中山大學孫逸仙紀念医院
SUN YAT-SEN MEMORIAL HOSPITAL,SUN YAT-SEN UNIVERSITY

RNA 3D structure prediction

GuangzhouRNA-human (183)
GuangzhouRNA-meta (417)

CASP16

Lin Huang

2024/12/03

The RNA structures we determined

4C40 4BW0 4C4W 4CS1
5FJ1 5FK1 5FK2 5FK3 5FK4 5FK5 5FK6 5FKD
5FKF 5FKH 5FJC 5FKE 5FKG 5FJ4 5FJ0 5G4T
5G4V 5G4U 5NDI 5NEX 5NEO 5NEP 5NOM 5NDH
5NEQ 5NEF 5LQO 5LR4 5LQT 5LR3 5LR5 5O62
5NZ3 5NZ6 5NZD 5NWQ 5NY8 5O69
6FZ0 6HCT 6HBT 6HBX 6HC5 6QN3 6Q8V 6Q8U
6YLB 6YMK 6YMM 6YL5 6YMJ 6YMI 6YML 6TF1 6TF3 6TFE 6TFG
6TFF 6TFH 6TF2 6TF0 6TB7
7EAG 7EAF 7V9E
8HB8 8HBA 8HB3 8HB1 8I3Z 8ITS

Released PDBs 75

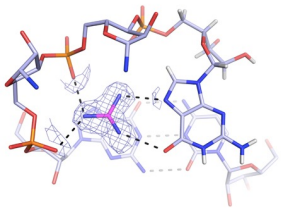
NEW RNA structures 15

8Z9K 8Z8U 8Z8Q 8KEB 8ZA0 8ZA4 8ZAU 8K7W
9G4R 9G4Q 9G4P

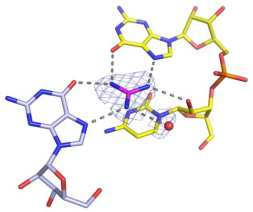
Deposited PDBs 11

NEW RNA structures 7

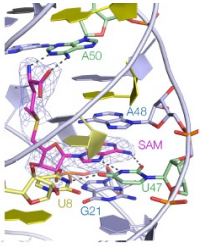
Riboswitches and ribozymes



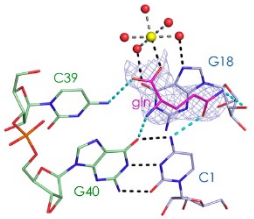
Guanidine-II



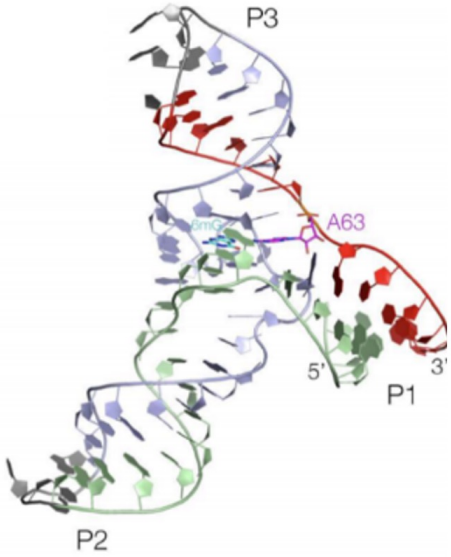
Guanidine-III



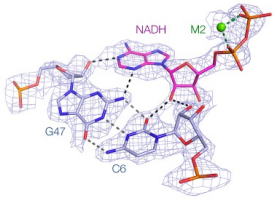
SAM-V



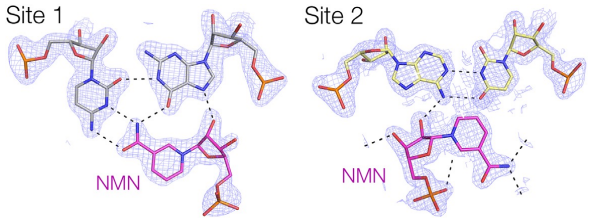
Glutamine-II



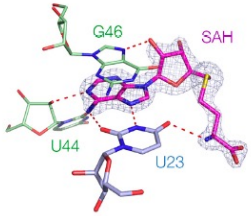
MTR1 Ribozyme



NAD⁺-I



NAD⁺- II



SAM/SAH

Two RNA modelling related meetings

Computational Approaches to RNA Structure and Function
Benasque, Spain
July 15-27, 2018

2nd RNA Puzzles meeting
Warsaw, Poland, December 6-8, 2018

RNA modeling and molecular replacement

19 June 2019

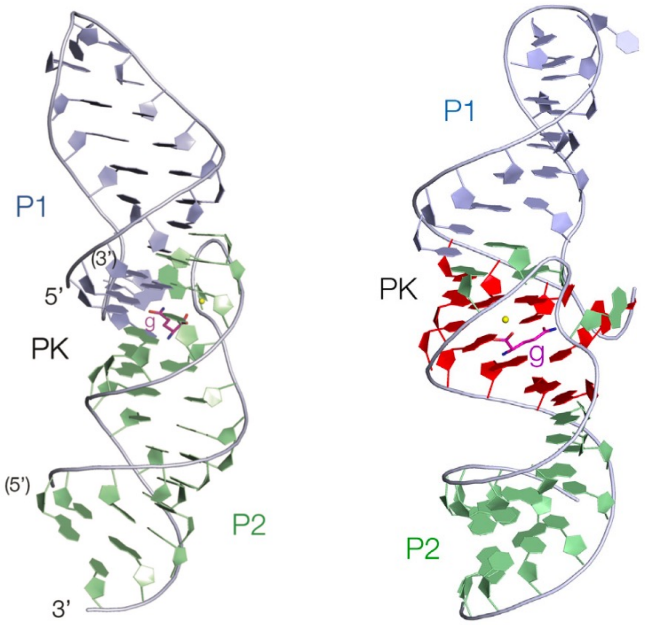
7666–7675 *Nucleic Acids Research, 2019, Vol. 47, No. 14*
doi: 10.1093/nar/gkz539
Published online 19 June 2019

Structure and ligand binding of the glutamine-II riboswitch

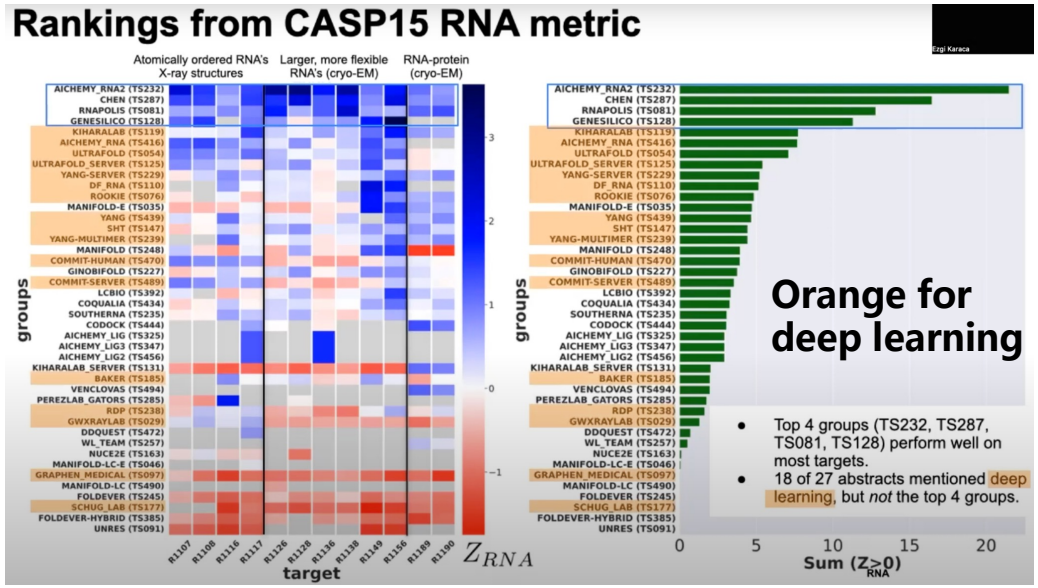
Lin Huang[1,†], Jia Wang[1,†], Andrew M. Watkins[2], Rhiju Das[2] and David M. J. Lilley[1,*]

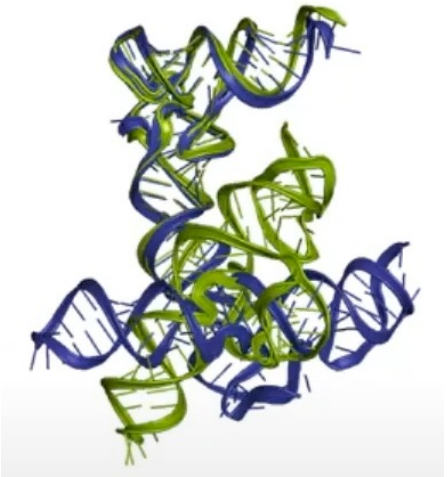
ABSTRACT

We have determined the structure of the glutamine-II riboswitch ligand binding domain using X-ray crystallography. The structure was solved using a novel combination of homology modeling and molecular replacement. The structure comprises three coaxial



Deep learning has not provided significant assistance in RNA structure prediction in CASP 15





Rhiju Das

Proteins. 2023 Dec;91(12):1747-1770

Main reasons for low prediction accuracy

1. Insufficient number of independently functional RNA structures, fewer than 200

16S and 23S rRNA,
18S and 28S rRNA,
group I and group II introns,
RNase P ribozymes,
tmRNAs,
tRNAs,
...

Small Ribozymes ≈ 10
Riboswitches ≈ 50
Aptamers ≈ 50

2. Research on RNA motifs is severely lacking

The problem: **Inadequate learning content for artificial intelligence and deep learning.**

Annual Review of Biophysics 2017 46: 483-503

RNA Structures：three databases

16S and 23S rRNA,
18S and 28S rRNA,
group I and group II introns,
RNase P ribozymes,
Small Ribozymes ≈ 10

tmRNAs,
tRNAs,
...

Riboswitches ≈ 50

Aptamers ≈ 50

D262–D268 *Nucleic Acids Research, 2023, Vol. 51, Database issue* *Published online 30 September 2022*
https://doi.org/10.1093/nar/gkac840

Ribocentre: a database of ribozymes

Jie Deng[1,†], Yaohuang Shi[2,3,†], Xuemei Peng[1,†], Yuanlin He[1], Xiaoxue Chen[1,4], Mengxiao Li[1], Xiaowei Lin[1,5], Wenjian Liao[1,5], Yuanyin Huang[1], Taijiao Jiang[3], David M.J. Lilley[6,*], Zhichao Miao[3,7,*] and Lin Huang[1,*]

Nucleic Acids Research, 2024, **52**, D265–D272
https://doi.org/10.1093/nar/gkad891
Advance access publication date: 19 October 2023
Database issue

Ribocentre-switch: a database of riboswitches

Fan Bu[1,2,†], Xiaowei Lin[3,4,†], Wenjian Liao[3,4,†], Zhizhong Lu[5], Yuanlin He[4], Yuhang Luo[4], Xuemei Peng[4], Mengxiao Li[4], Yuanyin Huang[4], Xiaoxue Chen[4], Bowen Xiao[2], Jiuhong Jiang[2], Jie Deng[4], Jian Huang[3,4], Tianxin Lin[3,4], Zhichao Miao[6,2,*] and Lin Huang[4,*]

Ribocentre-aptamer: A database of RNA aptamers

Zhizhong Lu[1,3], Ying Ao[2], Bo Fu[1], Shuang Zhu[3,*],Lin Huang[1,*], Zhichao Miao[2,*]

1. Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Guangdong-Hong Kong Joint Laboratory for RNA Medicine, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou 510120, China.
2. GMU-GIBH Joint School of Life Sciences, The Guangdong-Hong Kong-Macau Joint Laboratory for Cell Fate Regulation and Diseases,Guangzhou National Laboratory, Medical University, Guangzhou 510180, China.
3. School of Life Sciences and Biopharmaceutics, Guangdong Pharmaceutical University, Guangzhou 510006, China.
Email to : Lin Huang (huanglin36@mail.sysu.edu.cn), Zhichao Miao (miao_zhichao@gzlab.ac.cn)

Assess the structure models

---from a crystallographer point of view

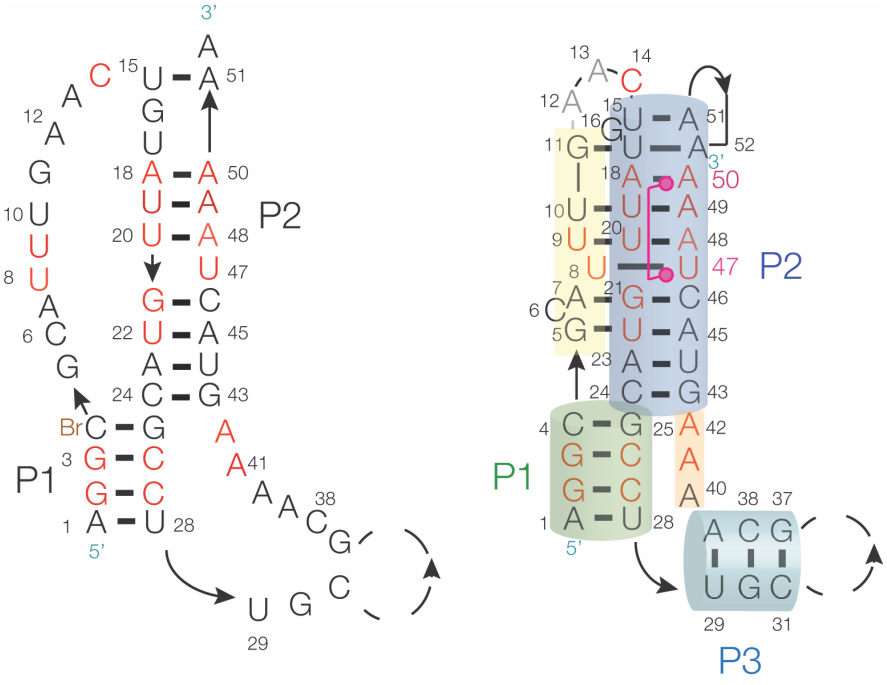
2nd RNA Puzzles meeting
Warsaw, Poland, December 6-8, 2018

1. Multiple sequence alignment is very important

2. Homology modelling, and fragment assemble

Assessment: multiple sequence alignment (MSA)

1. Most of RNA fold as the proposed secondary structure.
2. Could explain why some of the nucleotides are highly conserved.
3. Coaxial helical stacking follow empirical rules.

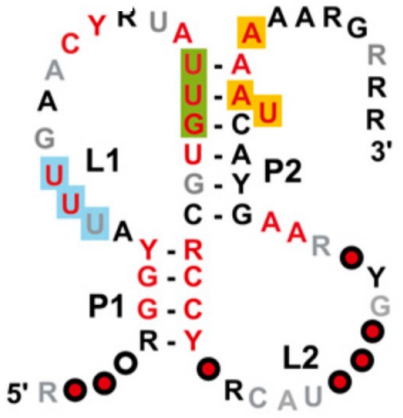


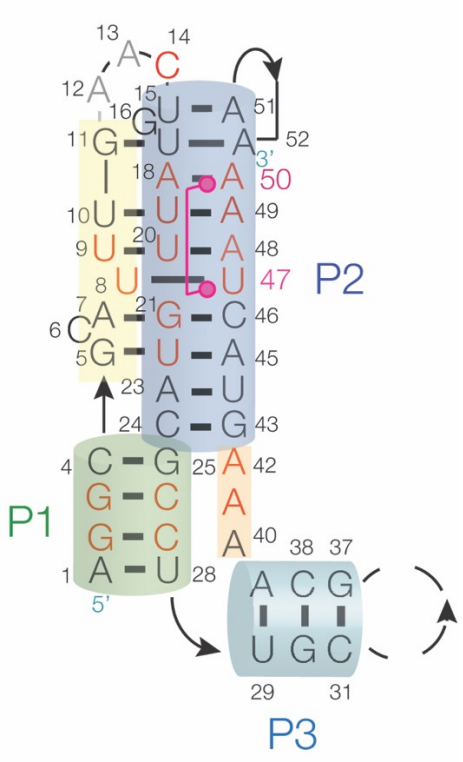
Lescoute and Westhof *RNA* 2006

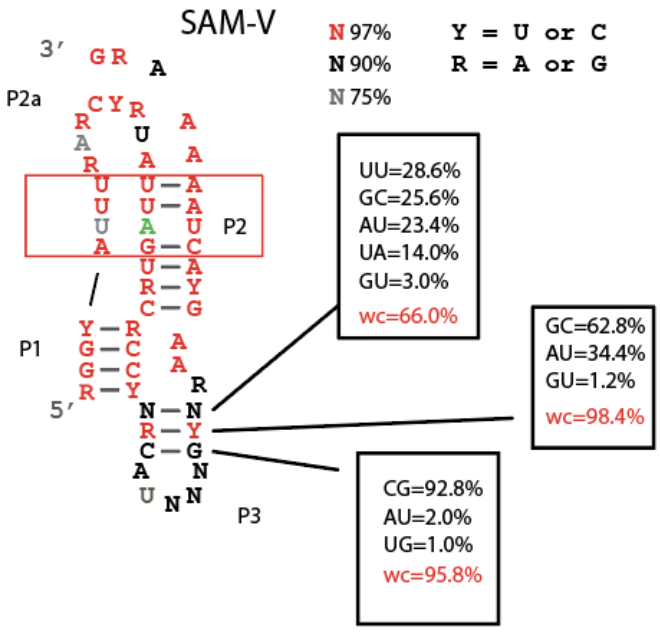
Huang & Lilley *Nucleic Acids res 2018*

Assessment: multiple sequence alignment (MSA)

4. Additional base-pair identified from the structure also supported by covariation analysis.

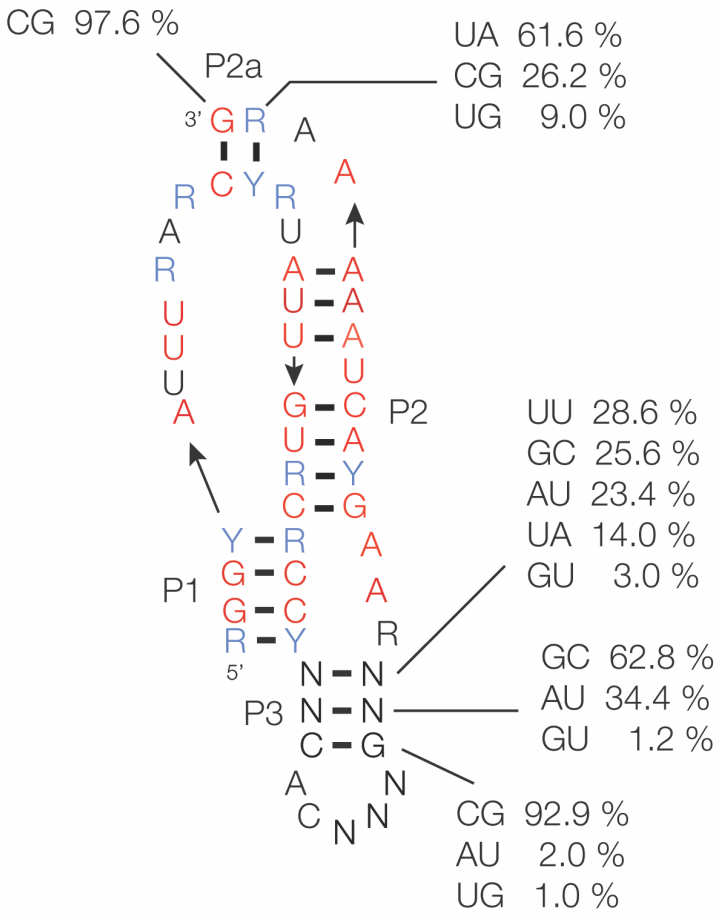






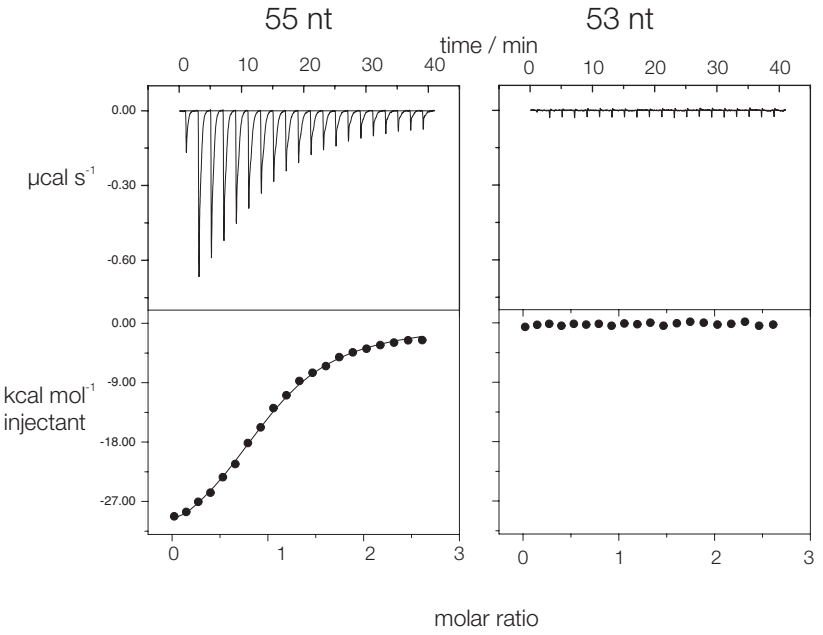
486 sequences analyzed

Assessment: multiple sequence alignment (MSA)



4. Additional base-pair identified from the structure also supported by covariation analysis.

At least 3 more structures we found additional base-pairs



Huang & Lilley *Nucleic Acids res 2018*

Homology modelling, and fragment assemble

19 June 2019

Rhiju Das

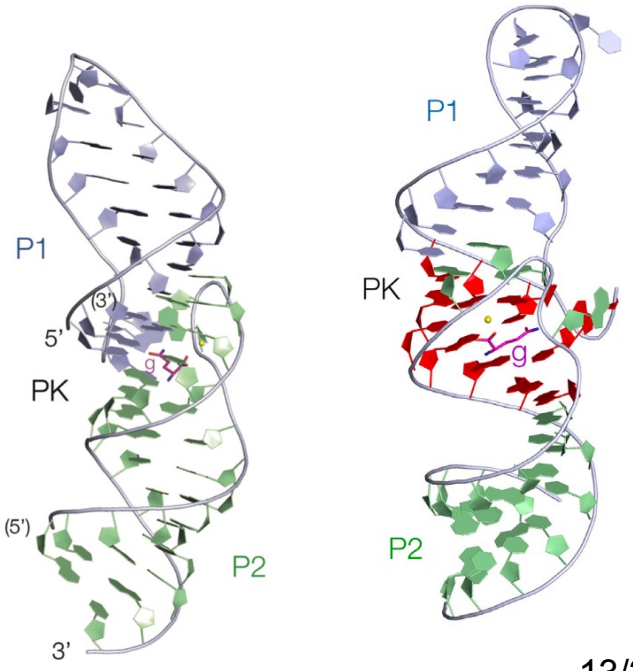
7666–7675 Nucleic Acids Research, 2019, Vol. 47, No. 14
doi: 10.1093/nar/gkz539
Published online 19 June 2019

Structure and ligand binding of the glutamine-II riboswitch

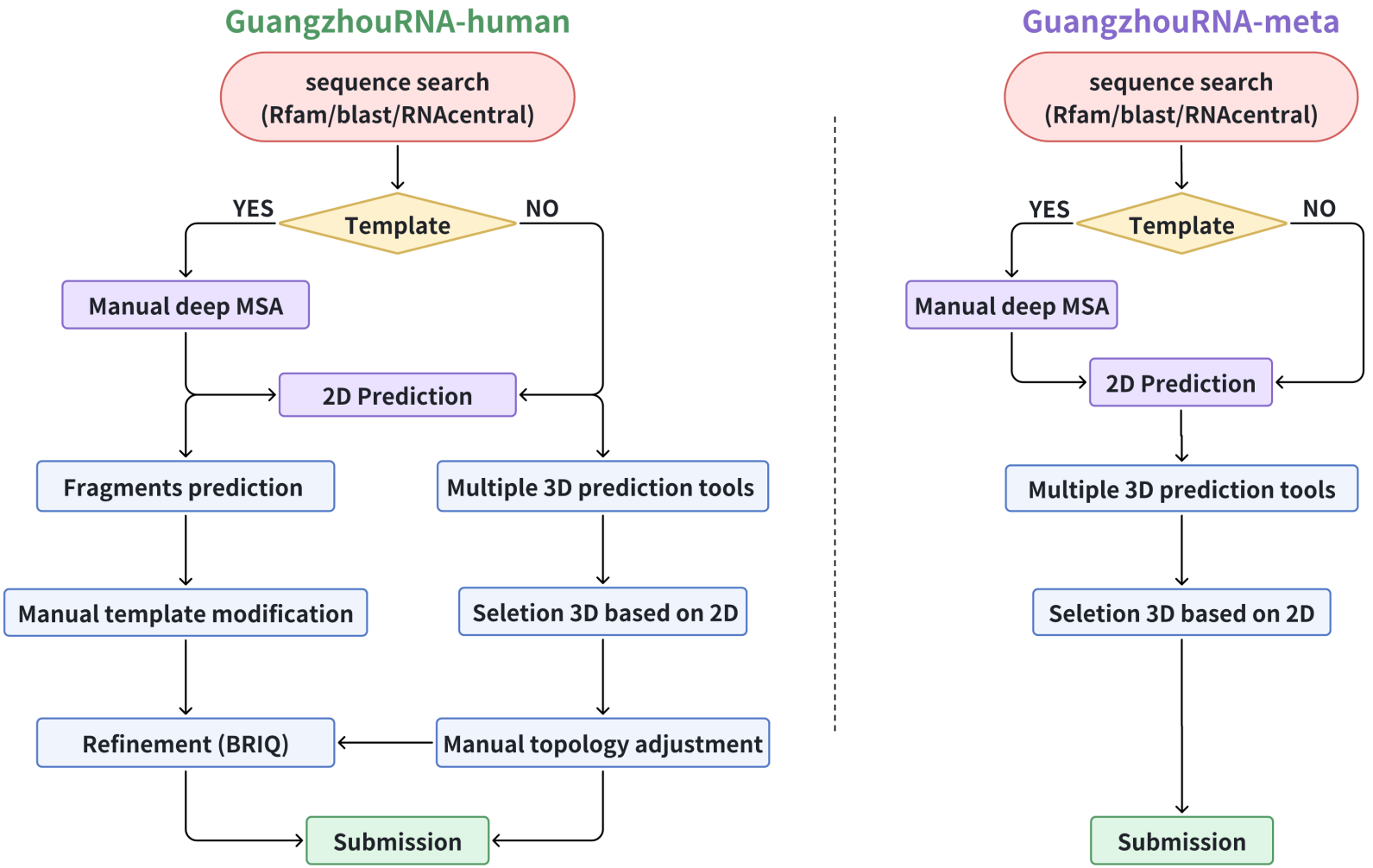
Lin Huang[1,†], Jia Wang[1,†], Andrew M. Watkins[2], Rhiju Das[2] and David M. J. Lilley[1,*]

ABSTRACT

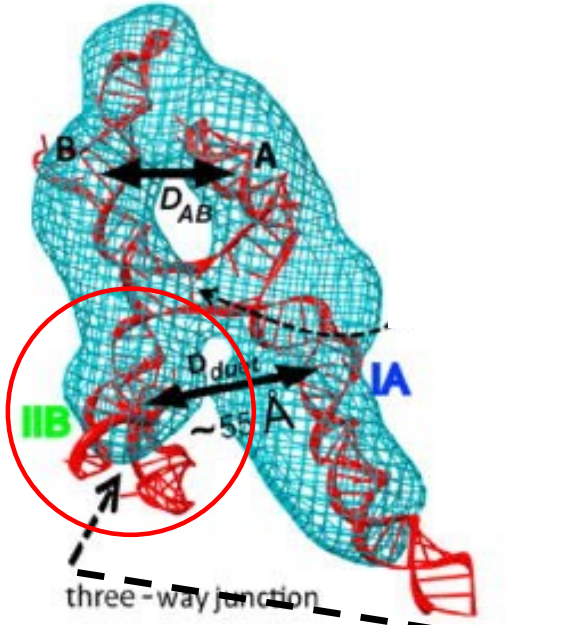
We have determined the structure of the glutamine-II riboswitch ligand binding domain using X-ray crystallography. The structure was solved using a novel combination of homology modeling and molecular replacement. The structure comprises three coaxial



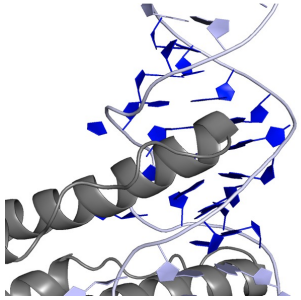
Workflow


GuangzhouRNA-human
sequence search (Rfam/blast/RNAcentral)
YES
Template
NO
Manual deep MSA
2D Prediction
Fragments prediction
Multiple 3D prediction tools
Manual template modification
Seletion 3D based on 2D
Refinement (BRIQ)
Manual topology adjustment
Submission
GuangzhouRNA-meta
sequence search (Rfam/blast/RNAcentral)
YES
Template
NO
Manual deep MSA
2D Prediction
Multiple 3D prediction tools
Seletion 3D based on 2D
Submission

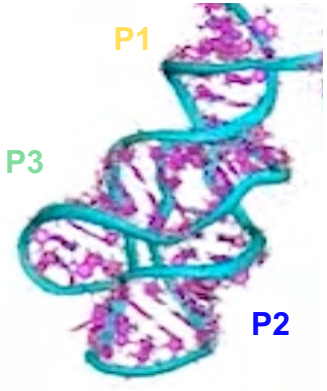
HIV-RRE stem loop II - R1209

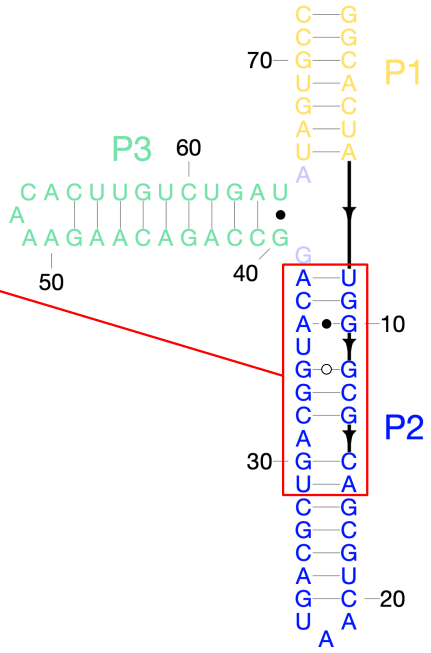


HIV-RRE in SAXS
(Fang X, et al. Cell. 2013)

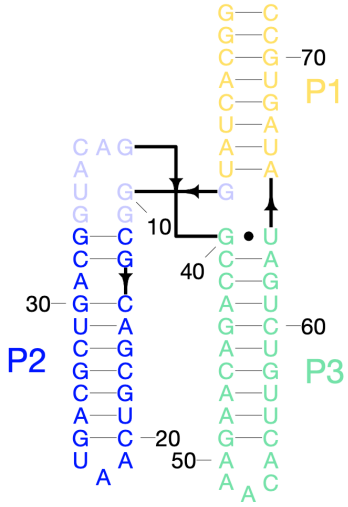


RRE RNA complex
(4PMI A:45-52 A:66-75)
X-ray 3.2 angstrom





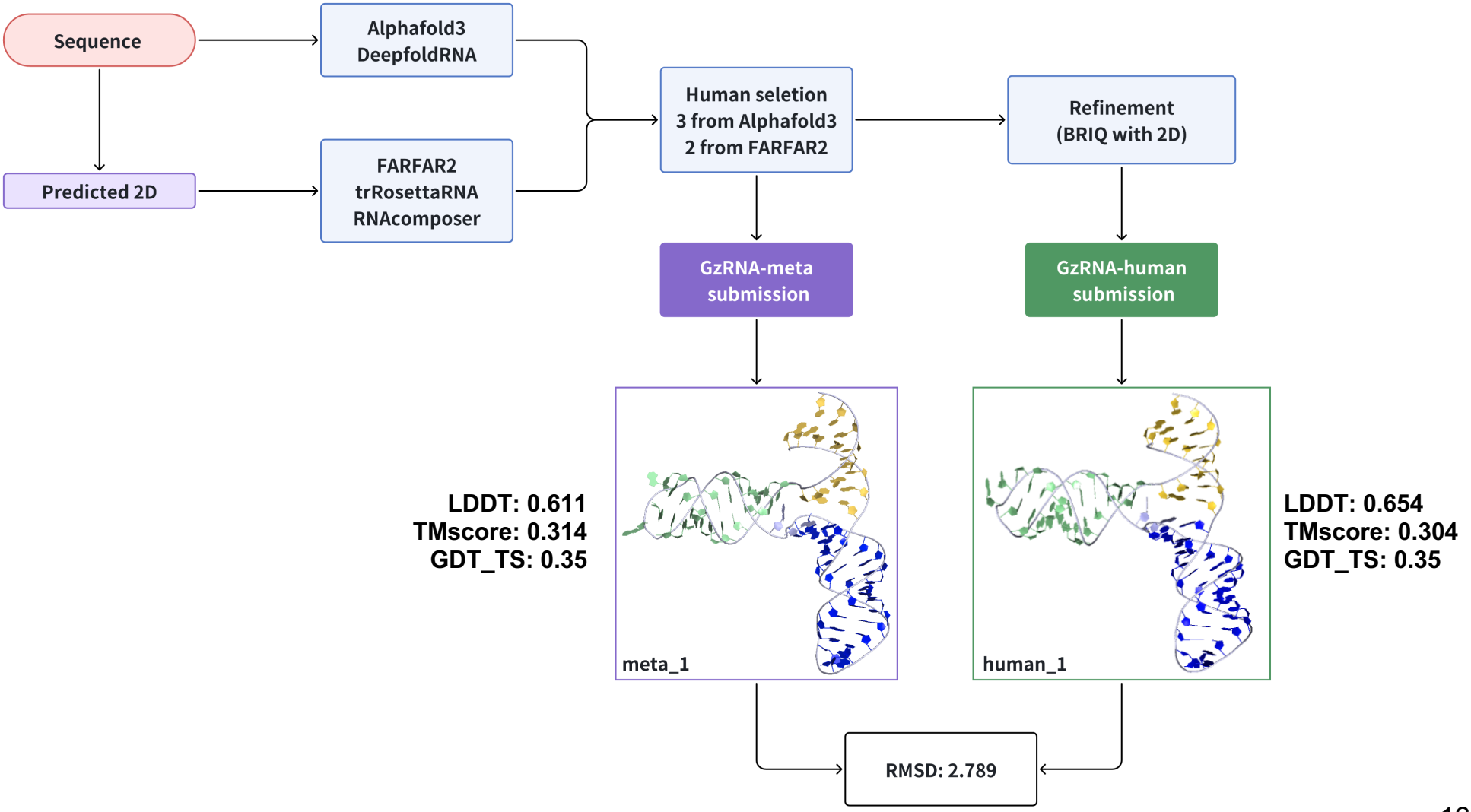
R1209-2D v1



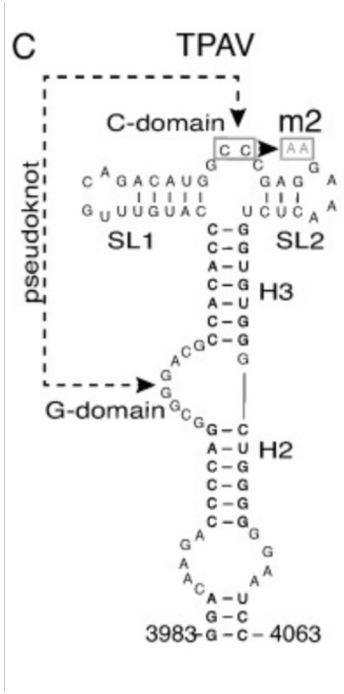
R1209-2D v2

```
R1209 GGCACUAUGGGCGCAGCGUCAAUGACGCUGACGGUACAGGCCAGACAAGAAACACUUGUCUGAUAUAGUGCC
      ((((((((((((((((((...))))))).))).)))).((((((((((.....))))))))))..)))))))
      (((((((...(((((((((...)))))))).))).....((((((((((.....))))))))))))))))
```

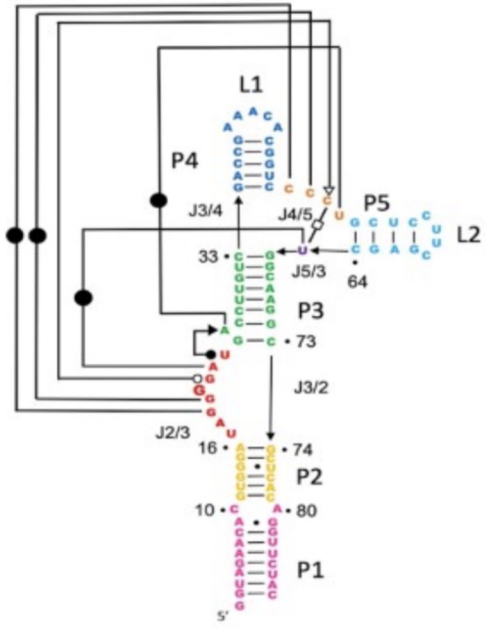
HIV-RRE stem loop II - R1209



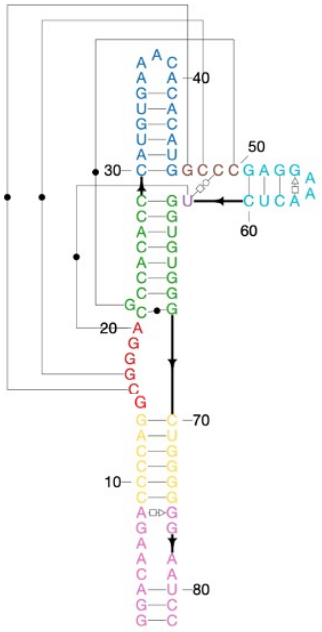
Translation enhancer motif - R1293



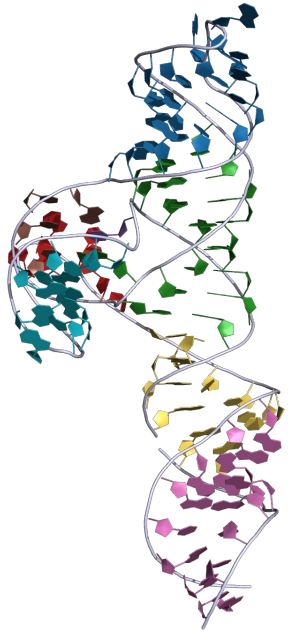
3'cap-independent translation enhancers (CITE) from paspalum asymptomatic virus (TPAV) (Jelena J.K. et al. Pathogens. 2019)



3'cap-independent translation enhancers (CITE) from Pea enation mosaic virus RNA 2 (PEMV2) (8SH5) (Anna Lewicka et al. NAR. 2023)



R1293

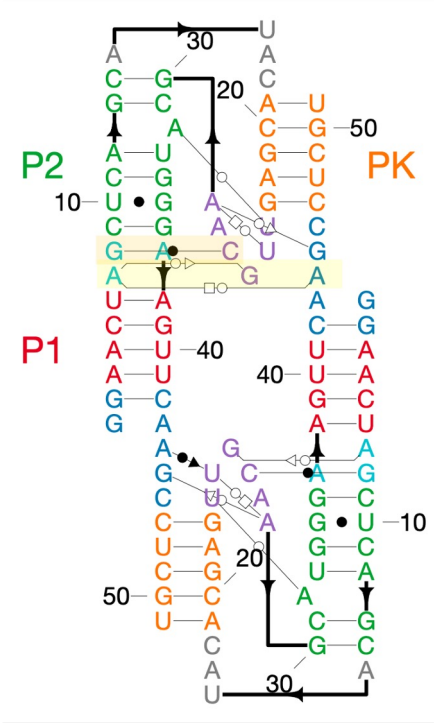


R1293_human_1 - 8SH5 edited

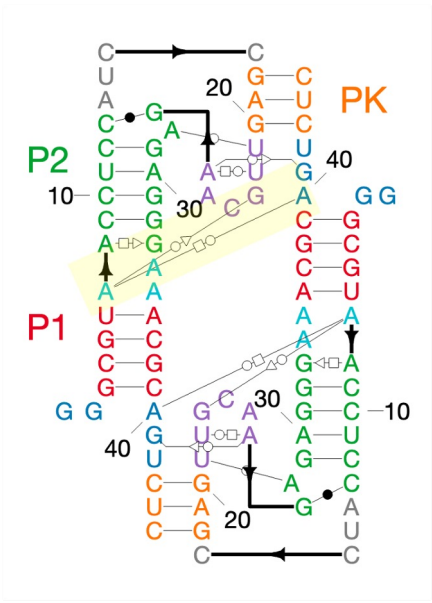
LDDT: 0.727
TMscore: 0.434
GDT_TS: 0.50

```
8SH5   GGUAGAAC----ACGUGGGAUAGGGGAUGACCUUGUCGACCG-AAACA-CGGUCCCCUGCUCCUUCGAGCUGGCAAGGCGCUCACAG---GUUCUAC
R1293  -----GGACAAGA-CCCCAGG-CGGGA-CGCCACACCCAUGUGAAACACACAUGGCCC-GAGGAAACUC-UGGUGUGGGCUGGGG-GGAAUCC----
       -----(((....(.((((((........(.((((((((((((((.....)))))))....((((....)))))..)))))))))))))).)...)))----
```

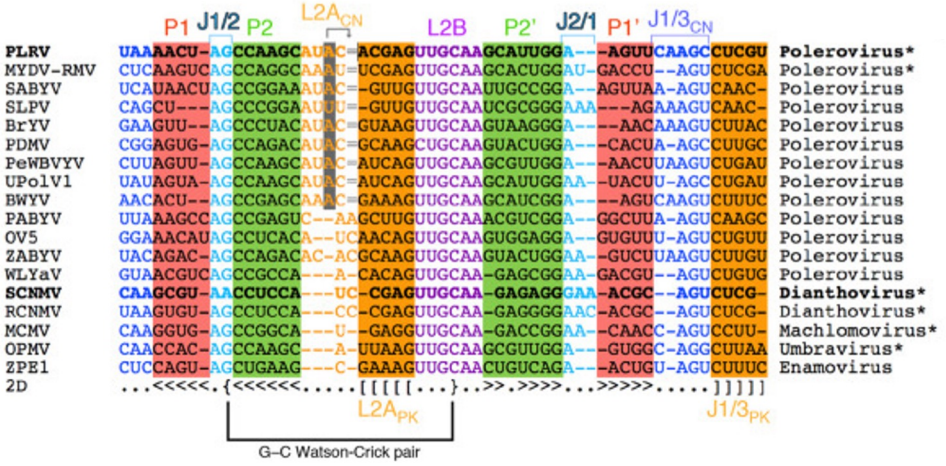
xrRNA - R1205



xrRNA from Potato leafroll virus (PLRV) (7JJU)



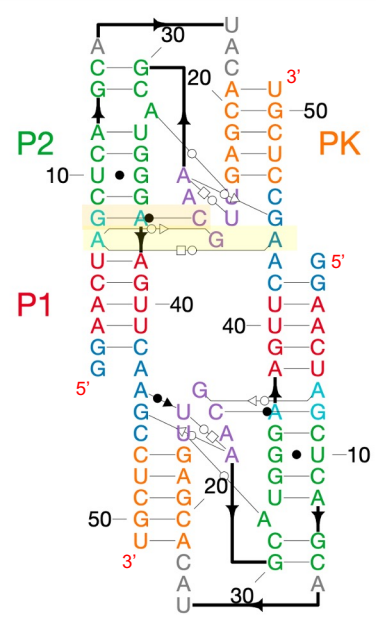
xrRNA from Sweet clover necrotic mosaic virus (SCNMV) (6D3P)



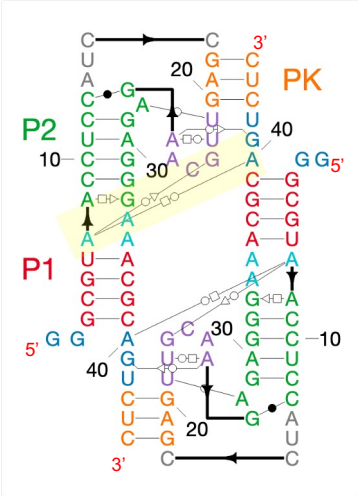
Alignment of xrRNA sequences from plant viruses (Steckelberg AL eta l. RNA. 2020)

```
>7JJU GGAACUAGCUCAGCAUAC---ACGAGUUGCAAGCAUGGG-A-AGUU-CAAGCCUCGU----------
>6D3P GGGCGUA-ACCUCCAUCC-----GAGUUGCAA-GAGAGGGAAACGC---AGUCUC------------
R1205 -A-AGUAC-CCUCCAAGCCCUACAGGUUGGAA-GAGGGG---GCU-AUCAGUCCUGUAGGCAGACUC
      ..(((((.{((((((....[[[[[[[[...}..)).)))))..))))).......]]]]]]]].......
```

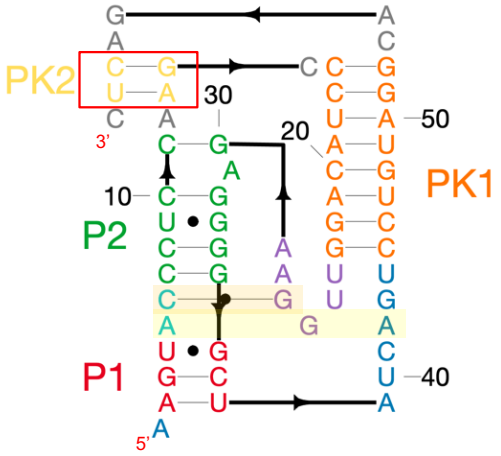
xrRNA - R1205



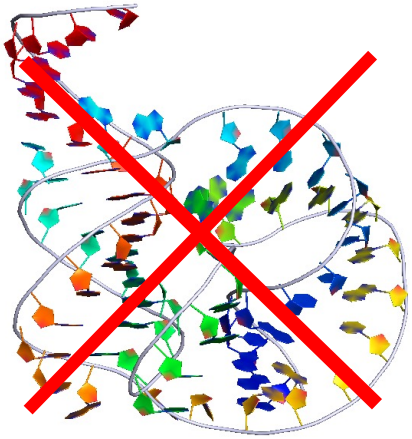
xrRNA from Potato leafroll virus (PLRV) (7JJU)



xrRNA from Sweet clover necrotic mosaic virus (SCNMV) (6D3P)



R1205



R1205_human_1

LDDT: 0.505
TMscore: 0.310
GDT_TS: 0.38

```
>7JJU GGAACUAGCUCAGCAUAC---ACGAGUUGCAAGCAUGGG-A-AGUU-CAAGCCUCGU----------
>6D3P GGGCGUA-ACCUCCAUCC-----GAGUUGCAA-GAGAGGGAAACGC---AGUCUC------------
R1205 -A-AGUAC-CCUCCAAGCCCUACAGGUUGGAA-GAGGGG---GCU-AUCAGUCCUGUAGGCAGACUC
      ..(((((.{(((((((.[[.[[[[[[[[...}..)).)))))..))))).....]]]]]]]]....]].
```

Summary of GuangzhouRNA in CASP16

What we did right:

1. Better 2D (literature search and deep MSA)
2. Homology modelling, and fragment assemble
3. Better selection based on 2D structure

What went wrong:

For some cases, even through we can get precise 2D of the target from deep MSA we believe, we still cannot get the prediction result which strictly follow the 2D we give.

Future directions:

1. Better MSA tools
2. More RNA/DNA Structures

Acknowledgements

CASP16 Modeler

Bowen Xiao
Yaohuang Shi
Baowei Huang
Jiajun Lin
Jiaxing Zhao

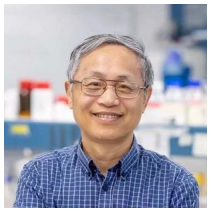
GuangzhouRNA-human (183)
GuangzhouRNA_AI (317)
GuangzhouRNA-meta (417)

Collaborators

David Lilley

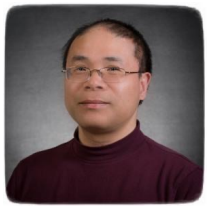
Eric Westhof

Rhiju Das



Yaoqi Zhou

Zhichao Miao

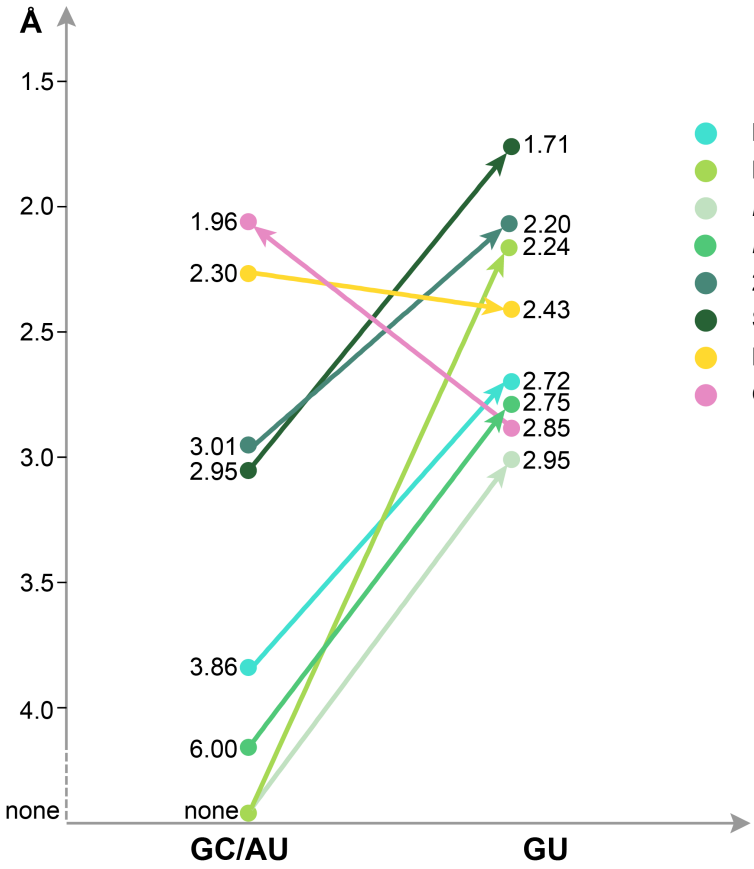


Shi-Jie Chen

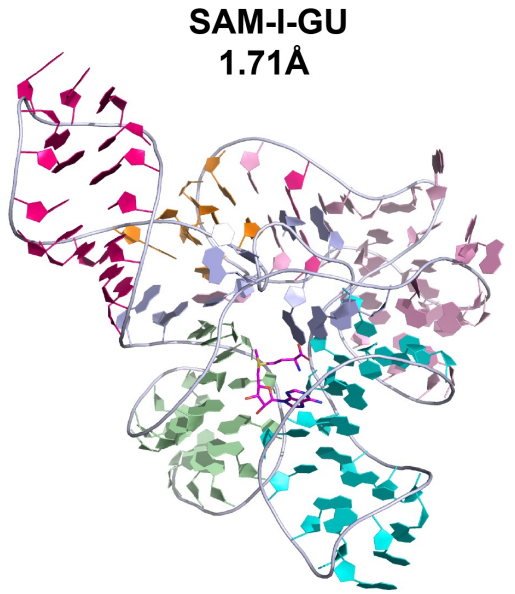
中山大學孙逸仙纪念医院
SUN YAT-SEN MEMORIAL HOSPITAL,SUN YAT-SEN UNIVERSITY

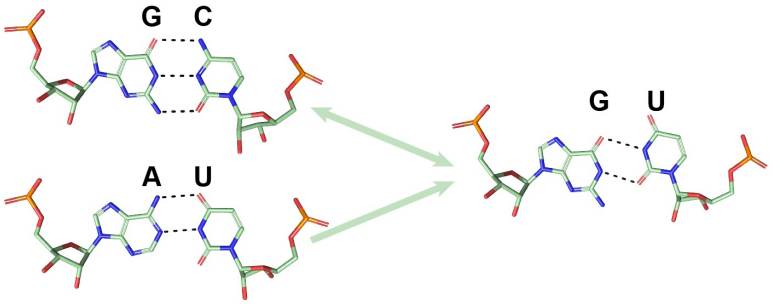
Thanks

A general strategy for engineering GU base pairs to facilitate RNA crystallization



- LINE-1-GC → GU
- Broccoli-AU → GU
- *Hm*Kt-7-GC → GU
- *Hm*Kt7-19ntX3-GC → GU
- 2'-dG-III-AU → GU
- SAM-I-GC → GU
- MTR1-GC/AU → GU
- OR4K15-GU → GC





Nucleic Acids Research, 2024, 1–13
https://doi.org/10.1093/nar/gkae1218
Methods

A general strategy for engineering GU base pairs to facilitate RNA crystallization

Yangyi Ren[1,†], Xiaowei Lin[1,2,3,†], Wenjian Liao[1,2], Xuemei Peng[1], Jie Deng[1], Zhe Zhang[4], Jian Zhan[4], Yaoqi Zhou[4], Eric Westhof[5,6], David M. J. Lilley[7], Jia Wang[8,*] and Lin Huang[1,*]

Nucleic Acids Research (In Press)