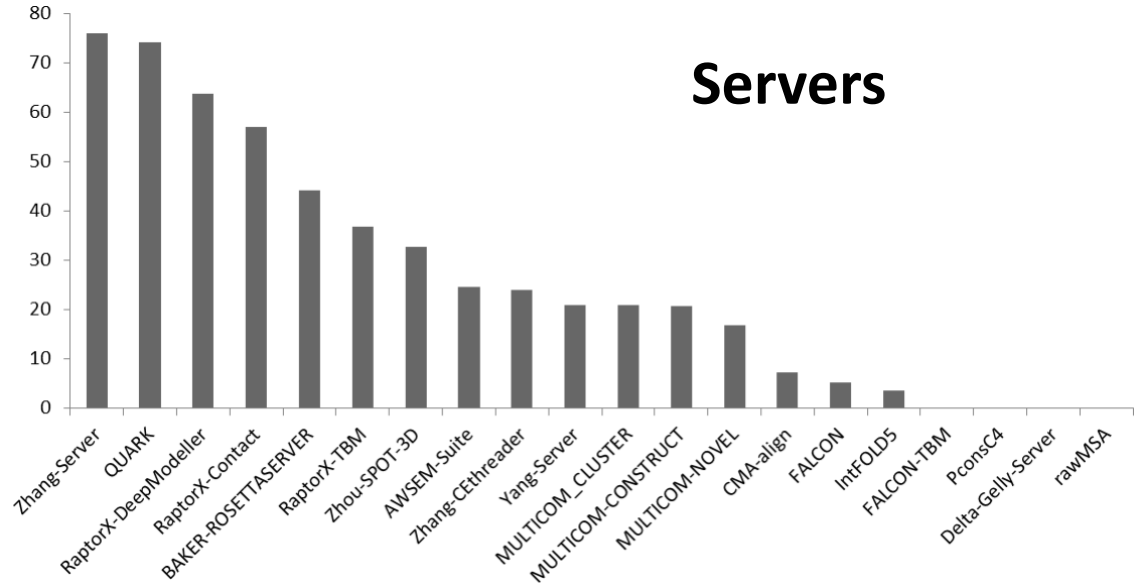
Part 2: Rankings

Ranking based on Z-scores of GDTTS & QCS

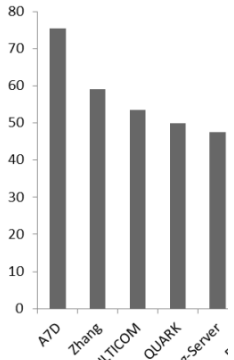
Ranking = **sum Z-scores** combined from **GDTTS & QCS** (as these are by far the two most informative scores to guide visual assessment) on all **models submitted as #1**, for **TBM/FM, FM and FM_sp** target EUs, and considering **sum of Z-score > -2**.

Ranking is very robust: scores with GDTTS only or QCS only return the same top groups.

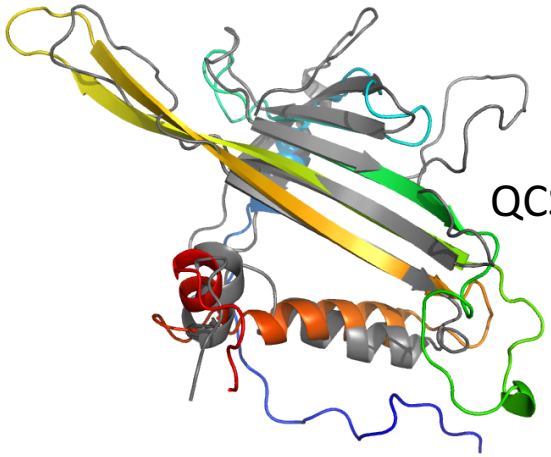




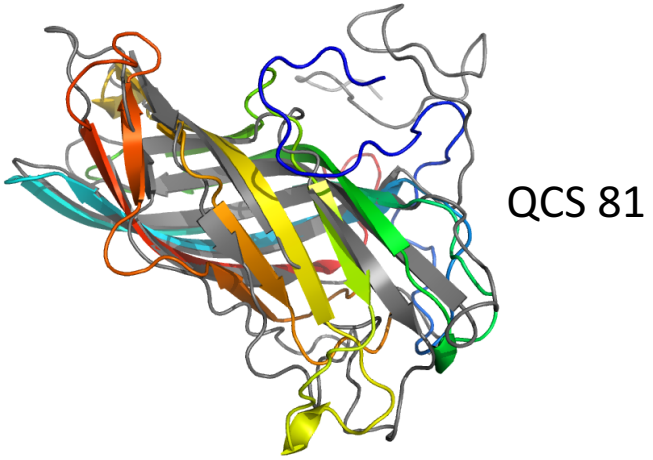
Notable highlights: groups not in top 5 who provided the only best models for some targets (upon visual evaluation)



- **ZHOU-SPOT for T0998-D1:** alone & quite better than runners-up



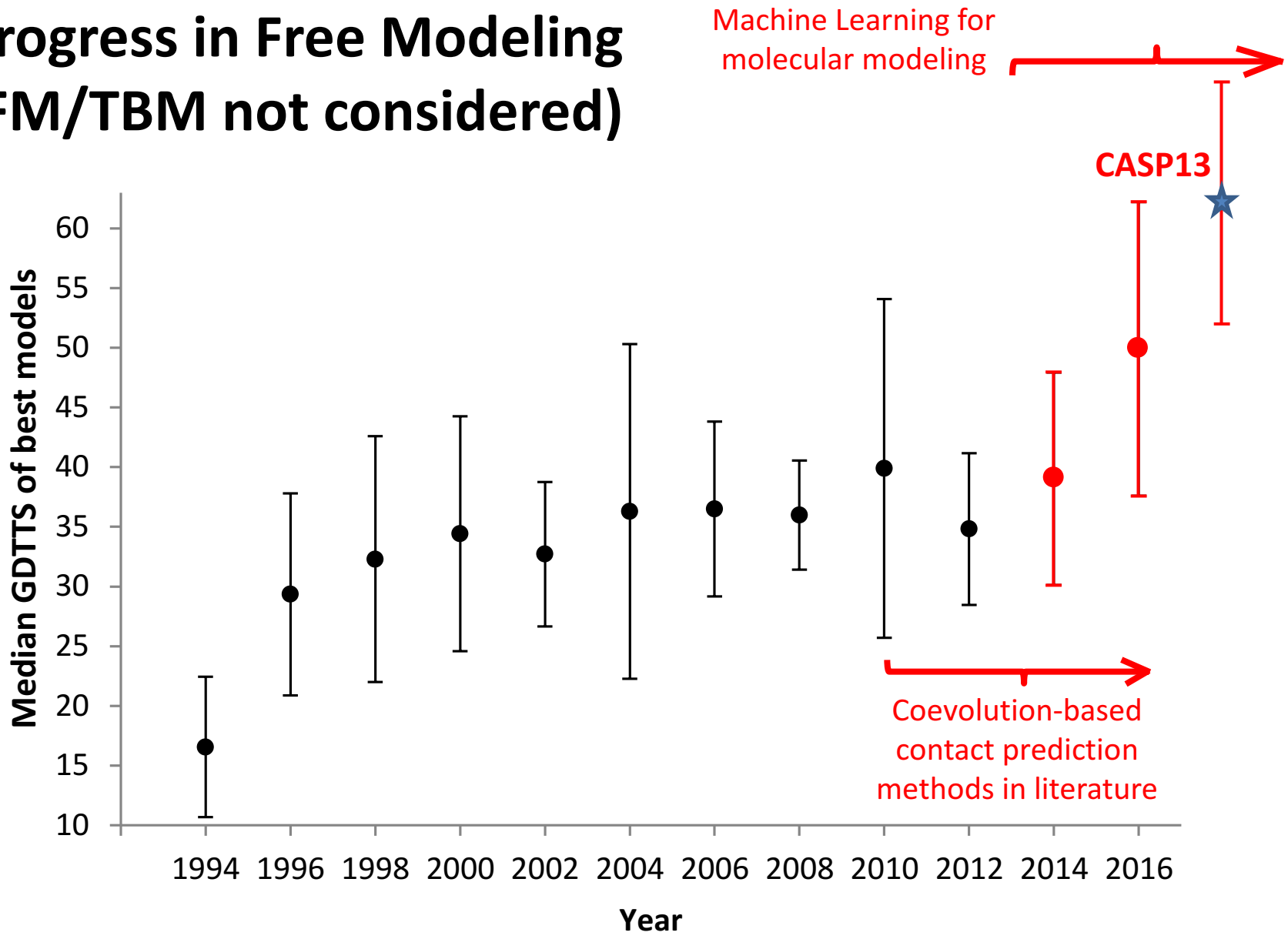
- **Jones-UCL for T1010-D1:** alone & quite better than runners-up
- **RaptorX-DeepModeller for T0949**
- **KIAS-Gdansk for T0957s1-D1**



- **BAKER for T0975-D1**
- **Venclovas for T0991-D1**

Part 3: Progress

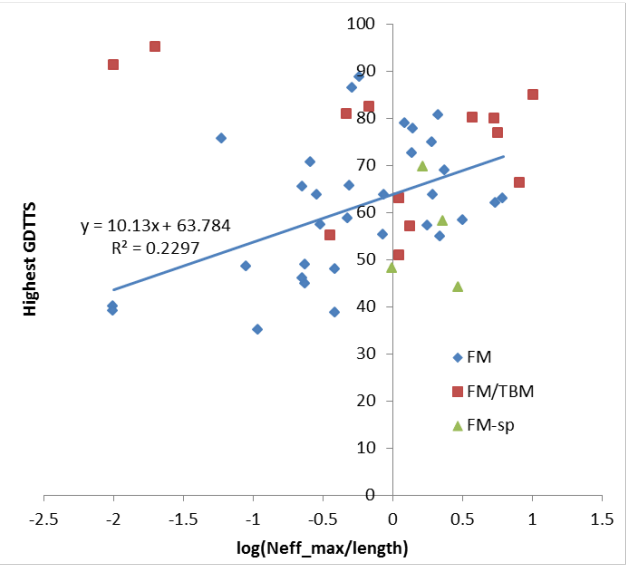
Progress in Free Modeling (FM/TBM not considered)

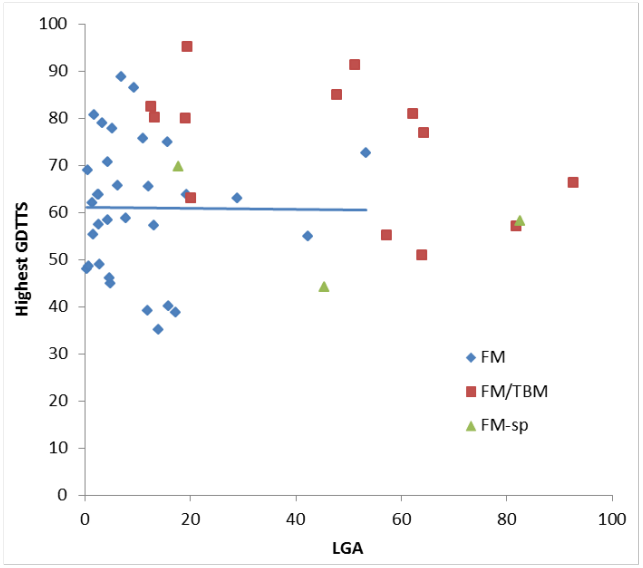


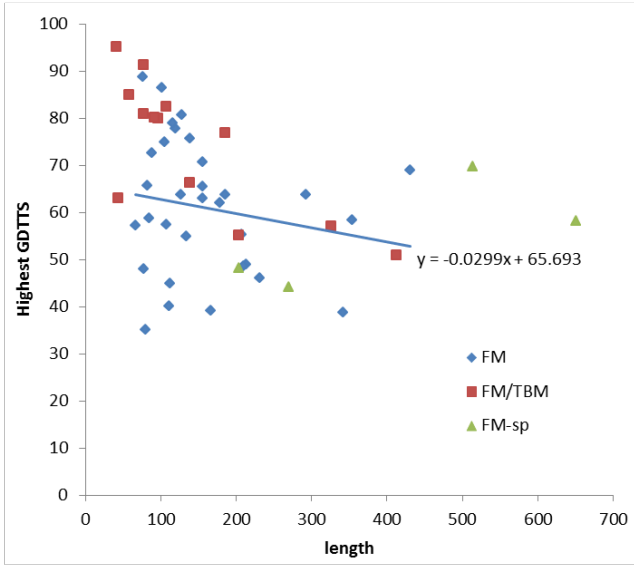
Notes:

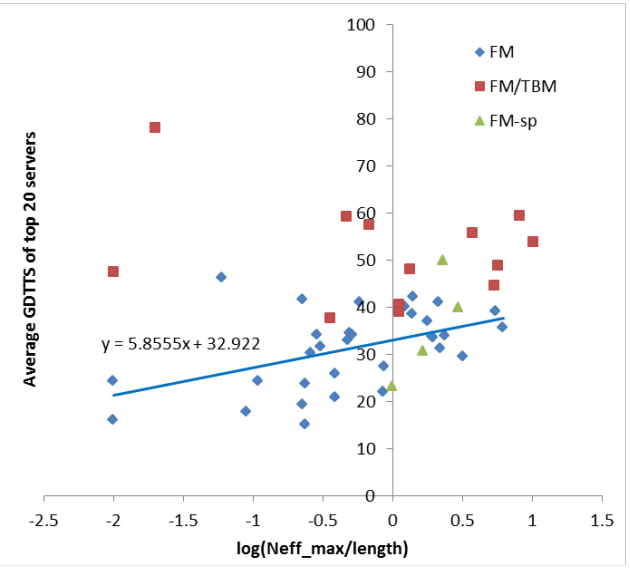
- Exact definition of FM EUs might vary from year to year
- CASP12 and CASP13 EUs of roughly of similar difficulty

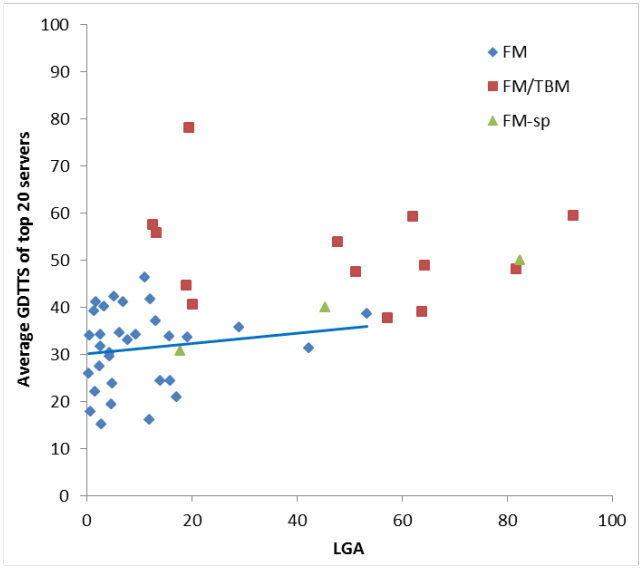
Possible sources of improvement:
alignment depth, existing templates, domain size?

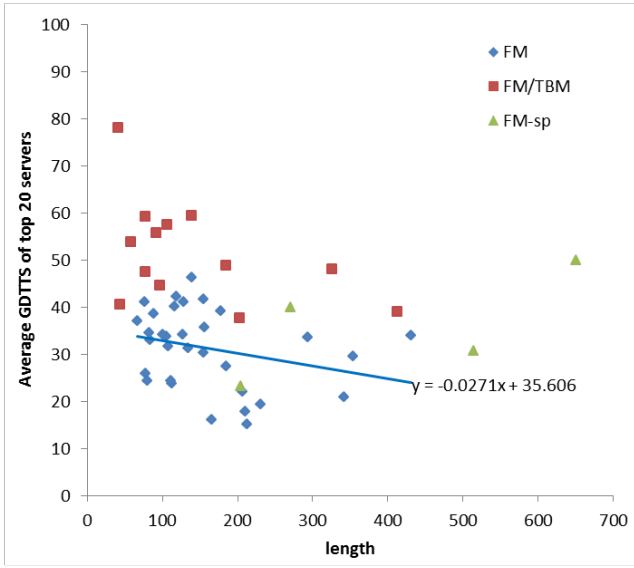




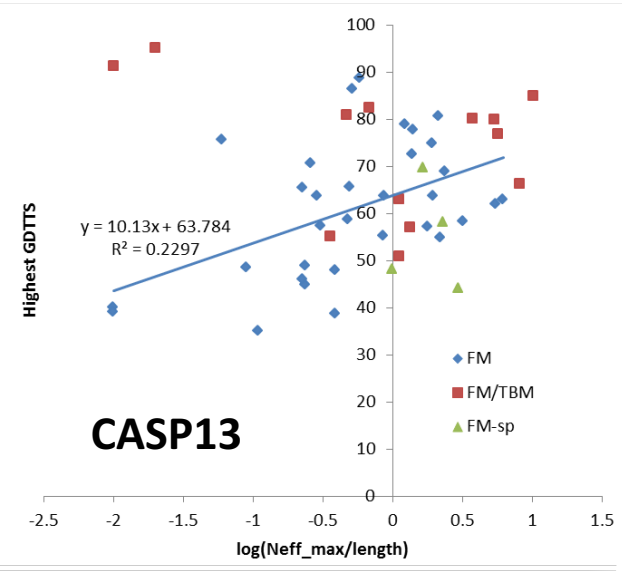


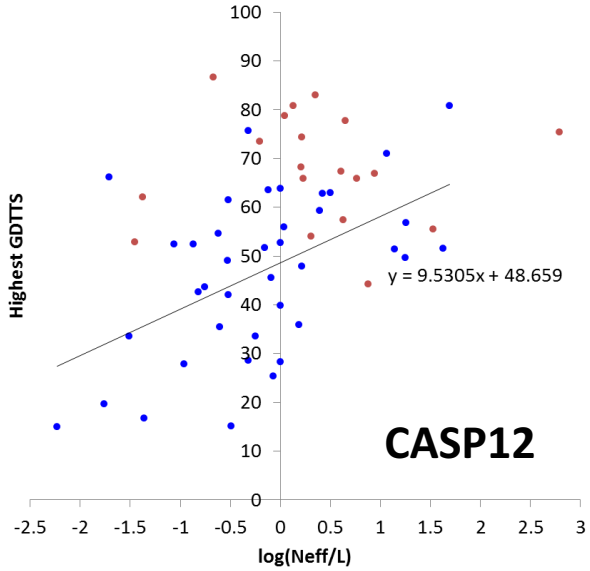






Possible sources of improvement: alignment depth, existing templates, domain size?





- From CASP12 to CASP13 significant improvement in performance
- Do some predictors have access to special, close metagenomics databases ?

Key conclusions from CASP13 on the tertiary structure prediction track

- Yet another significant improvement in prediction quality, mainly due to the rise of **machine learning methods** combined with **coevolution-based contact prediction**
- Reaching **nearly atomistic resolution** of the backbone for some very difficult EUs (< 150 residues) by **many groups!**
- Predictions are so good that **splitting EUs is in some cases not necessary**
- Alignment depth allows for better top models than in CASP12, but now seem to need lower numbers of sequences
- Templates of poor sequence similarity might be better identified than in CASP12
- Remaining limitations: **domain size** and **alignment depth**