

RNA-Puzzles

E. Westhof

Architecture et Réactivité de l'ARN
Institut de biologie moléculaire et cellulaire du CNRS

Université de Strasbourg

2015



RNA Puzzles

- *RNA-Puzzles* is a CASP-like **collective blind** experiment for the evaluation of **automatic** RNA three-dimensional structure prediction.
- It requires voluntary **RNA crystallographers**.
- It requires voluntary teams of **RNA modelers**.

Cruz, et al. RNA (2012), 18:610-625

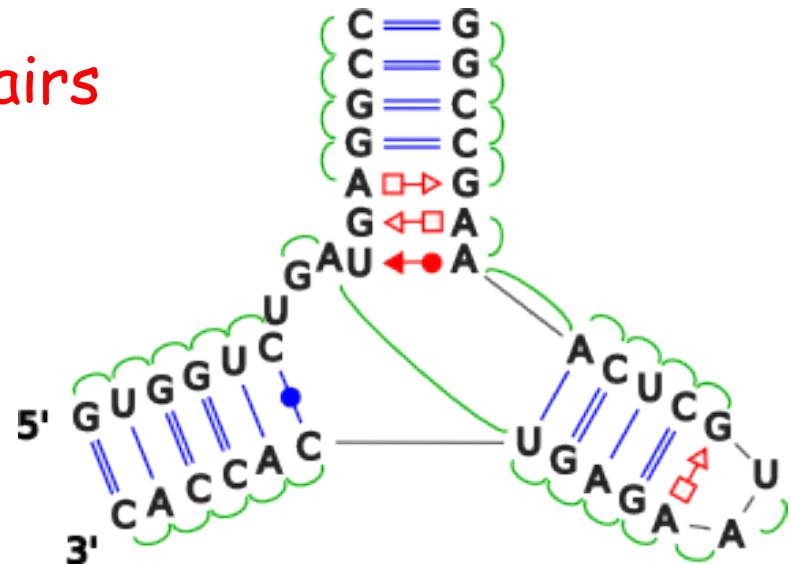
Miao, et al. RNA (2015), 21:1066-1084

Metrics for comparisons based on Interaction Networks

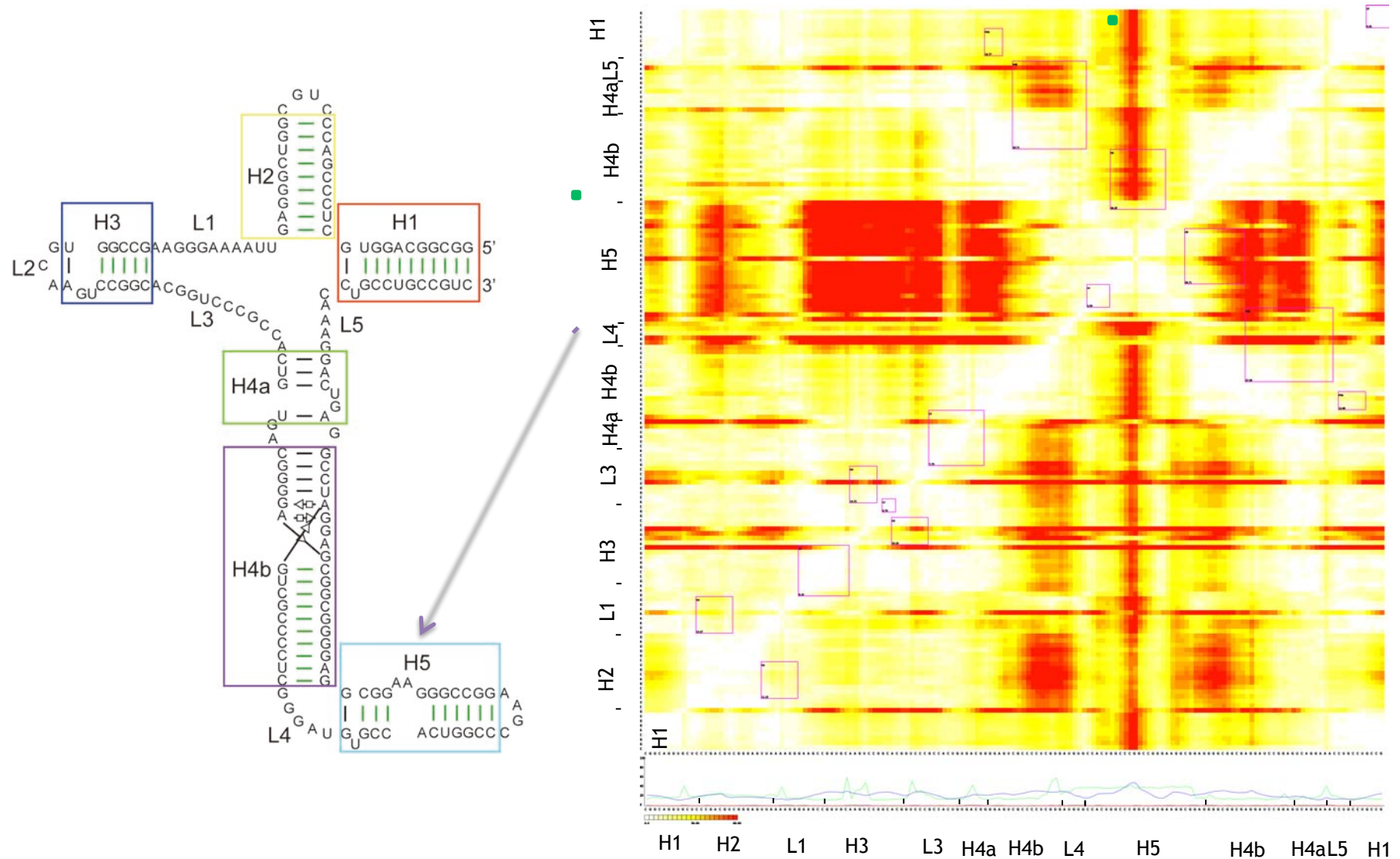
Set of base-base interactions constitute the elementary structural building units of RNA molecules

- Watson-Crick base pairs
- Non Watson-Crick base pairs
- Base stacking

Good Structure Prediction
||
v
Correct Interaction Network

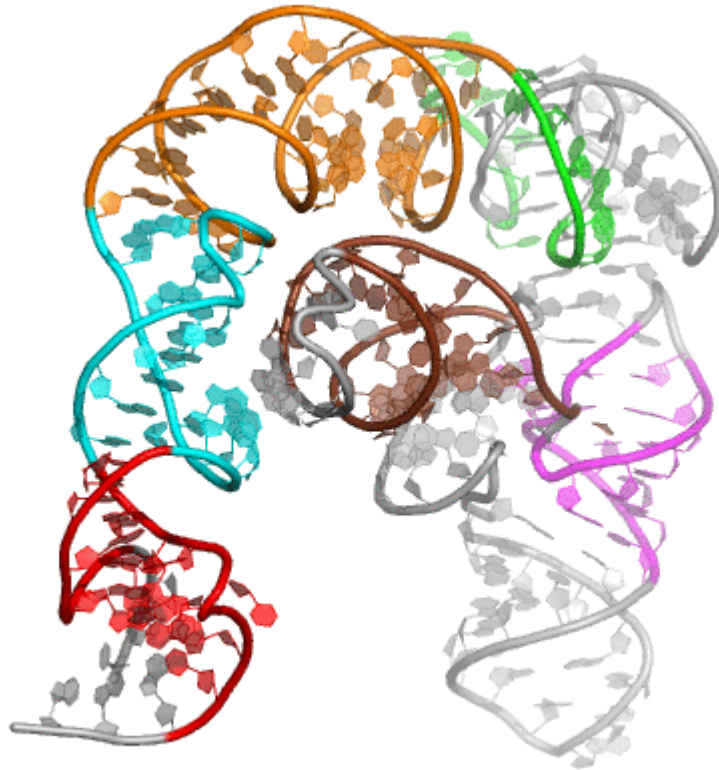


Deformation profile

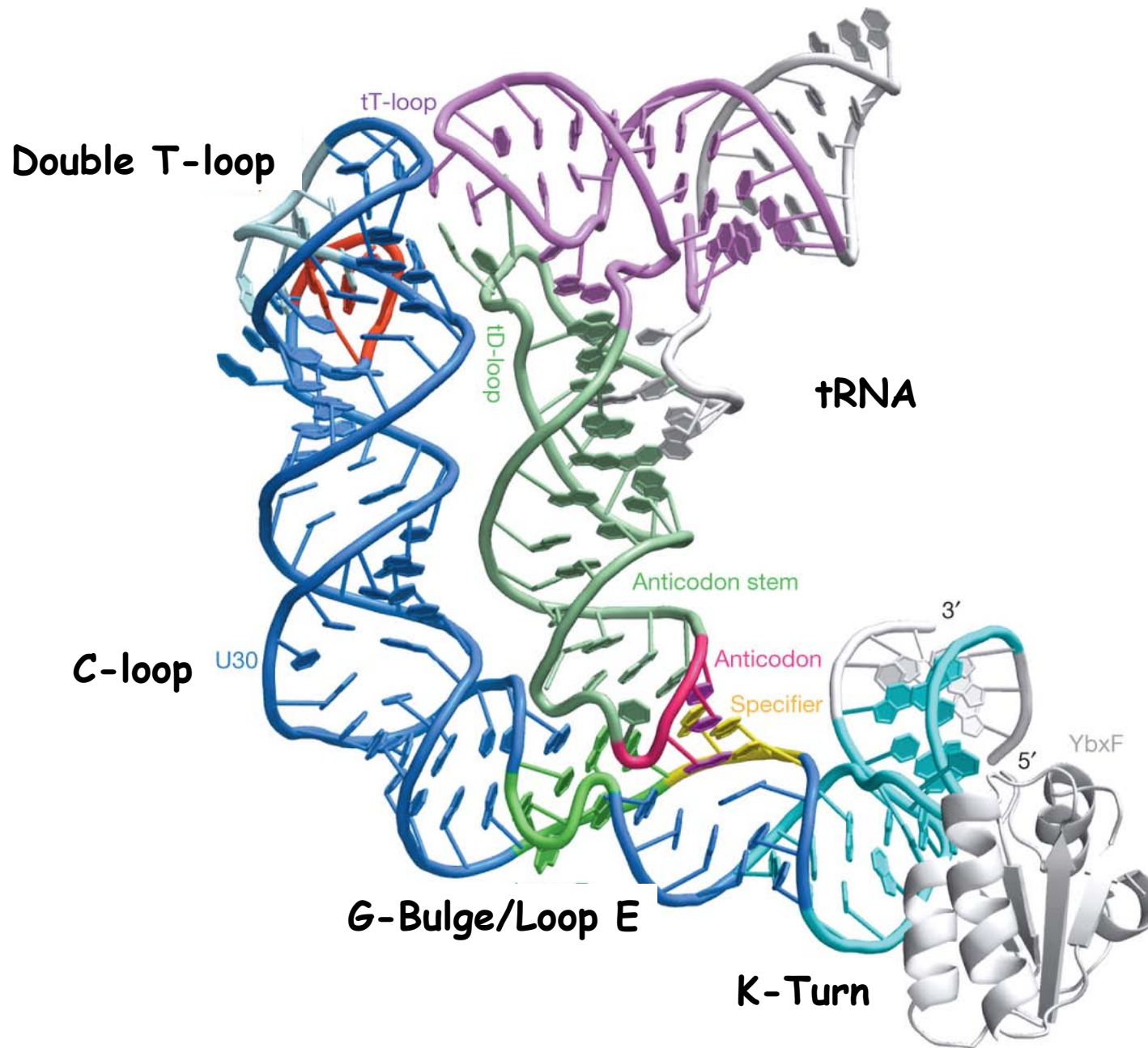


Problem 10

- The T-box riboswitch in complex with tRNA



- PDB: 4LCK
- Resolution: 3.20Å
- Avg B = 128 Å²
- tRNA: 75nt
- T-box : 96nt
- Clash score: 2.28
- Some small fragments solved by NMR before. tRNA structure known.



Overall structure of the T-box stem I in complex with tRNA.

Zhang, J. & Ferré-D'Amaré, A.R. *Nature* 500, 363-366 (2013)

Table of results (Problem 10 T-box)

Group ^a	Number ^b	RMSD ^c	Rank ^d	DI all ^e	Rank ^d	INF all ^f	Rank ^d	INF wc ^g	Rank ^d	INF nwc ^h	Rank ^d	INF stack ⁱ	Rank ^d	Clash Score ^j	Rank ^d	MCQ ^k	Rank ^d
Das	4	5.96	1	7.40	1	0.81	3	0.91	2	0.62	6	0.80	3	10.73	21	13.92	1
Das	3	6.51	2	8.18	2	0.80	5	0.90	4	0.65	4	0.79	9	11.09	24	15.39	5
Das	1	8.08	3	10.10	3	0.80	4	0.86	7	0.65	3	0.81	2	11.64	25	14.87	3
Das	5	9.32	4	11.24	4	0.83	1	0.91	3	0.74	1	0.82	1	10.91	22	15.08	4
Das	2	9.40	5	11.67	5	0.81	2	0.91	1	0.67	2	0.80	5	11.64	26	14.74	2
Bujnicki	4	10.62	6	16.64	15	0.64	20	0.81	20	0.00	25	0.64	18	0.91	2	21.77	21
Bujnicki	1	11.66	7	14.80	6	0.79	6	0.85	9	0.50	10	0.80	6	1.27	4	19.85	10
Bujnicki	8	11.69	8	14.86	7	0.79	7	0.85	10	0.56	8	0.79	7	1.45	7	18.32	6
Bujnicki	5	11.87	9	15.35	8	0.77	8	0.81	17	0.33	13	0.80	4	1.81	10	19.33	7
Bujnicki	7	11.88	10	16.14	14	0.74	14	0.81	18	0.30	14	0.75	14	1.45	6	20.07	11
Bujnicki	3	11.98	11	15.74	11	0.76	11	0.86	8	0.38	12	0.77	10	0.91	1	22.13	25
Bujnicki	6	12.00	12	15.95	13	0.75	13	0.79	24	0.17	15	0.79	8	1.27	5	19.46	8
Bujnicki	10	12.02	13	15.60	9	0.77	9	0.85	11	0.63	5	0.76	12	1.45	8	22.13	24
Bujnicki	9	12.02	14	15.83	12	0.76	12	0.81	19	0.59	7	0.76	13	0.91	3	20.43	12
Bujnicki	2	12.02	15	15.71	10	0.77	10	0.84	13	0.50	9	0.77	11	1.63	9	19.56	9
Dokholyan	2	12.29	16	19.44	18	0.63	23	0.80	22	0.00	17	0.62	20	8.89	16	20.90	14
Dokholyan	3	12.67	17	18.75	17	0.68	16	0.84	12	0.00	18	0.67	16	7.80	13	21.65	20
Chen	1	13.01	18	18.53	16	0.70	15	0.77	26	0.42	11	0.73	15	7.81	14	22.07	23
Dokholyan	8	13.22	19	21.39	20	0.62	24	0.83	14	0.00	23	0.59	25	8.71	15	21.63	19
Dokholyan	9	13.36	20	21.96	22	0.61	25	0.78	25	*	*	0.59	24	9.80	19	21.24	17
Dokholyan	10	13.47	21	20.90	19	0.65	19	0.88	6	0.00	24	0.61	22	10.16	20	21.98	22
Dokholyan	4	14.14	22	22.23	23	0.64	22	0.82	15	0.00	19	0.62	21	9.07	18	21.60	18
Dokholyan	1	14.42	23	21.93	21	0.66	17	0.80	21	0.00	16	0.66	17	11.07	23	20.92	15
Dokholyan	5	15.99	24	26.69	25	0.60	26	0.82	16	0.00	20	0.57	26	8.89	17	22.41	26
Dokholyan	7	16.87	25	26.47	24	0.64	21	0.79	23	0.00	22	0.64	19	6.35	11	20.64	13
Dokholyan	6	17.93	26	27.69	26	0.65	18	0.88	5	0.00	21	0.61	23	6.71	12	21.21	16
Mean		12.09		17.35		0.72		0.84		0.31		0.71		6.32		19.74	
Standard deviation		2.76		5.35		0.08		0.04		0.29		0.09		4.26		2.67	
											X-ray model			2.28			

Table of results (Problem 10 tRNA)

TABLE 5. Summary of the results for Puzzle 10 tRNA

Group ^a	Number ^b	RMSD ^c	Rank ^d	DI all ^e	Rank ^d	INF all ^f	Rank ^d	INF wc ^g	Rank ^d	INF nwc ^h	Rank ^d	INF stack ⁱ	Rank ^d	Clash Score ^j	Rank ^d	MCQ ^k	Rank ^d
Bujnicki	7	2.49	1	2.75	1	0.91	8	0.97	10	0.91	7	0.88	8	1.45	6	19.03	5
Bujnicki	9	2.50	2	2.76	2	0.91	9	0.97	11	0.91	8	0.88	9	0.91	3	19.17	6
Bujnicki	4	2.57	3	2.80	3	0.92	3	0.97	8	1.00	2	0.89	7	0.91	2	20.56	15
Bujnicki	2	2.60	4	2.85	4	0.91	4	0.97	9	0.91	6	0.89	4	1.63	9	19.63	10
Bujnicki	6	2.63	5	2.89	5	0.91	5	0.94	13	1.00	4	0.89	6	1.27	5	19.42	8
Bujnicki	8	2.65	6	2.98	10	0.89	10	0.94	14	0.80	15	0.88	10	1.45	7	20.39	14
Bujnicki	10	2.67	7	2.93	6	0.91	6	0.97	12	0.85	14	0.90	2	1.45	8	19.39	7
Bujnicki	5	2.68	8	2.96	9	0.91	7	0.92	16	1.00	3	0.89	5	1.81	10	19.62	9
Bujnicki	3	2.69	9	2.93	7	0.92	2	0.97	7	0.89	9	0.90	1	0.91	1	19.82	12
Bujnicki	1	2.71	10	2.95	8	0.92	1	1.00	5	0.91	5	0.89	3	1.27	4	20.03	13
Das	3	2.92	11	3.39	11	0.86	15	1.00	2	0.85	11	0.81	15	11.09	24	18.78	4
Das	5	2.96	12	3.41	12	0.87	12	1.00	4	0.85	13	0.83	11	10.91	22	18.45	3
Das	1	3.01	13	3.50	13	0.86	14	1.00	1	0.85	10	0.81	14	11.64	25	19.82	11
Das	2	3.21	14	3.69	14	0.87	11	0.97	6	1.00	1	0.81	12	11.64	26	16.19	1
Das	4	3.32	15	3.85	15	0.86	13	1.00	3	0.85	12	0.81	13	10.73	21	17.12	2
Chen	1	3.66	16	4.59	16	0.80	16	0.91	19	0.60	16	0.78	17	7.81	14	24.33	26
Dokholyan	6	4.63	17	6.52	18	0.71	24	0.90	20	0.26	20	0.67	25	6.71	12	22.38	18
Dokholyan	5	4.87	18	7.30	22	0.67	26	0.84	24	*	*	0.63	26	8.89	17	23.74	25
Dokholyan	9	4.88	19	6.93	20	0.71	25	0.82	26	0.00	24	0.72	22	9.80	19	23.39	24
Dokholyan	3	5.04	20	6.44	17	0.78	17	0.92	18	0.20	23	0.80	16	7.80	13	22.96	21
Dokholyan	1	5.27	21	7.01	21	0.75	19	0.92	17	0.22	22	0.74	19	11.07	23	23.38	23
Dokholyan	7	5.27	22	7.32	23	0.72	22	0.84	25	0.45	18	0.71	23	6.35	11	22.05	16
Dokholyan	4	5.30	23	6.91	19	0.77	18	0.92	15	0.45	17	0.73	21	9.07	18	23.18	22
Dokholyan	10	5.52	24	7.47	24	0.74	21	0.89	22	0.00	25	0.74	20	10.16	20	22.89	20
Dokholyan	8	5.74	25	7.63	25	0.75	20	0.87	23	0.26	21	0.75	18	8.71	15	22.15	17
Dokholyan	2	6.90	26	9.62	26	0.72	23	0.89	21	0.26	19	0.69	24	8.89	16	22.79	19
Mean		3.80		4.78		0.83		0.94		0.65		0.81		6.32		20.79	
Standard deviation		1.33		2.15		0.08		0.05		0.34		0.08		4.26		2.18	
											X-ray model			2.28			

Values in each row correspond to a predicted model.

^aName of the research group that submitted the model.

^bNumber of the model among all models from the same group.

^cRMSD of the model compared with the accepted structure (in Å).

^dColumns indicate the rank of the model with respect to the left-hand column metric.

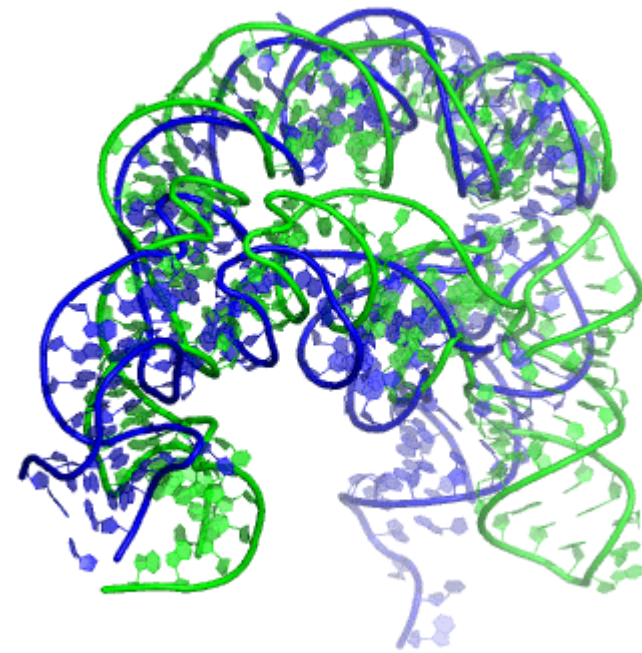
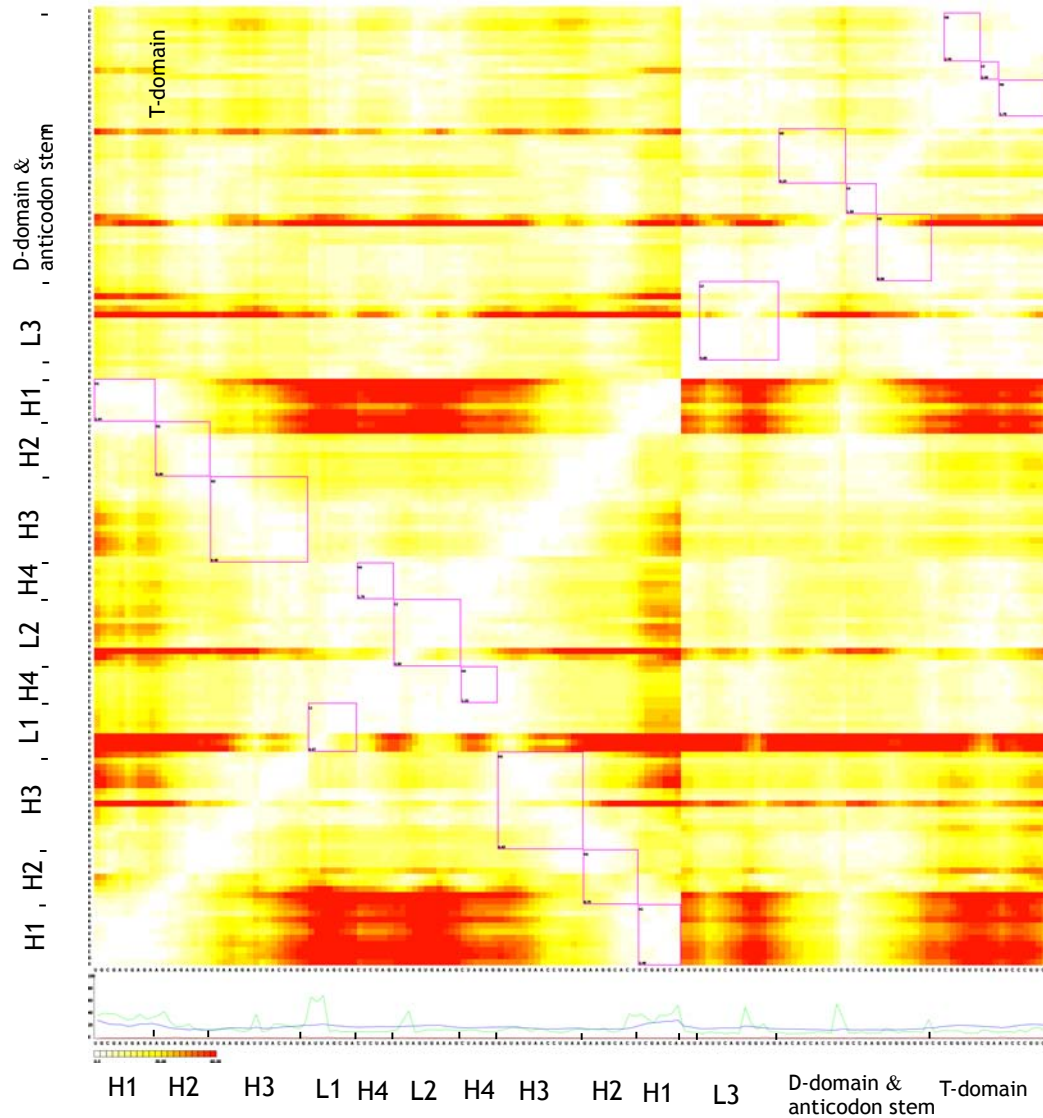
^eDI all is the Deformation Index taking into account all interactions (stacking, Watson–Crick, and non-Watson–Crick).

^fINF all is the Interaction Network Fidelity taking into account all interactions.

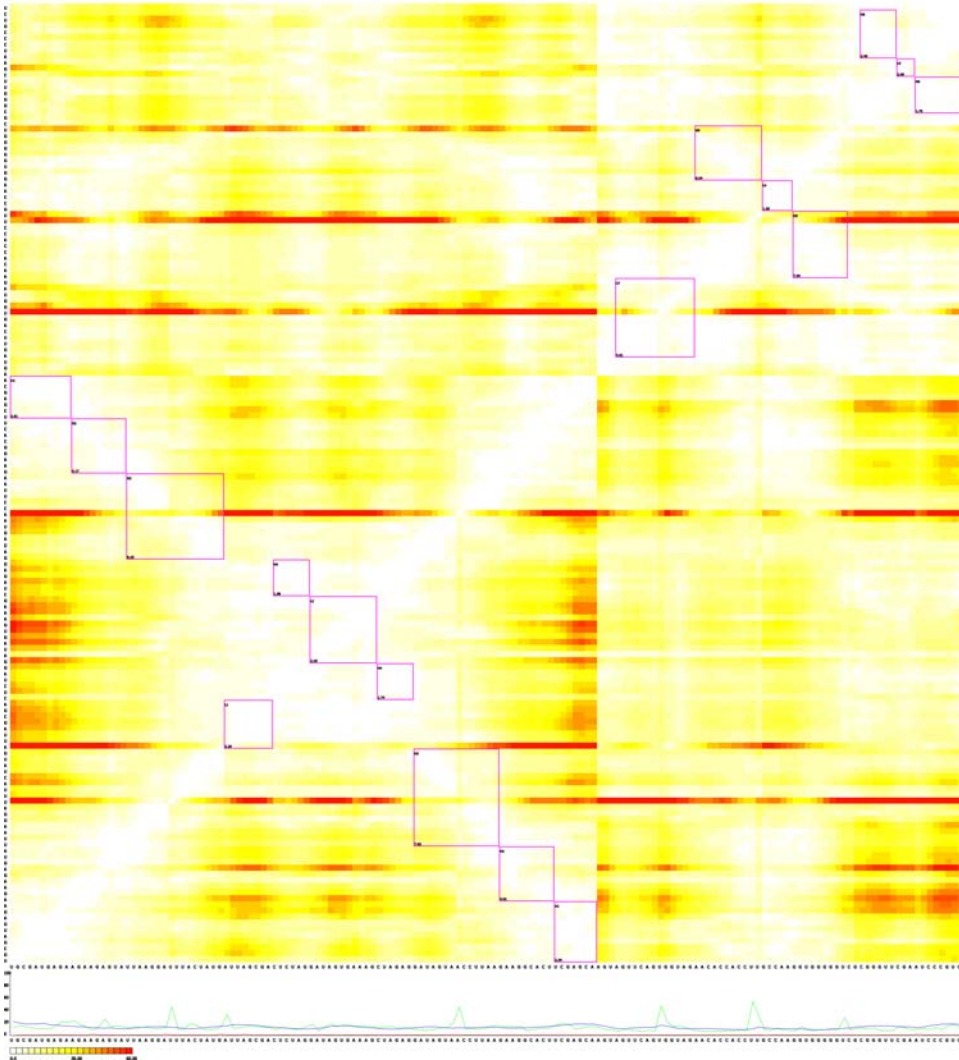
^gINF wc is the Interaction Network Fidelity taking into account only Watson–Crick interactions.

^hINF nwc is the Interaction Network Fidelity taking into account only non-Watson–Crick interactions.

Deformation profile Das model #1



2D map of deformation profile

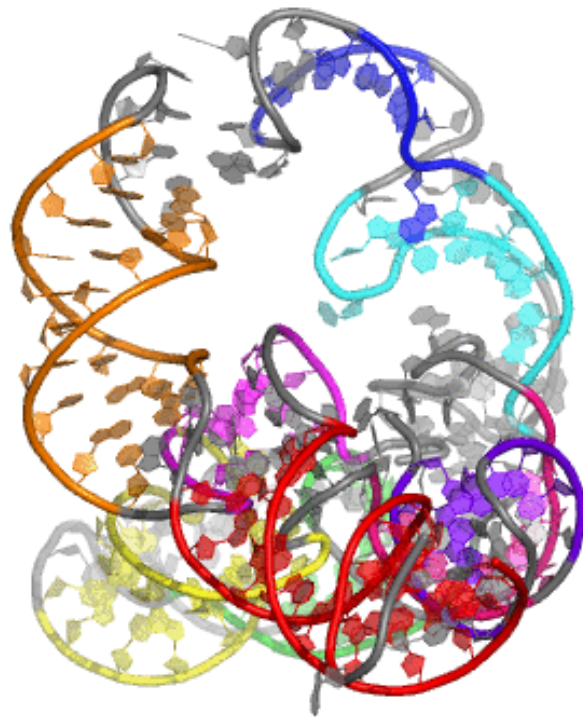


- Most of the key regions are well predicted in the best model.

Deformation profile of the Das 3 model .

Problem 5

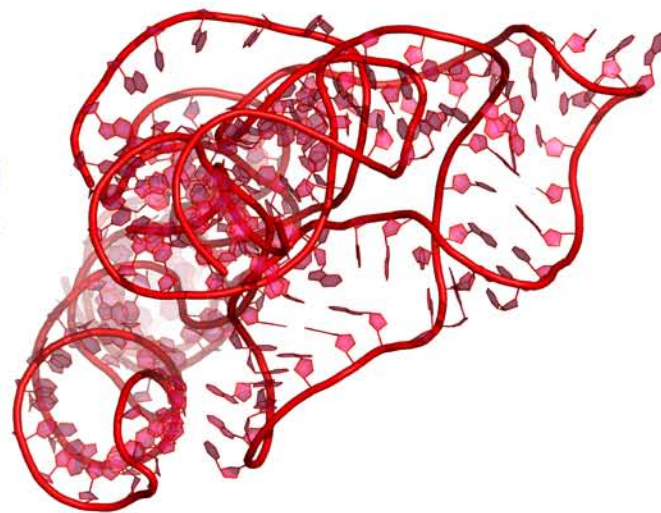
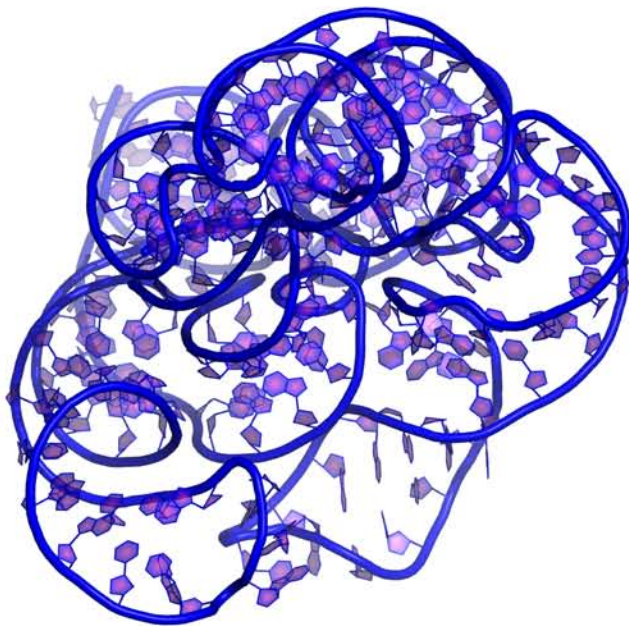
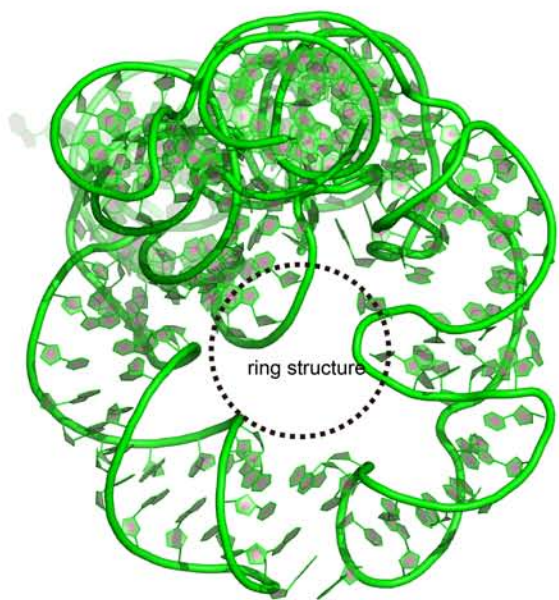
- Lariat capping ribozyme (group I intron)



- 4P95/4P9R
- Resolution : 2.5/2.7Å
- Avg B = 18/45 Å²
- 189nt
- Clash score: 5.86
- No homologous structure

Table of results (Problem 5)

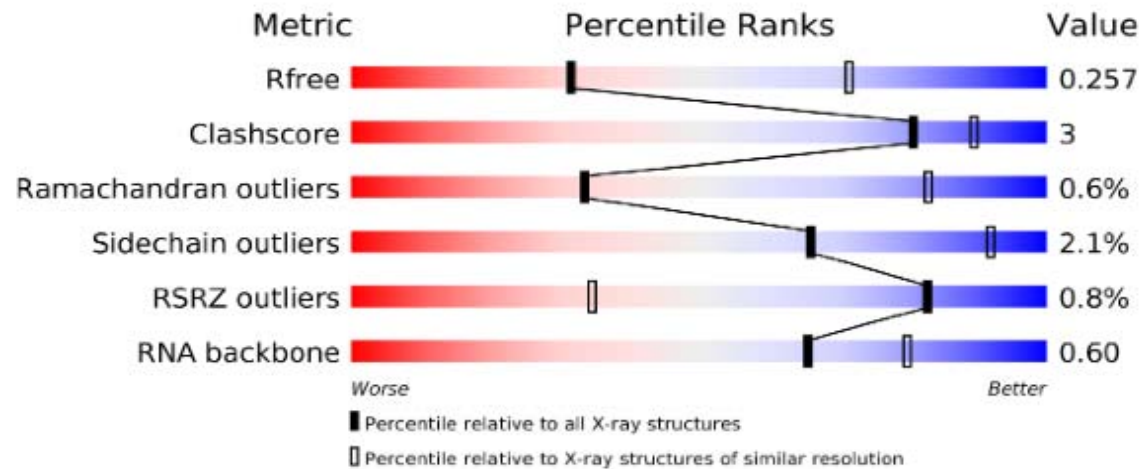
Group ^a	Number ^b	RMSD ^c	Rank ^d	DI all ^e	Rank ^d	INF all ^f	Rank ^d	INF wc ^g	Rank ^d	INF nwc ^h	Rank ^d	INF stack ⁱ	Rank ^d	Clash Score ^j	Rank ^d	MCQ ^k	Rank ^d
Das	2	9.15	1	12.02	1	0.76	1	0.91	3	0.33	1	0.75	2	6.79	12	10.57	2
Das	1	9.95	2	13.15	2	0.76	2	0.92	1	0.26	2	0.75	1	9.44	21	11.63	4
Adamiak	1	16.78	3	23.96	3	0.70	13	0.90	5	0.00	5	0.66	20	11.10	23	13.30	13
Dokholyan	2	19.90	4	29.14	6	0.68	19	0.84	13	*	*	0.66	22	7.62	14	13.87	17
Dokholyan	4	19.90	5	29.14	7	0.68	20	0.84	14	*	*	0.66	23	7.62	15	13.87	18
Bujnicki	2	20.77	6	28.09	4	0.74	3	0.86	9	0.00	11	0.74	5	1.66	8	12.14	6
Dokholyan	6	20.82	7	28.65	5	0.73	4	0.91	2	0.00	7	0.69	12	10.43	22	15.81	23
Dokholyan	7	21.44	8	30.61	8	0.70	14	0.91	4	0.00	8	0.66	19	8.77	19	14.63	22
Bujnicki	4	22.24	9	32.80	11	0.68	21	0.79	22	0.00	13	0.68	15	1.16	6	10.34	1
Bujnicki	3	22.37	10	36.99	16	0.61	25	0.72	24	0.00	12	0.62	25	5.30	10	12.75	12
Dokholyan	5	23.01	11	32.55	9	0.71	11	0.87	8	*	*	0.68	17	6.46	11	13.68	16
Dokholyan	3	23.69	12	32.64	10	0.73	5	0.88	7	0.00	6	0.71	10	8.61	18	14.61	21
Bujnicki	5	23.81	13	34.15	13	0.70	16	0.86	10	0.00	14	0.68	16	0.66	5	11.73	5
Dokholyan	8	24.01	14	33.84	12	0.71	10	0.89	6	0.00	9	0.67	18	9.11	20	14.40	20
Bujnicki	1	24.78	15	36.69	15	0.68	22	0.79	21	0.00	10	0.69	13	1.32	7	11.23	3
Chen	2	25.67	16	36.33	14	0.71	12	0.80	20	0.00	16	0.73	6	7.78	16	13.36	14
Chen	7	26.26	17	39.50	19	0.67	24	0.76	23	0.00	21	0.69	14	0.00	3	13.39	15
Dokholyan	1	27.20	18	38.03	17	0.72	8	0.84	12	*	*	0.70	11	7.45	13	14.33	19
Chen	1	27.24	19	38.90	18	0.70	15	0.81	18	0.00	15	0.71	9	4.64	9	12.19	7
Chen	3	28.71	20	41.69	20	0.69	18	0.71	25	0.00	17	0.74	4	11.42	24	12.35	10
Chen	5	31.39	21	44.06	22	0.71	9	0.81	19	0.00	19	0.73	7	13.41	25	12.19	8
Chen	4	31.52	22	43.50	21	0.73	6	0.83	15	0.00	18	0.74	3	8.28	17	12.25	9
Chen	6	31.55	23	44.07	23	0.72	7	0.83	16	0.00	20	0.73	8	0.00	2	12.47	11
Xiao	2	32.53	24	48.54	24	0.67	23	0.83	17	0.20	4	0.64	24	0.17	4	19.90	25
Xiao	1	36.54	25	52.90	25	0.69	17	0.86	11	0.21	3	0.66	21	0.00	1	19.71	24
Mean		24.05		34.48		0.70		0.84		0.05		0.69		5.97		13.47	
Standard deviation		4.91		7.11		0.03		0.05		0.06		0.03		4.30		2.30	
											X-ray model			5.86			



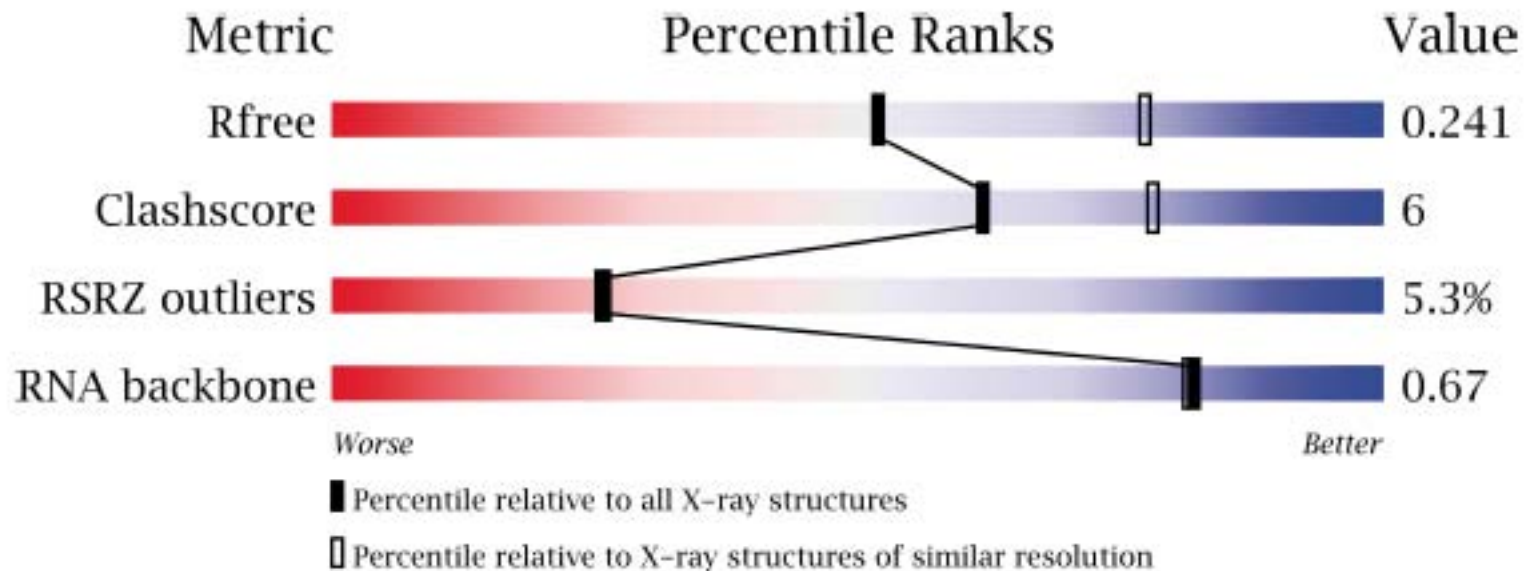
Starting hypothesis : the X-ray structure
is correct and accurate.

But, is it always sure?

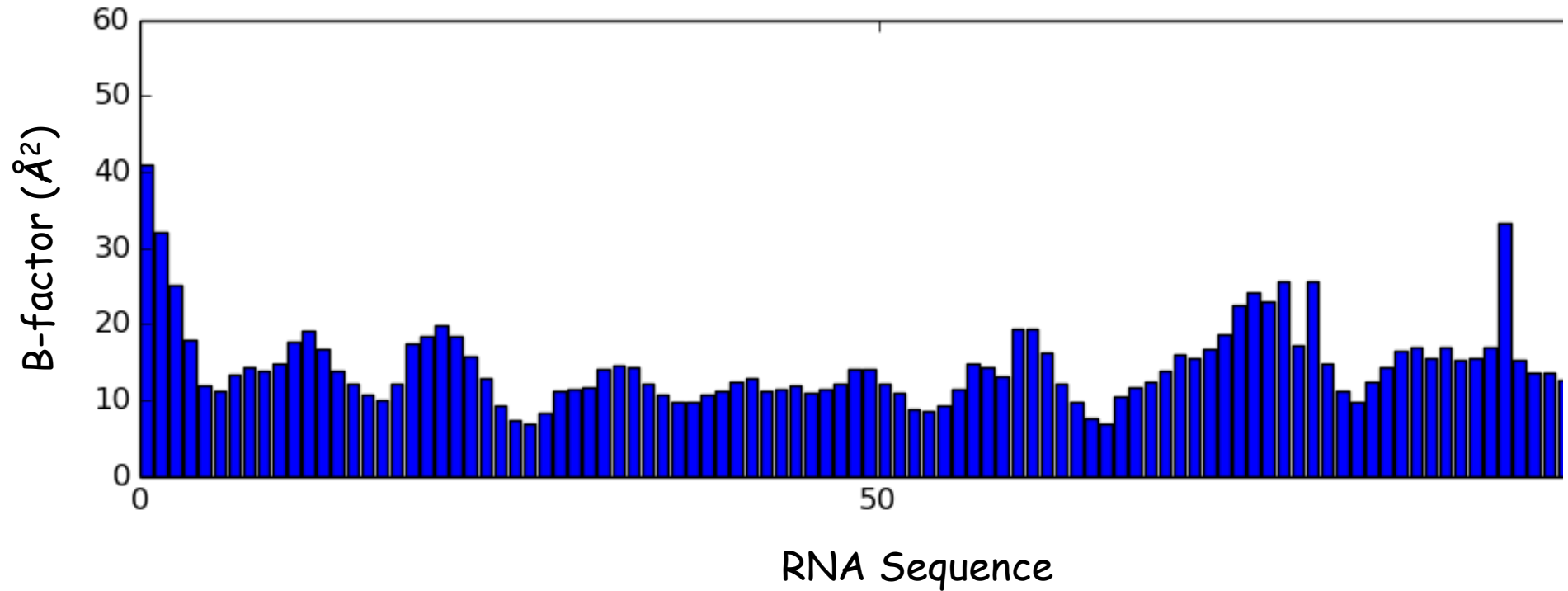
The T-box riboswitch in complex with tRNA



The lariat ribozyme



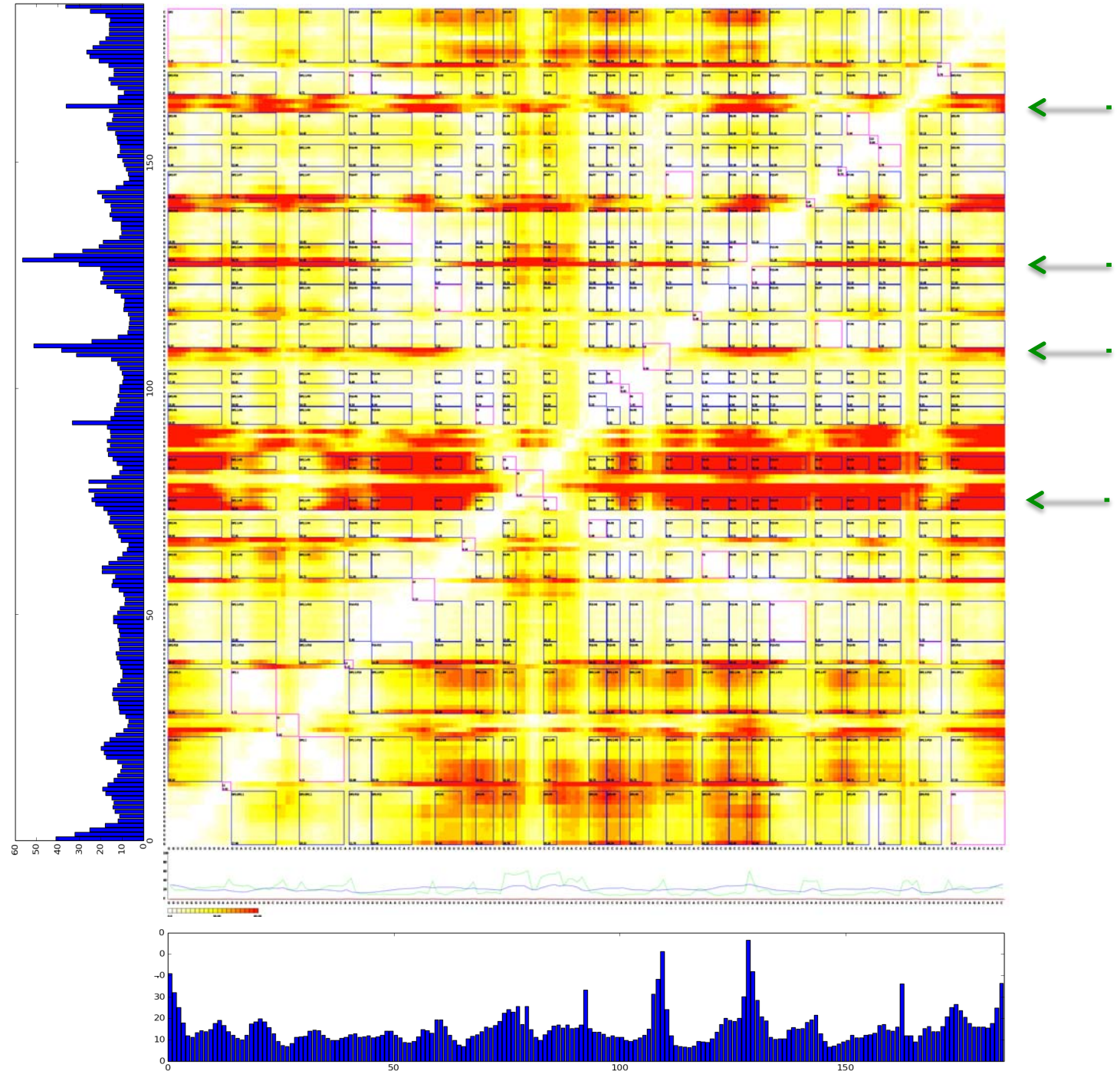
What about B-factors?



P5

“Best
model”

Das 2



General trends

- General correlations between the various metrics (except with clash score).
- Top models have better deformation profile
- And better nwc-INF (**non-Watson-Crick** Interaction Network Fidelity)
- The rankings proposed by the modellers do not correspond to the best metric scores

Discussion

- Non-Watson-Crick interaction and RNA modules
 - Not derived from 2D
 - Key for 3D
 - Should be detected before 3D modelling
- Clashes exist in models
 - How to build model with right topology and fewer clashes?
 - Build the topology -> optimize the model
 - Build and optimize at the same time

Discussion

- How to incorporate RNA dynamics in the comparisons?
- How to deduce a theoretical measure of "local mobility" that could act as a proxy for B-factors.

RNA-Puzzles

- New Competition process will allow the assessment of more aspects in RNA 3D structure prediction
- How accurate automatic prediction can be (with known 2D or not)?
- How well participants can rank/assess prediction models (their own and those of others).



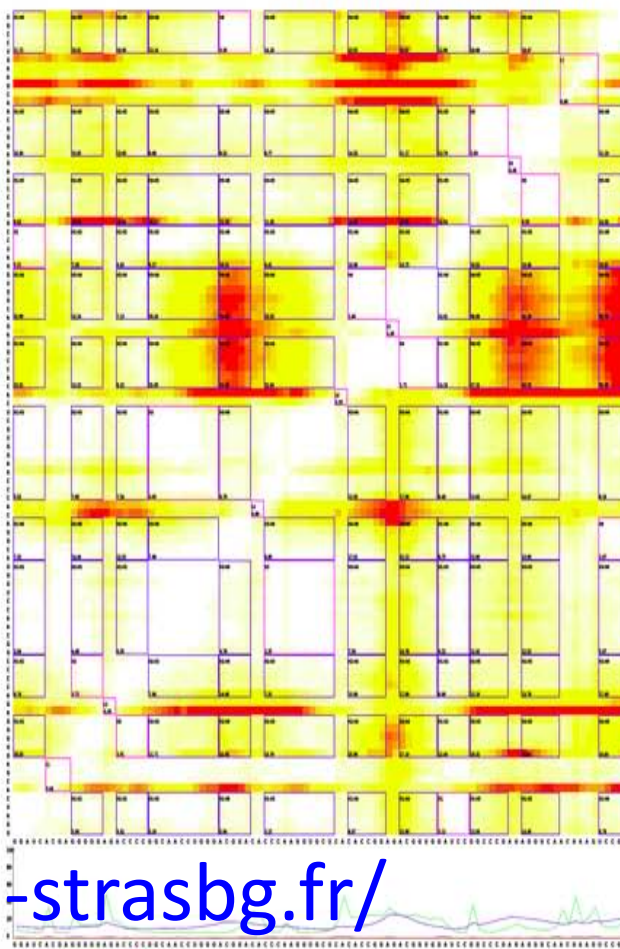
- Dry run: test of web servers
- Web server: 2 days
- Web server given 2D structure: 2 days
- Human experts: 30 days (with the help of experiments)
- Assessment of prediction models: participants predict ranks of all the prediction models

Problem #8 Result: #7 from Bujnicki lab

Result summary

RMSD: 6.711 (# 3)
 P-value: 1.789e-09(# 3)
 DI all: 8.934 (# 3)
 INF all: 0.751 (# 20)
 INF wc: 0.906 (# 24)
 INF nwc: 0.289 (# 22)
 INF stack: 0.711 (# 20)
 Clash Score:0.0 (# 1)
 P-value: 1.789e-09(# 3)
 MCQ: 15.751 (# 16)

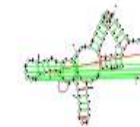
Deformation Profile Matrix



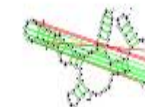
Click figure to zoom in.



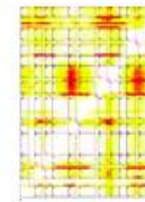
<< show big



<< show big



<< show big



<< show big

<http://ahsoka.u-strasbg.fr/rnapuzzles/>

The co-authors

- **Crystallographers:**
 - (Alexander Serganov, Alla Peselis (NYU, New-York))
 - Adrian Ferre-D'Amare, Jinwei Zhang (NIH)
 - Benoît Masquida, Mélanie Meyer (CNRS, Strasbourg)
- **Groups of Modellers:**
 - François Major (Montreal, Canada)
 - Rhiju Das (Stanford, US)
 - Shi-Jie Chen (Missouri, US)
 - Janusz M. Bujnicki (Warsaw, Poland)
 - Adamiak Ryszard (Poznan, Poland)
 - Nikolay V. Dokholyan (North Carolina, US)
 - Feng Ding (South Carolina, US)

UPR 9002 du CNRS,
Architecture et Réactivité de l'ARN,
Institut de Biologie Moléculaire et Cellulaire,
Université de Strasbourg



José Cruz

Zhichao Miao

French Government grant ANR-10-
BINF-02-02 " BACNET "

