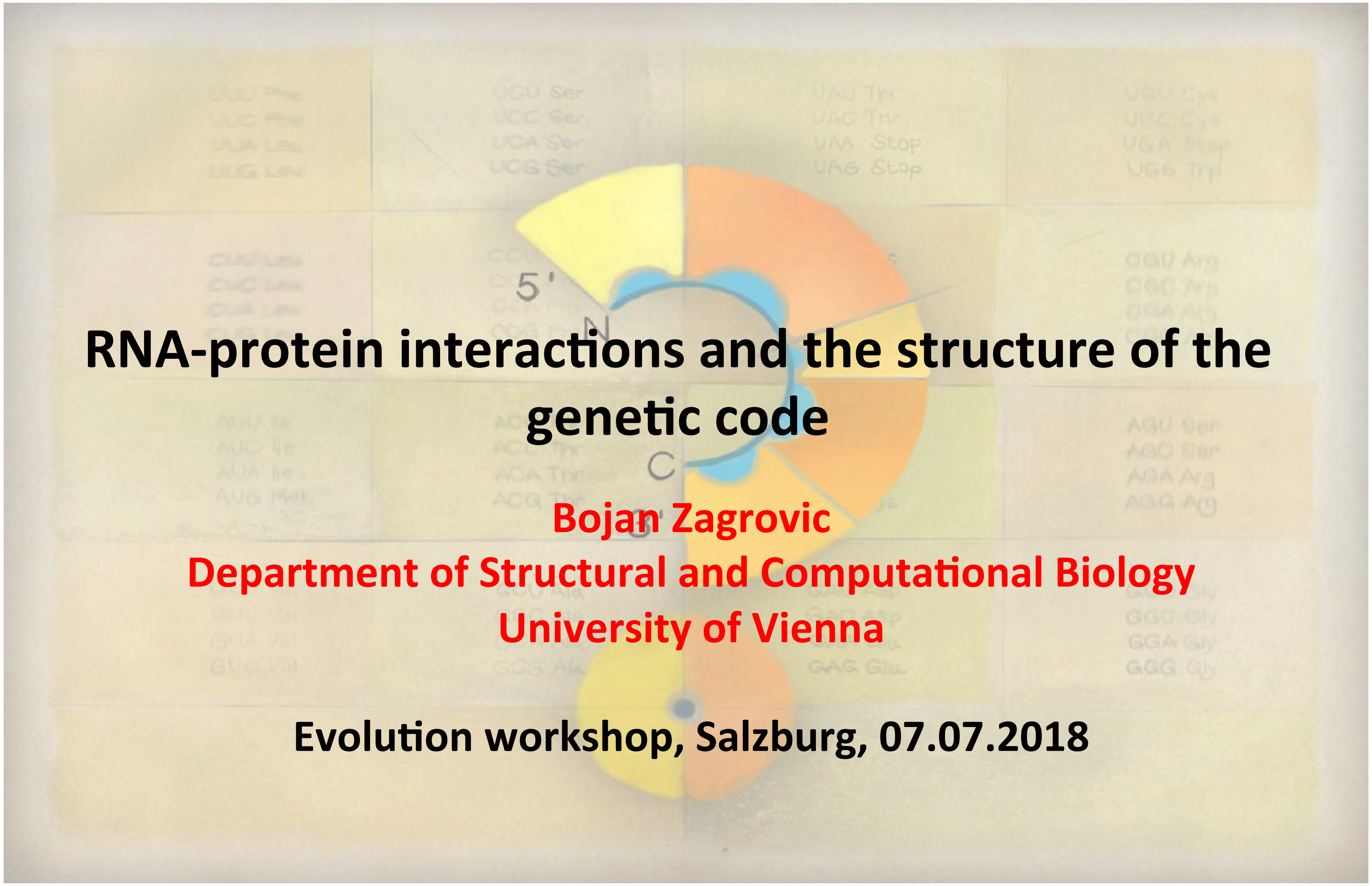


RNA-protein interactions and the structure of the genetic code

Bojan Zagrovic
Department of Structural and Computational Biology
University of Vienna

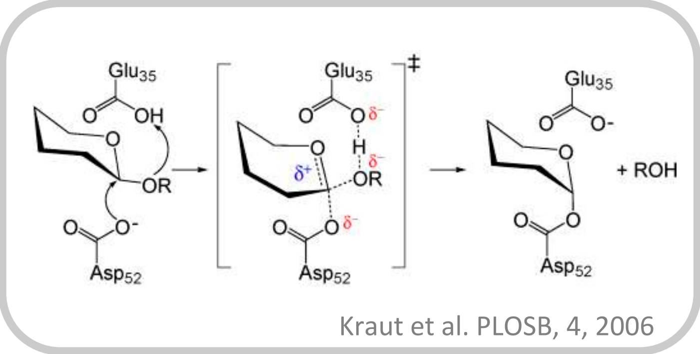
Evolution workshop, Salzburg, 07.07.2018

Physicochemical complementarity

one of the most powerful paradigms in biology

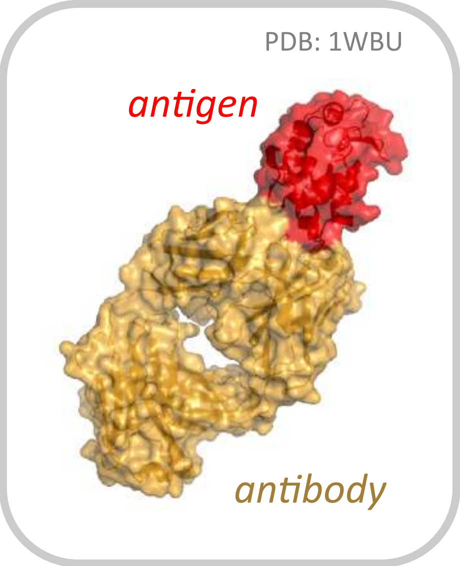

PDB: 1QPH

gene duplication



Kraut et al. PLOSB, 4, 2006

enzymatic catalysis


PDB: 1WBU

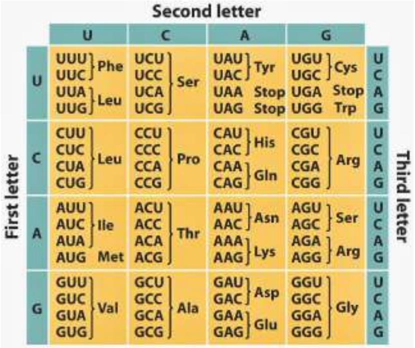
immune response

a novel complementarity?

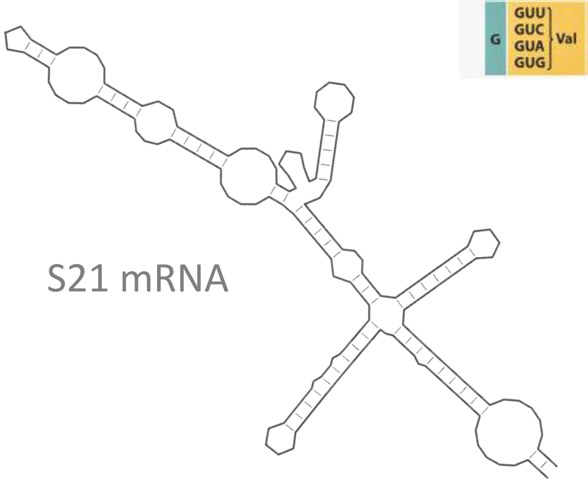
messenger RNA-protein relationship:

the cornerstone of molecular biology

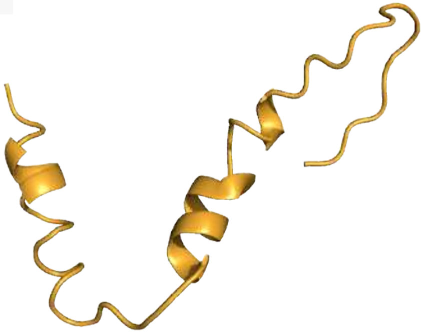
linked by the **universal genetic code**



	Second letter								
First letter	U		C		A		G		Third letter
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U
	UUC		UCC		UAC		UGC		C
	UUA	Leu	UCA		UAA	Stop	UGA	Stop	A
	UUG		UCG		UAG	Stop	UGG	Trp	G
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	Gln	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	Lys	AGA	Arg	A
	AUG	Met	ACG		AAG		AGG		G
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	Glu	GGA		A
	GUG		GCG		GAG		GGG		G



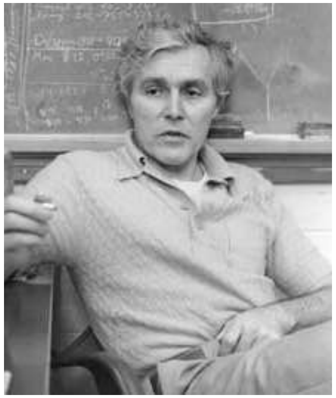
S21 mRNA



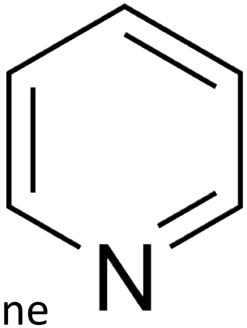
ribosomal protein S21
(PDB: 4GD2)

physicochemical connection?

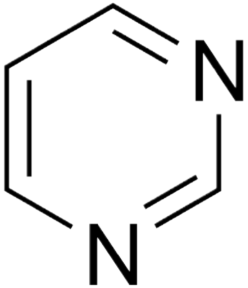
What is the affinity of amino acids for pyrimidines?



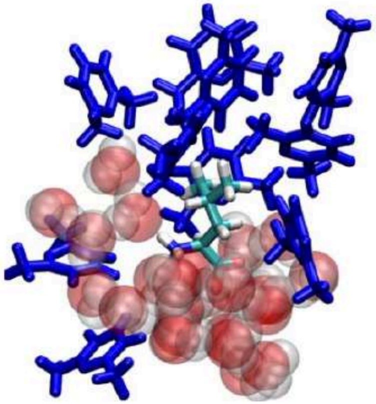
Carl Woese



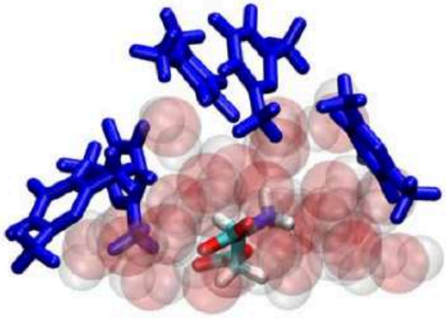
pyridine



pyrimidine



LEU in DMP:H_2O



ASP in DMP:H_2O

*DMP: 2,6-dimethylpyridine

C	L	F	W	I	M	P	V	T	A	S	Y	H	R	Q	G	N	K	D	E
4.3	4.4	4.5	4.9	5.0	5.0	6.1	6.2	6.2	6.5	7.5	7.7	7.9	8.6	8.9	9.0	9.6	10.2	12.2	13.6

a. u.

Woese CR et al., **PNAS**, 55:966-974 (1966)
Mathew DC & Luthey-Schulten Z, **J Mol Evol**, 66:519-528 (2008)

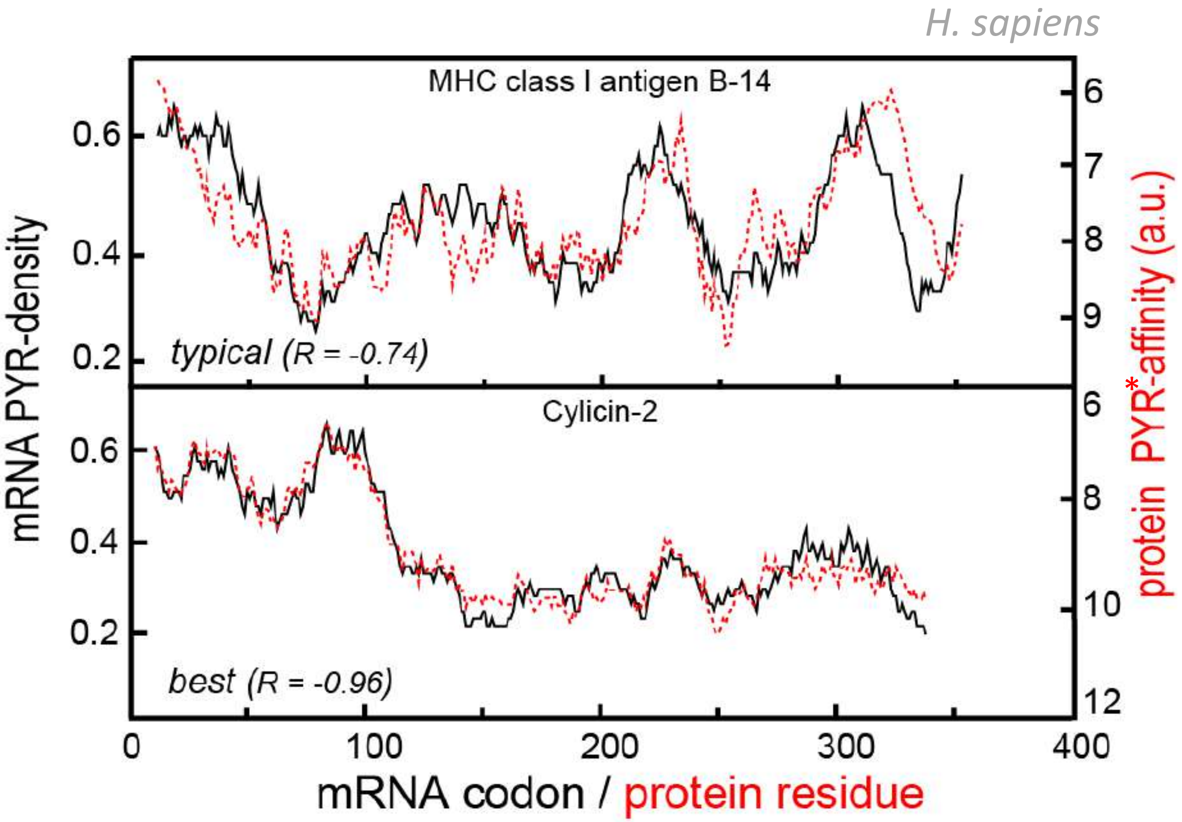
THE CENTRAL FINDING

PYR-density = 0.7

mRNA ...UGGGCC CUGGGCUUCUACCCU GCGGAG...

protein ...Trp - Ala - Leu - Gly - Phe - Tyr - Pro - Ala - Glu...

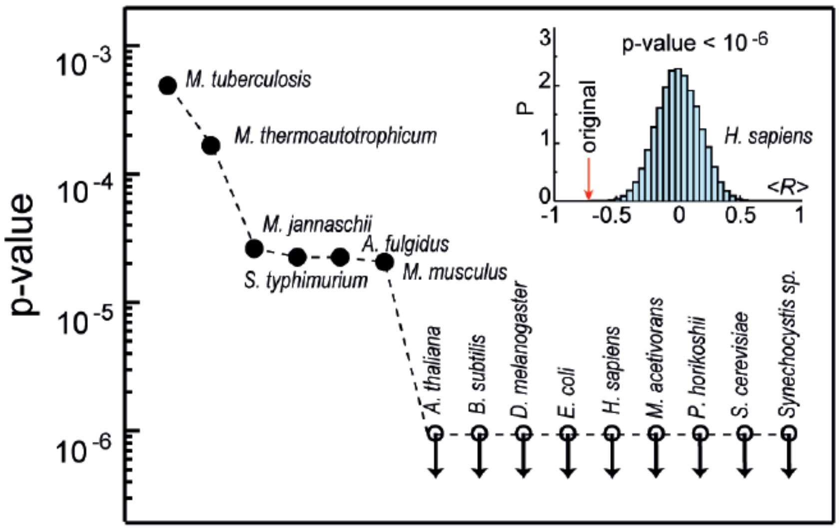
PYR*-affinity = 6.3



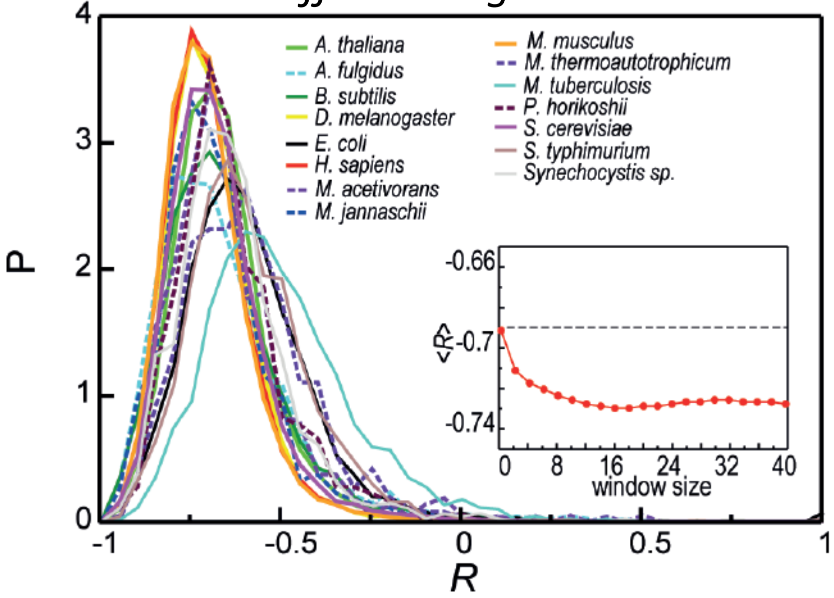
Hlevnjak M, Polyansky AA & Zagrovic B, **Nucleic Acids Research**, 40, 2012
Polyansky AA, Hlevnjak M & Zagrovic B, **RNA Biology**, 10, 2013

A statistically robust finding of universal validity

genetic code shuffling



different organisms



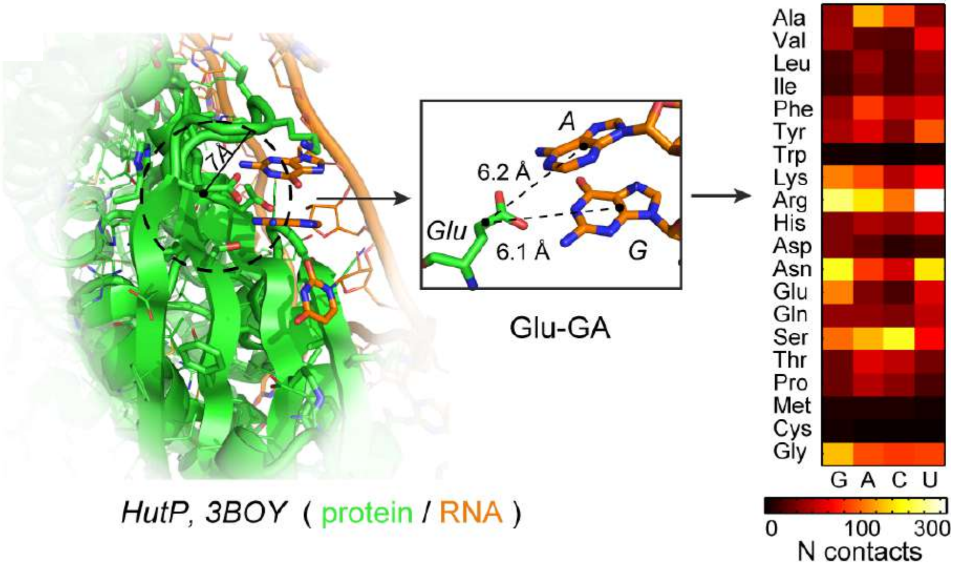
Hlevnjak M, Polyansky AA & Zagrovic B, **Nucleic Acids Research**, 40, 2012
Polyansky AA, Hlevnjak M & Zagrovic B, **RNA Biology**, 10, 2013

What about natural bases and, in particular, purines?

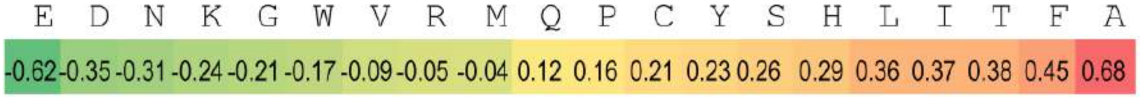
- 3D structures of ~300 protein-RNA complexes

free energy proxies from contact statistics:

$$\varepsilon^{ij} = -ln\frac{N^{ij}_{obs}}{N^{ij}_{exp}} = -ln\frac{N^{ij}_{obs}}{X_iX_jN^{TOT}_{obs}}$$

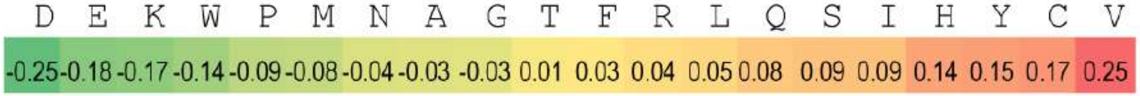


HutP, 3BOY (protein / RNA)



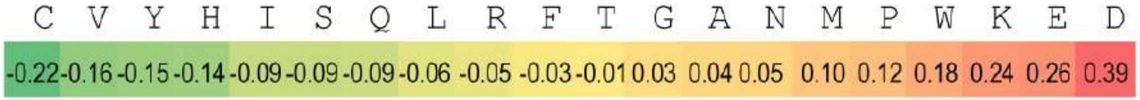
E	D	N	K	G	W	V	R	M	Q	P	C	Y	S	H	L	I	T	F	A
-0.62	-0.35	-0.31	-0.24	-0.21	-0.17	-0.09	-0.05	-0.04	0.12	0.16	0.21	0.23	0.26	0.29	0.36	0.37	0.38	0.45	0.68

G-preference scale



D	E	K	W	P	M	N	A	G	T	F	R	L	Q	S	I	H	Y	C	V
-0.25	-0.18	-0.17	-0.14	-0.09	-0.08	-0.04	-0.03	-0.03	0.01	0.03	0.04	0.05	0.08	0.09	0.09	0.14	0.15	0.17	0.25

PUR-preference scale



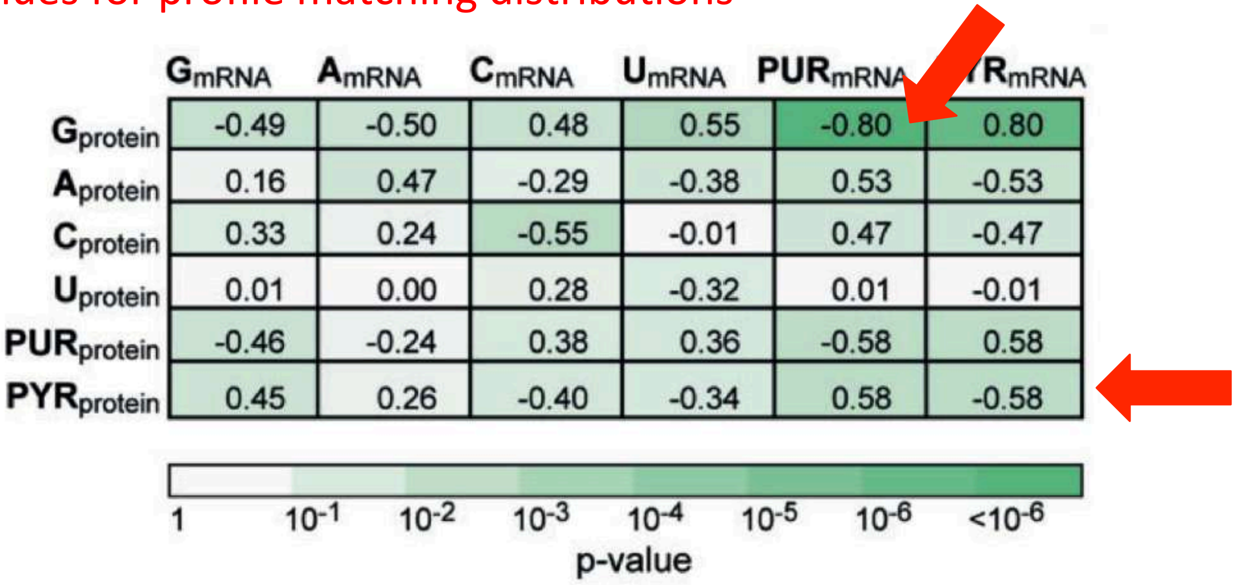
C	V	Y	H	I	S	Q	L	R	F	T	G	A	N	M	P	W	K	E	D
-0.22	-0.16	-0.15	-0.14	-0.09	-0.09	-0.09	-0.06	-0.05	-0.03	-0.01	0.03	0.04	0.05	0.10	0.12	0.18	0.24	0.26	0.39

PYR-preference scale

*in units of k_BT

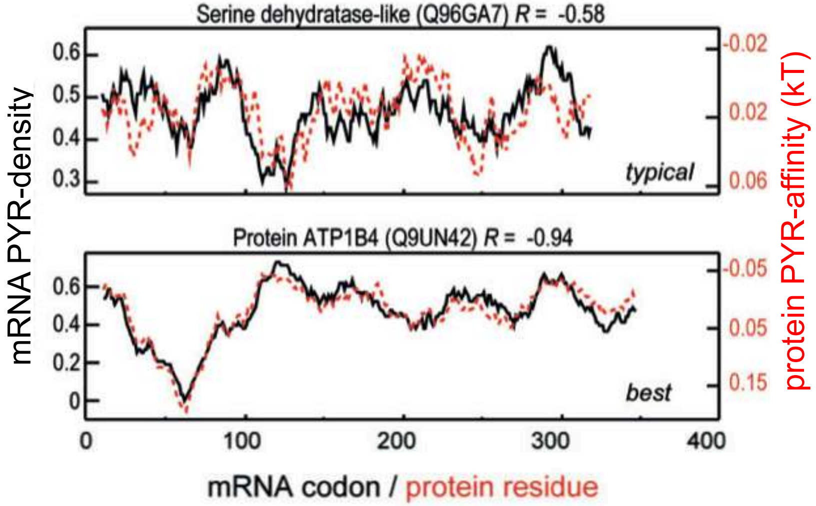
Polyansky AA & Zagrovic B, **NAR**, 18: 8434-8443 (2013)

median R-values for profile matching distributions

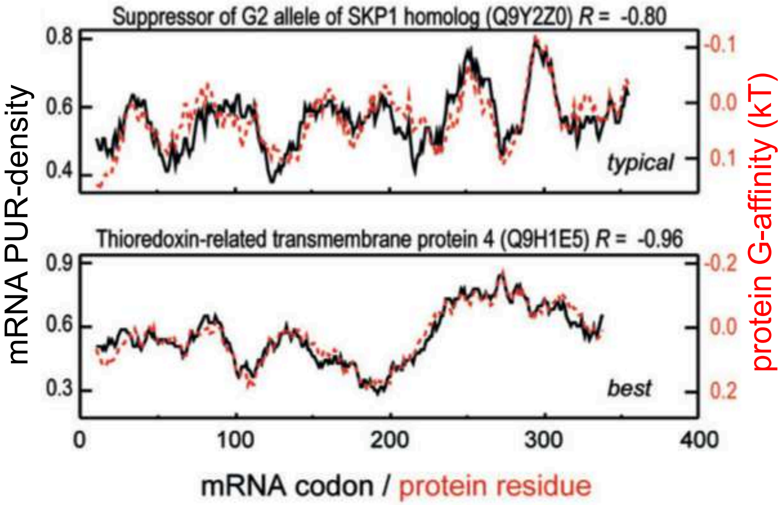


	G_{mRNA}	A_{mRNA}	C_{mRNA}	U_{mRNA}	PUR_{mRNA}	YR_{mRNA}
$\text{G}_{\text{protein}}$	-0.49	-0.50	0.48	0.55	-0.80	0.80
$\text{A}_{\text{protein}}$	0.16	0.47	-0.29	-0.38	0.53	-0.53
$\text{C}_{\text{protein}}$	0.33	0.24	-0.55	-0.01	0.47	-0.47
$\text{U}_{\text{protein}}$	0.01	0.00	0.28	-0.32	0.01	-0.01
$\text{PUR}_{\text{protein}}$	-0.46	-0.24	0.38	0.36	-0.58	0.58
$\text{PYR}_{\text{protein}}$	0.45	0.26	-0.40	-0.34	0.58	-0.58

mRNA PYR-density vs. protein PYR-affinity

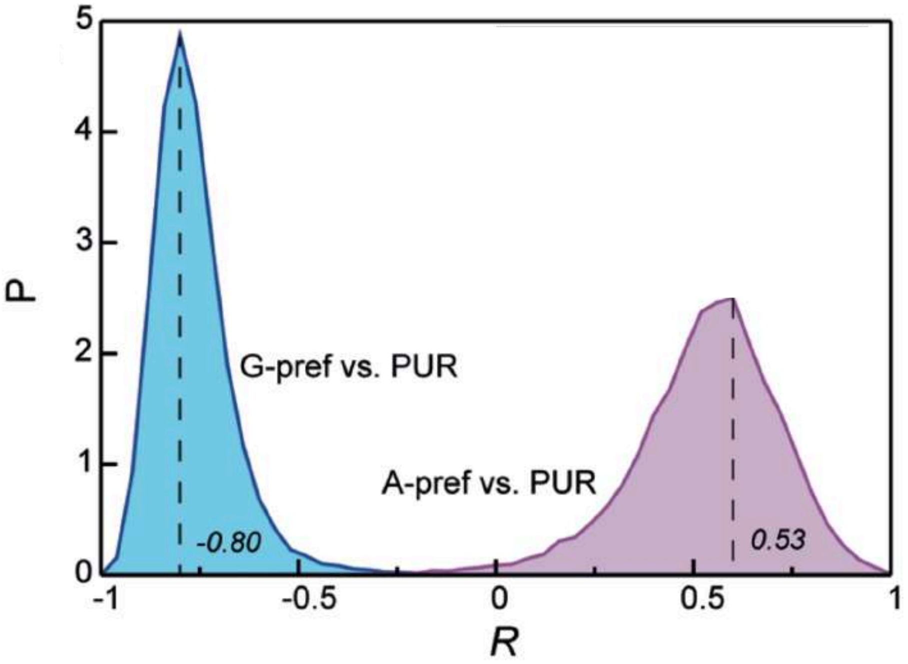


mRNA PUR-density vs. protein G-affinity



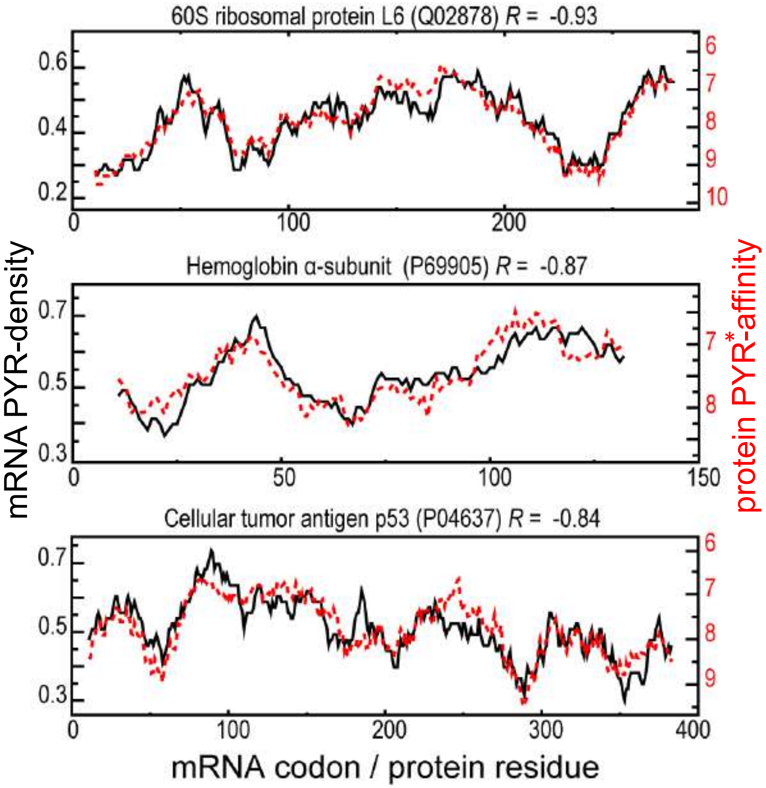
An important exception:

ADE-preferring amino acids tend to be encoded by PYR-rich stretches and vice versa

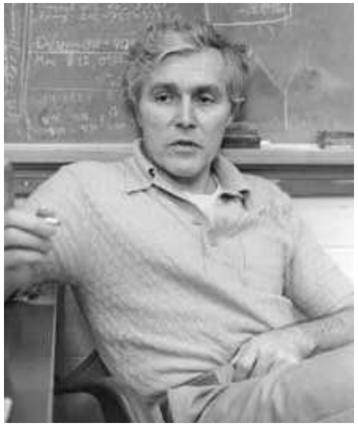


Polyansky AA & Zagrovic B, **NAR**, 18: 8434-8443 (2013)

Hypotheses



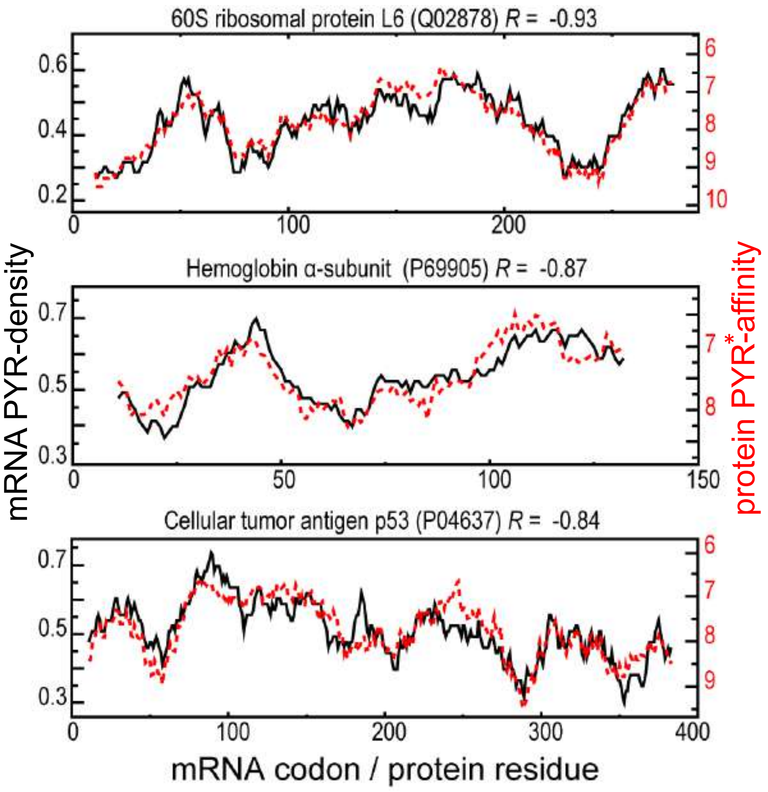
1. our findings support the possibility of a connection between nucleobase/amino-acid binding and genetic encoding (such as in the **direct templating** mechanism of ancient translation)



stereochemical hypothesis
(Woese CR, 1965):
"coding from binding"

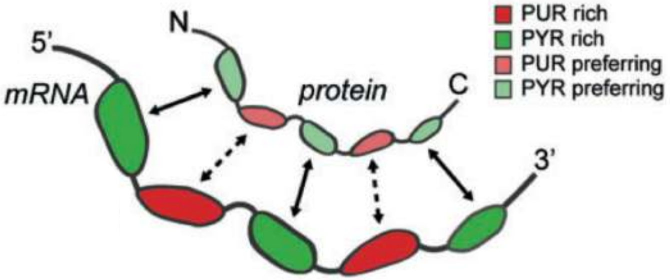
Hlevnjak, Polyansky & Zagrovic, **NAR**, 40, 2012; Polyansky & Zagrovic, **NAR**, 41, 2013; Polyansky AA, Hlevnjak M & Zagrovic B, **RNA Biology,** 10, 2013; Hajnic, Osorio & Zagrovic, **NAR**, 42, 2014; Hajnic, Osorio & Zagrovic, **PCCP**, 17, 2015; de Ruiter & Zagrovic, **NAR**, 43, 2015; Hlevnjak & Zagrovic, **NAR**, 43, 2015; Bartonek & Zagrovic**, PLOS CB**, 13, 2017

Hypotheses



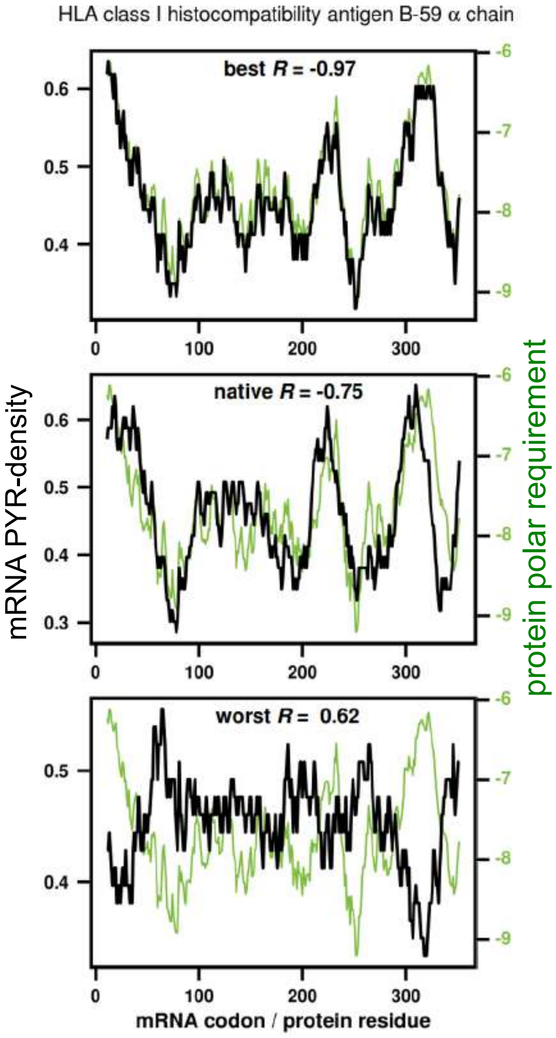
1. our findings support the possibility of a connection between nucleobase/amino-acid binding and genetic encoding (such as in the **direct templating** mechanism of ancient translation)

2. today's mRNAs and cognate proteins may be **complementary** to each other and bind, especially if unstructured. Complementarity is negatively regulated by mRNA ADE content.

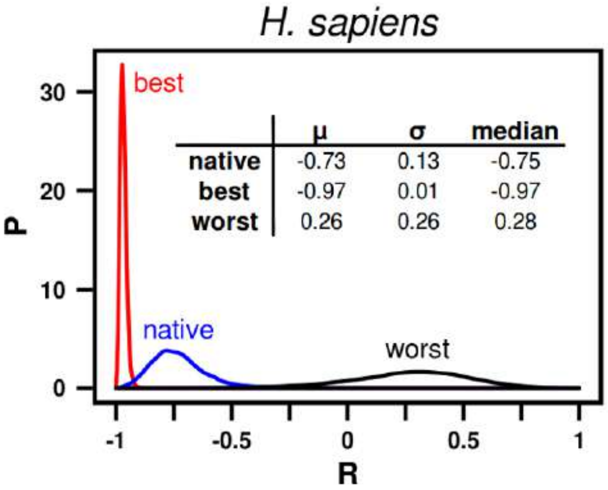


Hlevnjak, Polyansky & Zagrovic, **NAR**, 40, 2012; Polyansky & Zagrovic, **NAR**, 41, 2013; Polyansky AA, Hlevnjak M & Zagrovic B, **RNA Biology,** 10, 2013; Hajnic, Osorio & Zagrovic, **NAR**, 42, 2014; Hajnic, Osorio & Zagrovic, **PCCP**, 17, 2015; de Ruiter & Zagrovic, **NAR**, 43, 2015; Hlevnjak & Zagrovic, **NAR**, 43, 2015; Bartonek & Zagrovic**, PLOS CB**, 13, 2017

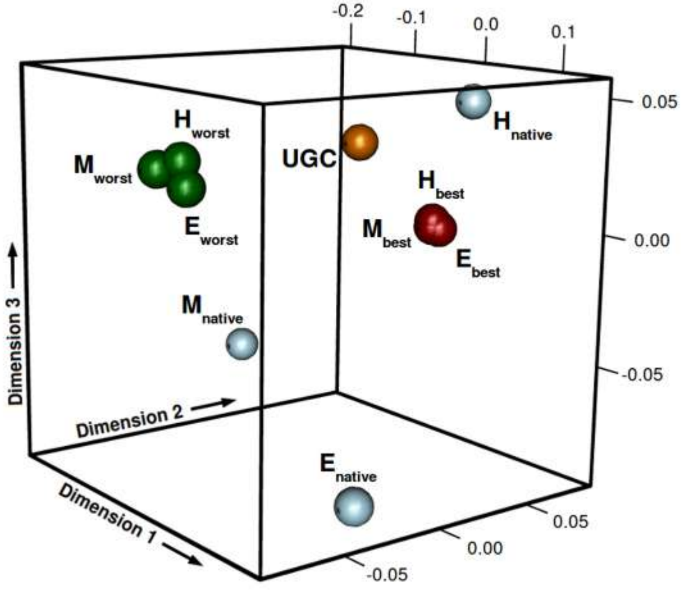
mRNA recoding vs. cognate matching



Woese pyridine scale



	μ	σ	median
native	-0.73	0.13	-0.75
best	-0.97	0.01	-0.97
worst	0.26	0.26	0.28



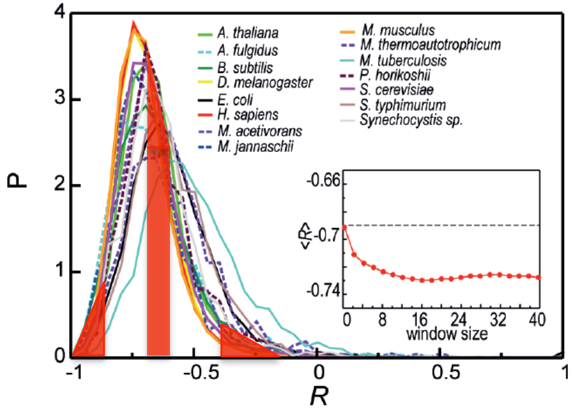
codon usage bias

Hlevnjak M & Zagrovic B, **NAR**, 43, 3012-3021 (2015)

Functional significance

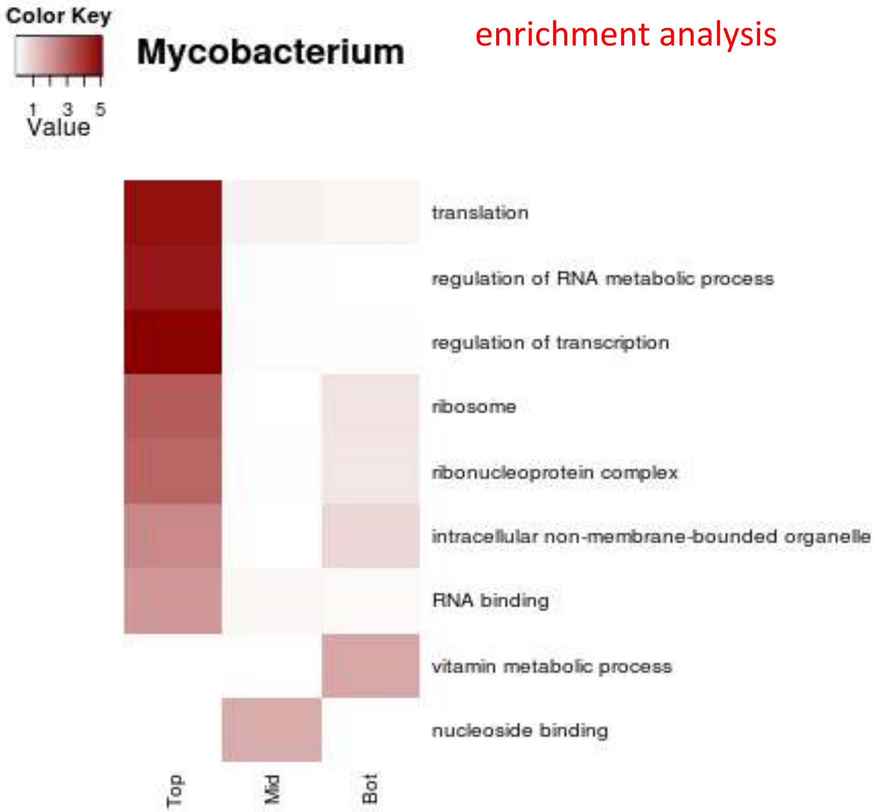
?

- Potentially relevant in many areas of RNA/protein biology:
  - translation regulation
  - viral capsid assembly
  - structure of non-membrane-bound cellular compartments
  - protein interactions with DNA, lncRNA etc.



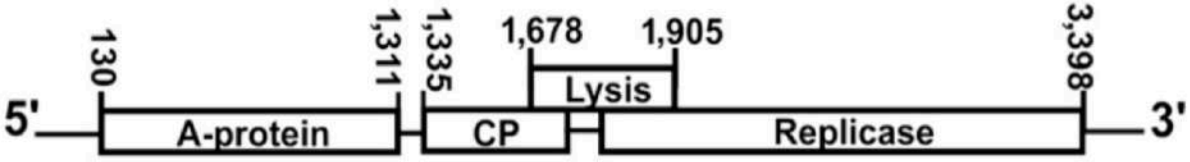
PR/PYR matching

Polyansky AA, Hlevnjak M & Zagrovic B, *in preparation*



enrichment analysis

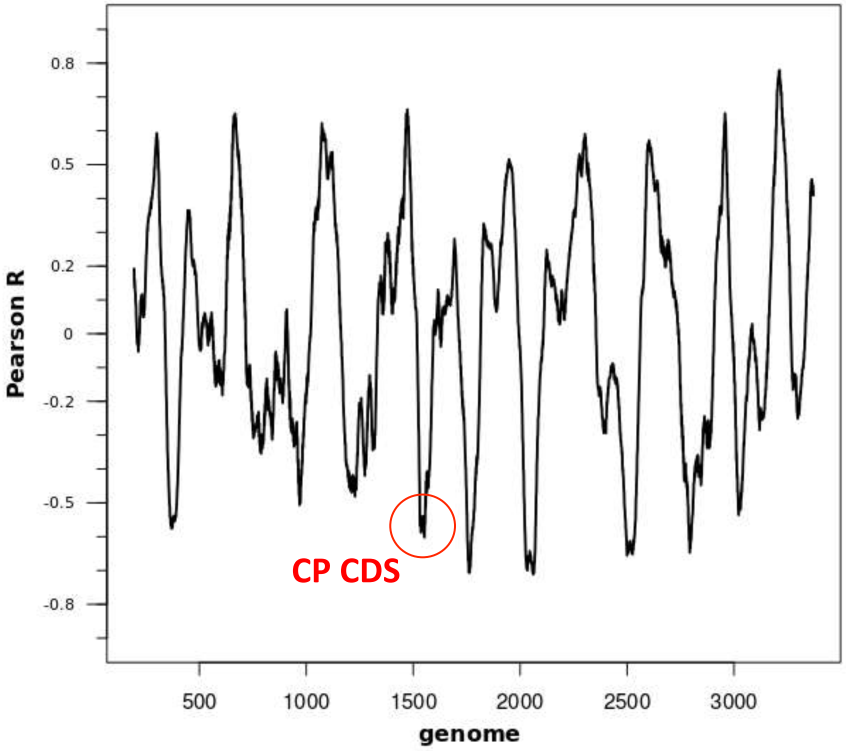
A case study: assembly of the MS2 bacteriophage



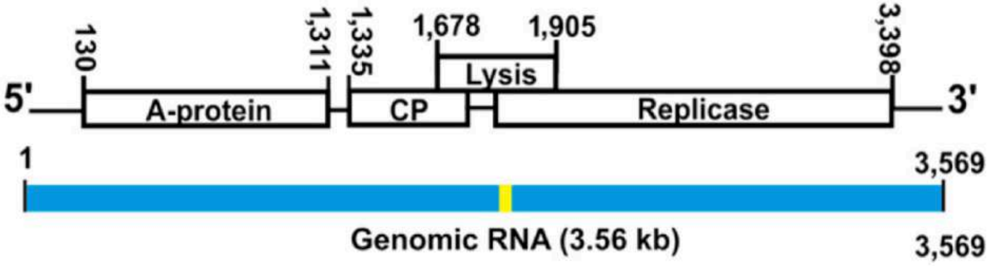
IDEA:

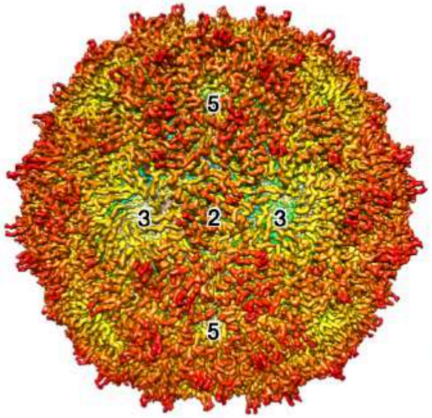
genome %PYR profile

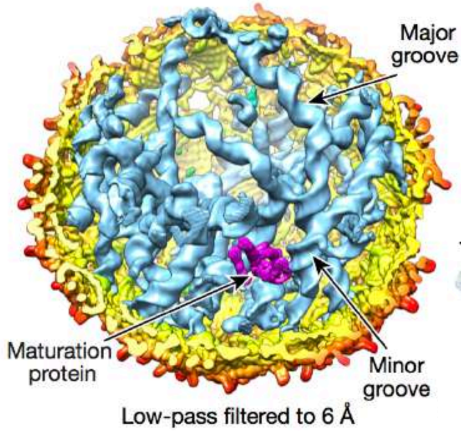
coat protein PYR affinity profile

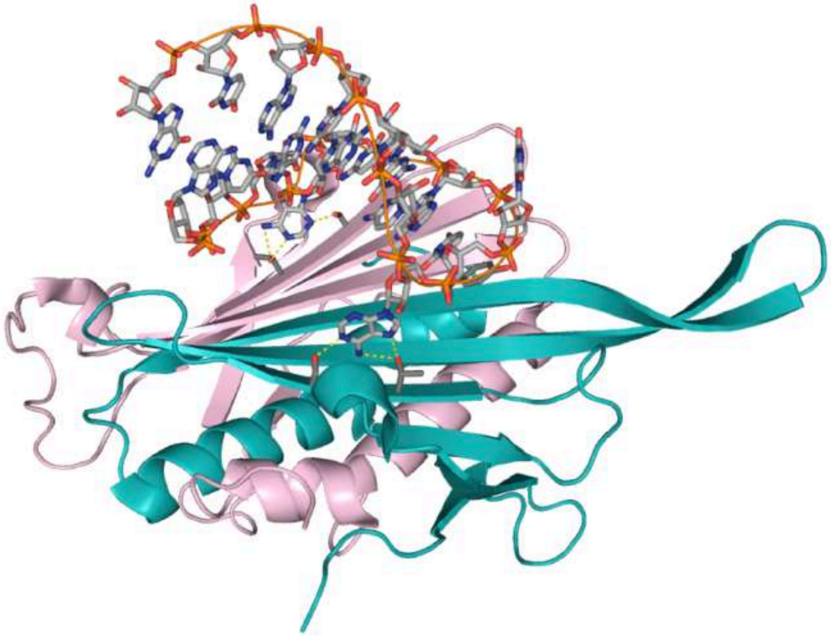


Comparison against a 3.6 Å cryo EM structure of MS2





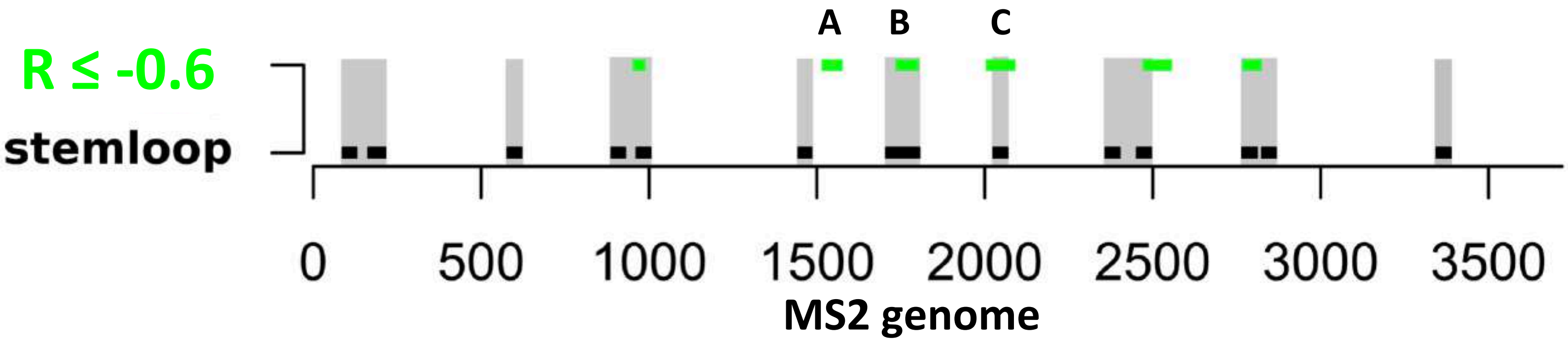


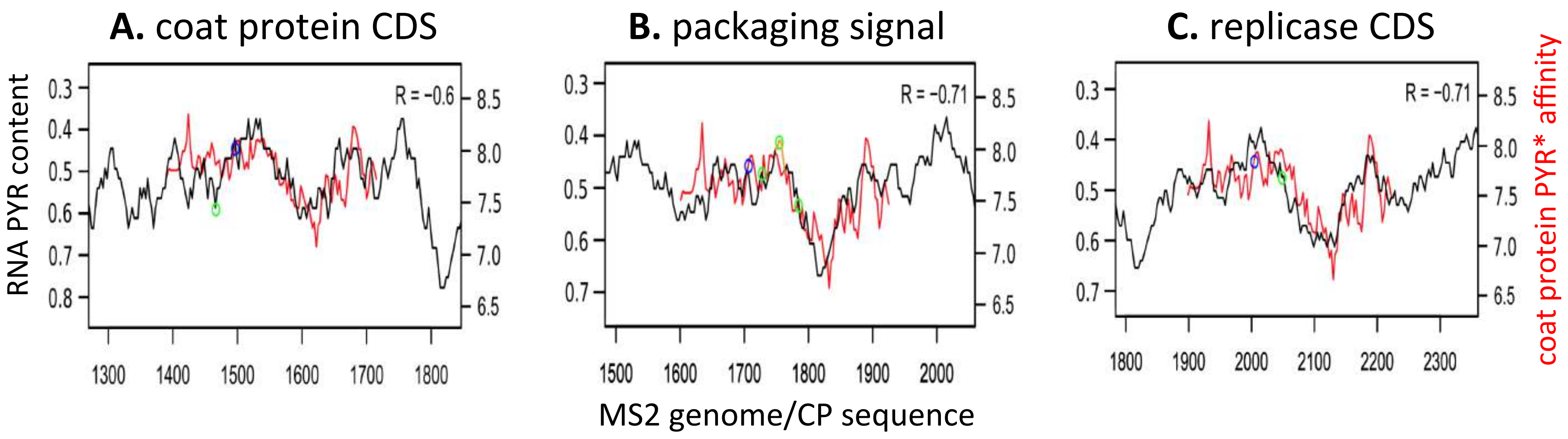


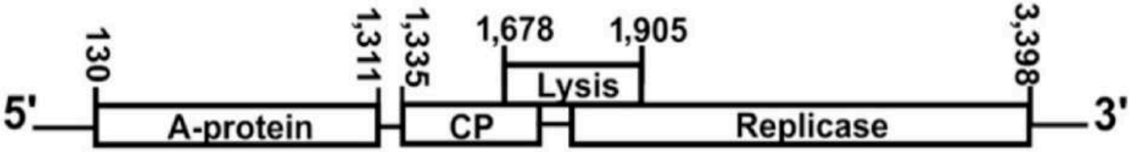
16 high-resolution RNA stem loops with capsid protein

Dai et al. Nature, 541, 2017

Profile-based predictions match experimental locations of stem loops

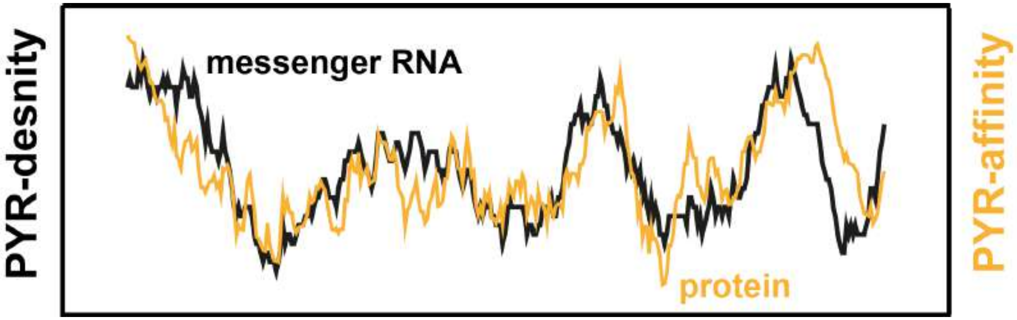






Poetsch F, Hlevnjak M, Polyansky AA & Zagrovic B, in preparation

mRNA-PROTEIN COMPLEMENTARITY



Origin of the genetic code
coding from binding?

RNA-protein interactions
a novel mechanistic perspective

New paradigms?
- translation regulation
- viral capsid assembly
- cellular networks…

OPEN CHALLENGES

geometry; ΔGs; functional significance; experimental testing

Thanks

Present members

- **Anton A. Polyansky**
- Lukas Bartonek
- Markus Fleck
- **Florian Pötsch**
- Daniel Braun
- Daniel Hoffmann

Former members

- **Matea Hajnic**
- **Anita de Ruiter**
- **Mario Hlevnjak**
- **Juan Osorio Iregui**
- **Theres Friesacher**
- Gerald Platzer
- Antonija Kuzmanic
- Drazen Petrov
- Daniela Kruschel
- Christian Margreitter
- Matija Piskorec
- Mijo Simunovic
- Ruben Zubac
- Natalie Romanov
- Lily Chen
- Raphael Peer
- Thomas Malzac
- Bianca Mladek
- Maud Formanek
- Alexander Beribisky
- Mathias Kreuter

Collaborators

- Robert Konrat, MFPL
- Renee Schroeder, MFPL
- Kristina Djinovic, MFPL
- Philippe H. Huenenberger, ETH Zurich
- Navraj S. Pannu, Leiden University
- Carrie Bernecky, ISTA
- Chris Oostenbrink, BOKU Vienna

Funding

UNITY THROUGH KNOWLEDGE FUND

nzz

FWF

erc