

RNA-based drugs: discovery and design

From aptamers to oligonucleotides: a perspective for
in silico approaches

Fabrice Leclerc

Institute for Integrative Biology of the Cell (I2BC)

CNRS-CEA-Univ. Paris Saclay

team: RNA sequence structure function, Dept. Genome Biology

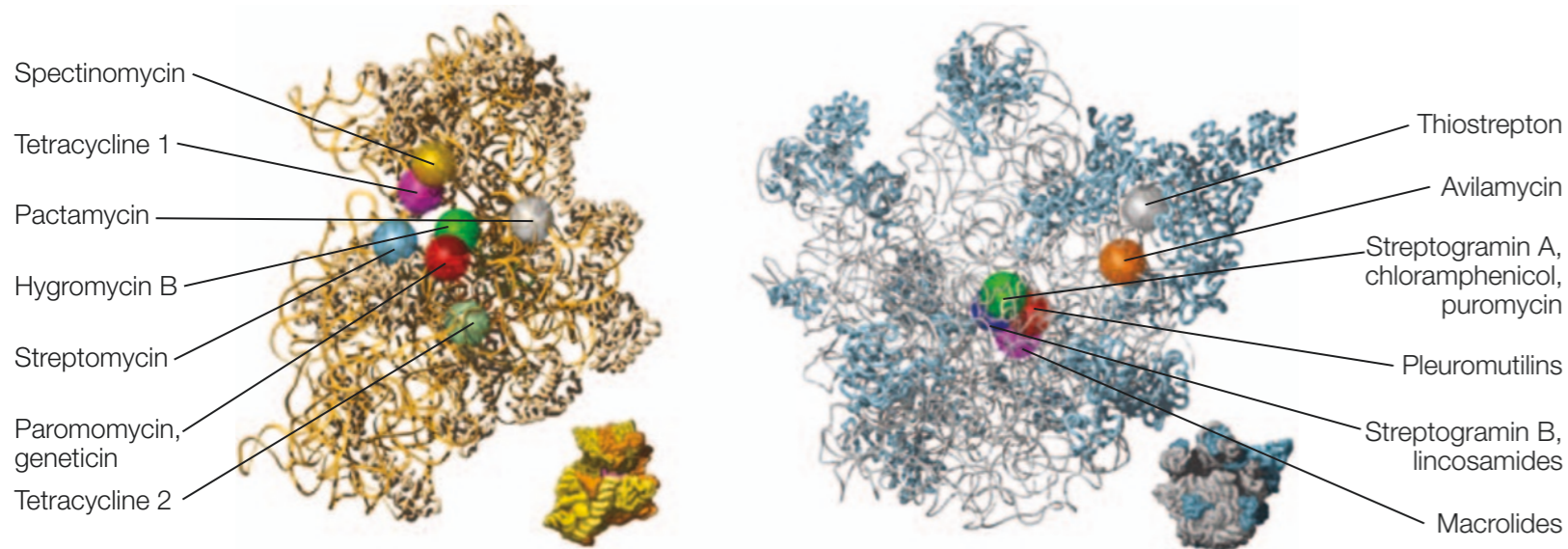
Outline

- **Introduction to RNA Therapy and Strategies: Target/Drug**
 - RNA as a target (e.g. ribosome, other folded RNAs, mRNAs, ...)
 - RNA as a drug: silencing, interference, editing, binding, etc
- **Approaches & Technologies: RNA(DNA) as a drug/tool**
 - from aptamers to SOMAmers (drug target: protein, or RNA, etc)
 - from antisense oligonucleotides (ASO) to gapmers (drug target: RNA)
- **Between pseudo-aptamers and ASOs:** in silico design of short oligonucleotides against protein targets

RNA as a drug target

structure-dependent approaches

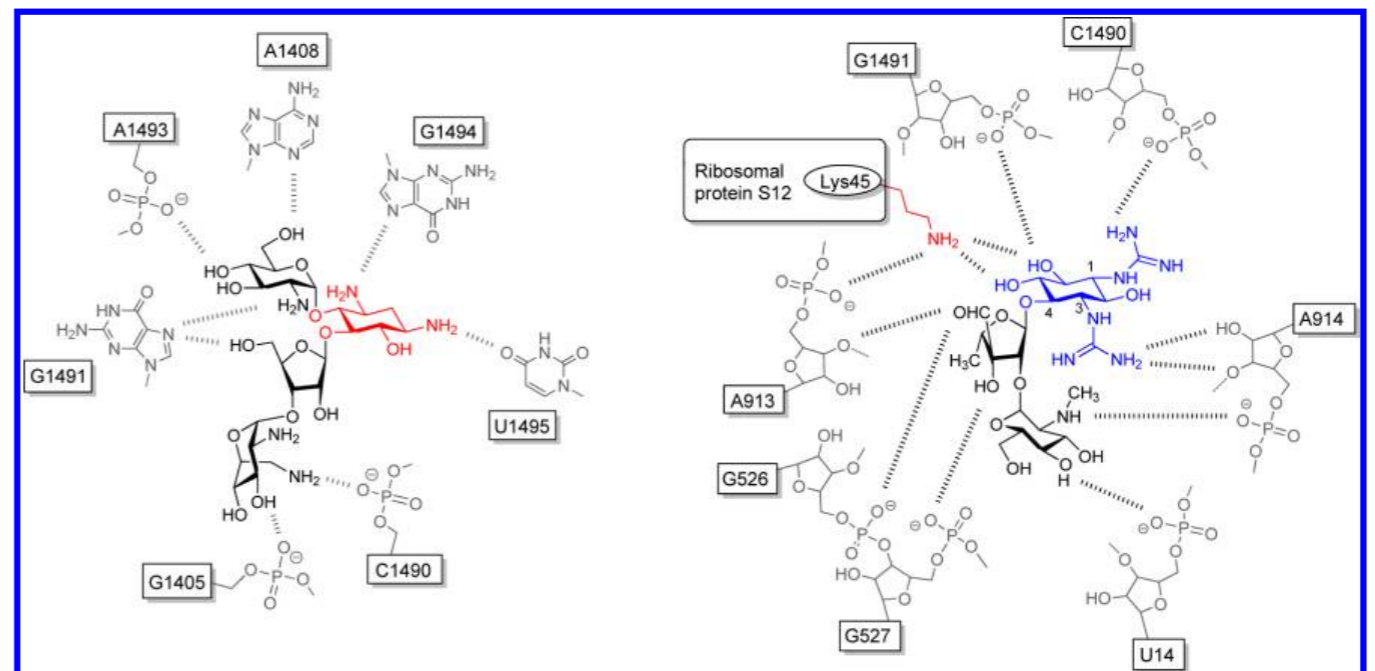
RNAs as targets: from cellular machineries to folded RNAs



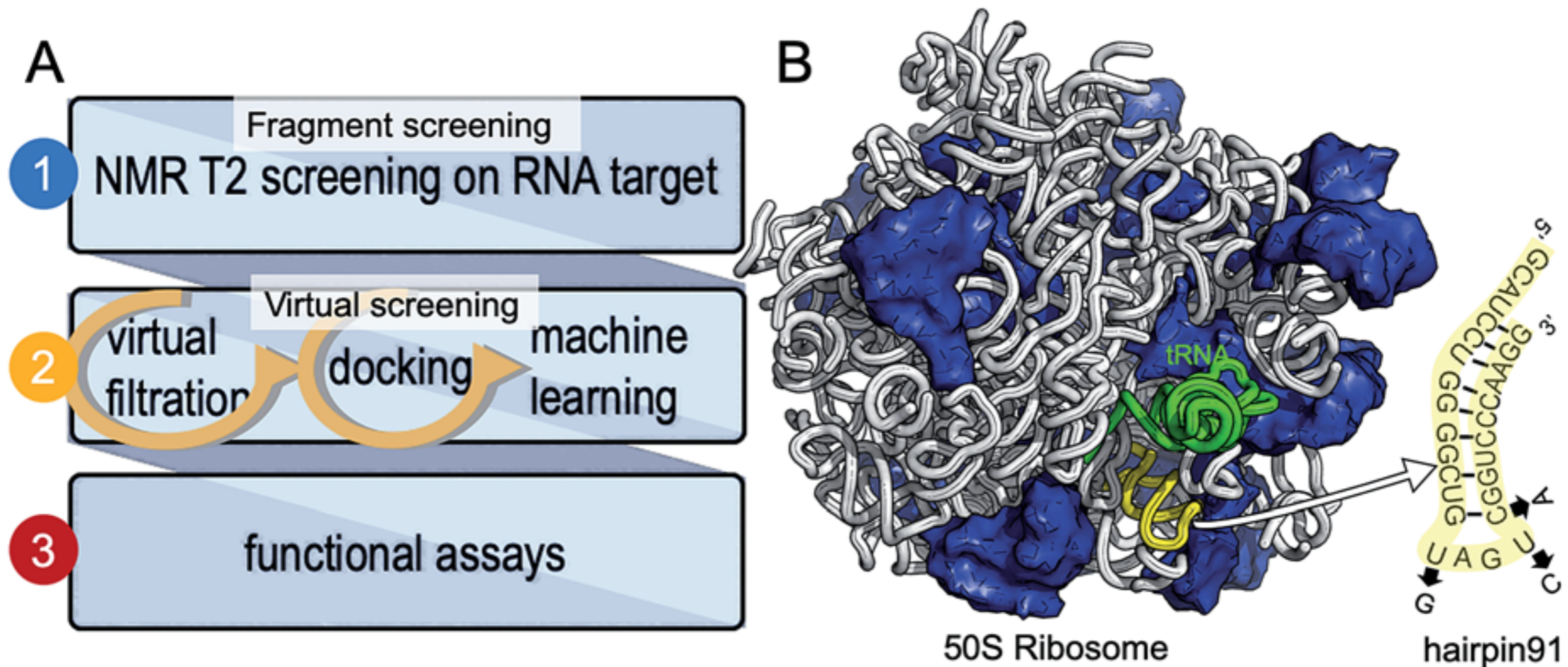
- ribosome as translation machinery
- ribosome biogenesis
- ribosome assembly ...

Poehlsgaard & Douthwaite, 2005

- aminoglycosides bound to 16S rRNA
- other RNA targets:
 - viral RNAs, riboswitch, ribozyme, etc
- Off-target effects (OTEs) & toxicity:
 - mitochondrial rRNA



Ligand-Drug Design against RNA targets



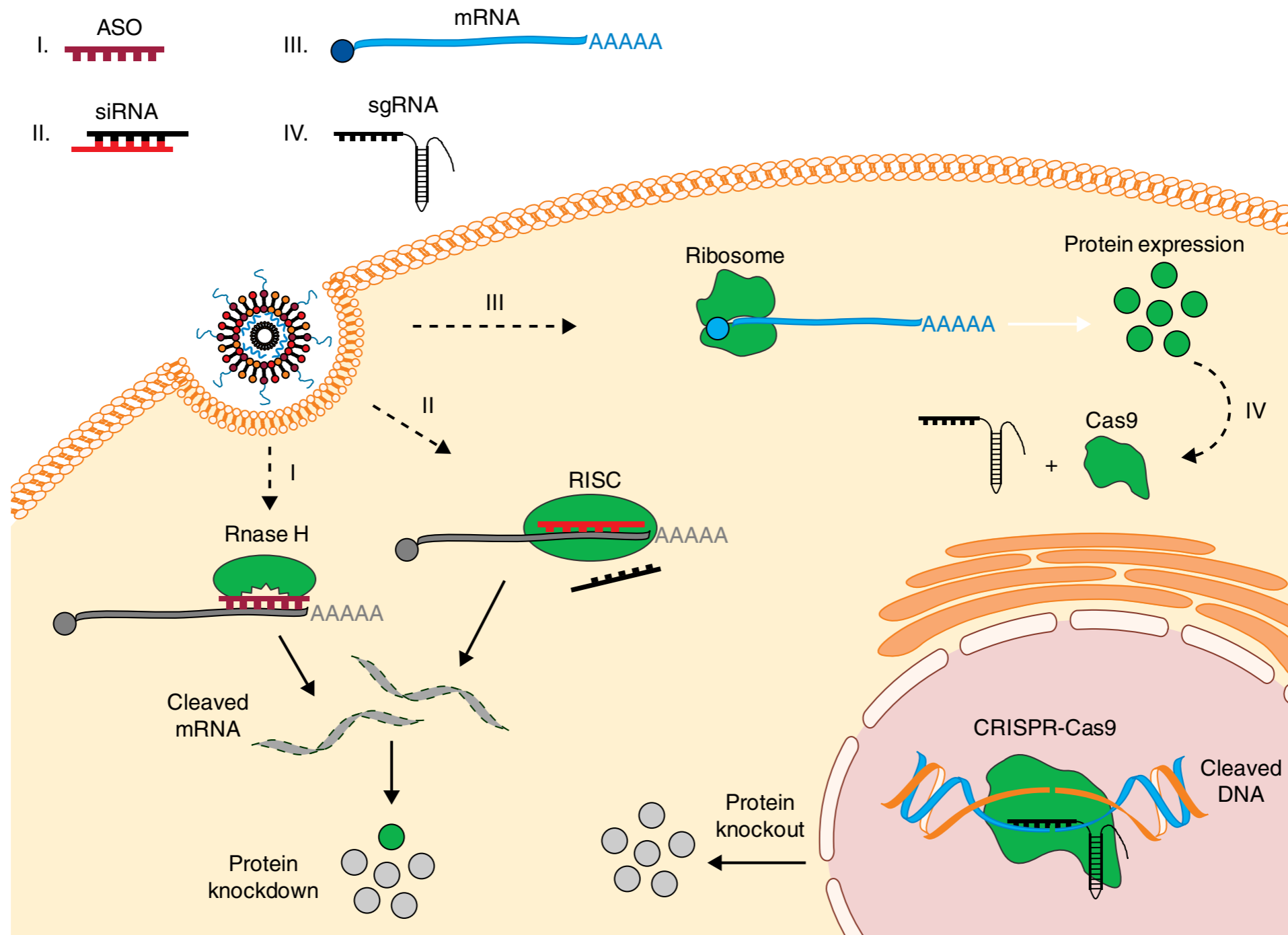
structure-based screening, design, optimization of ligands (experimental & *in silico*)

RNA as a drug

sequence-dependent approaches

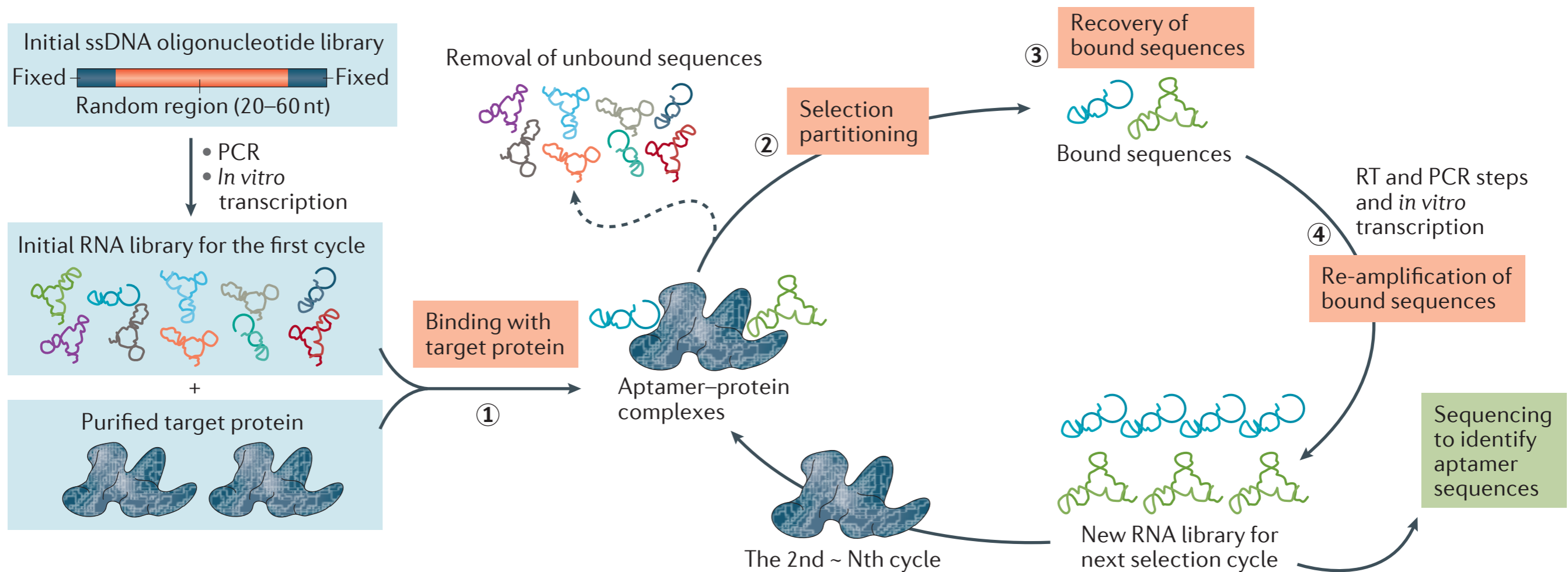
- hybridisation-dependent approaches: ASOs, siRNA, miR, CRISPR, splice modifying oligos, etc
- hybridisation-independent approaches: aptamers, immunostimulatory (CpG) oligos, mRNA, etc

RNAs as drugs: using intracellular mechanisms



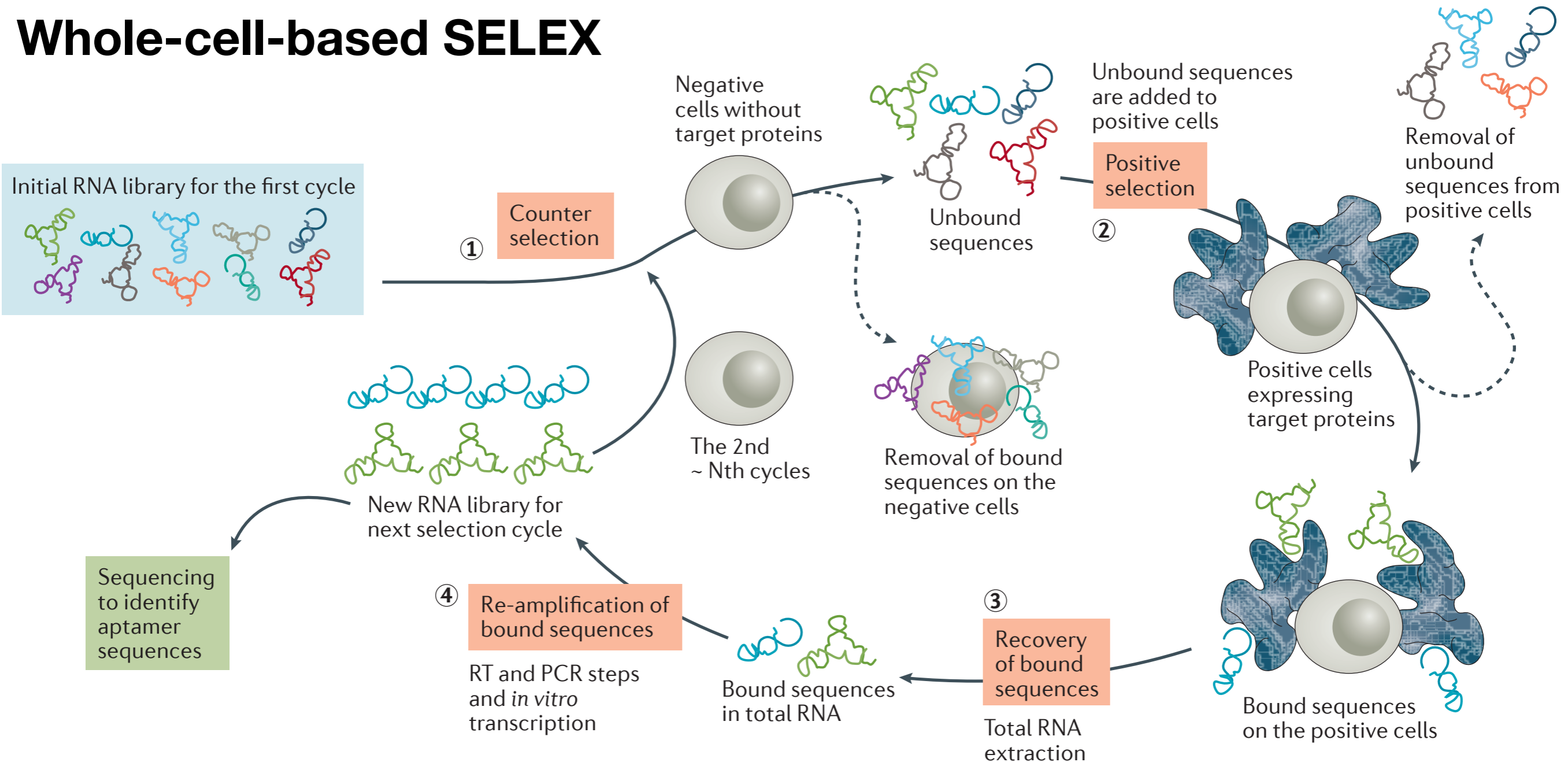
RNAs as drugs/tools: using in vitro generated aptamers (1)

Purified-protein-based SELEX



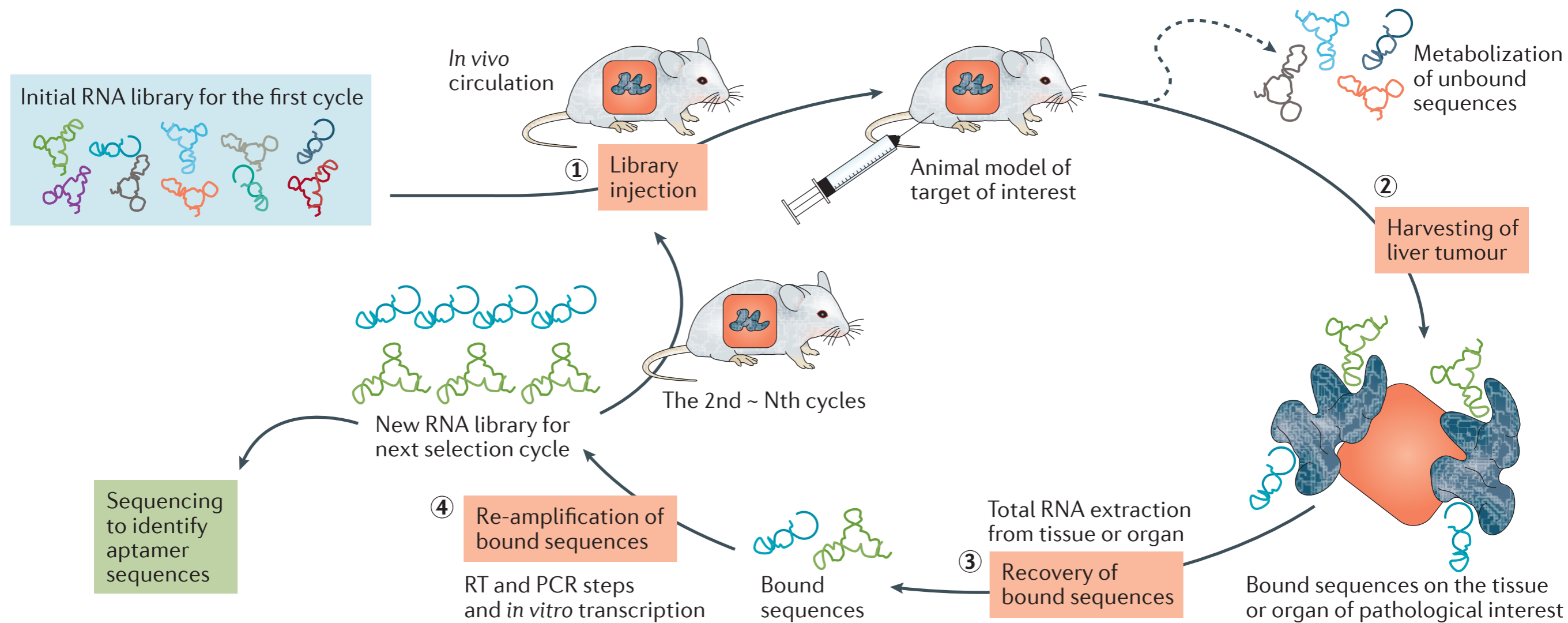
RNAs as drugs/tools: using in vitro generated aptamers (2)

Whole-cell-based SELEX



RNAs as drugs/tools: using in vitro generated aptamers (3)

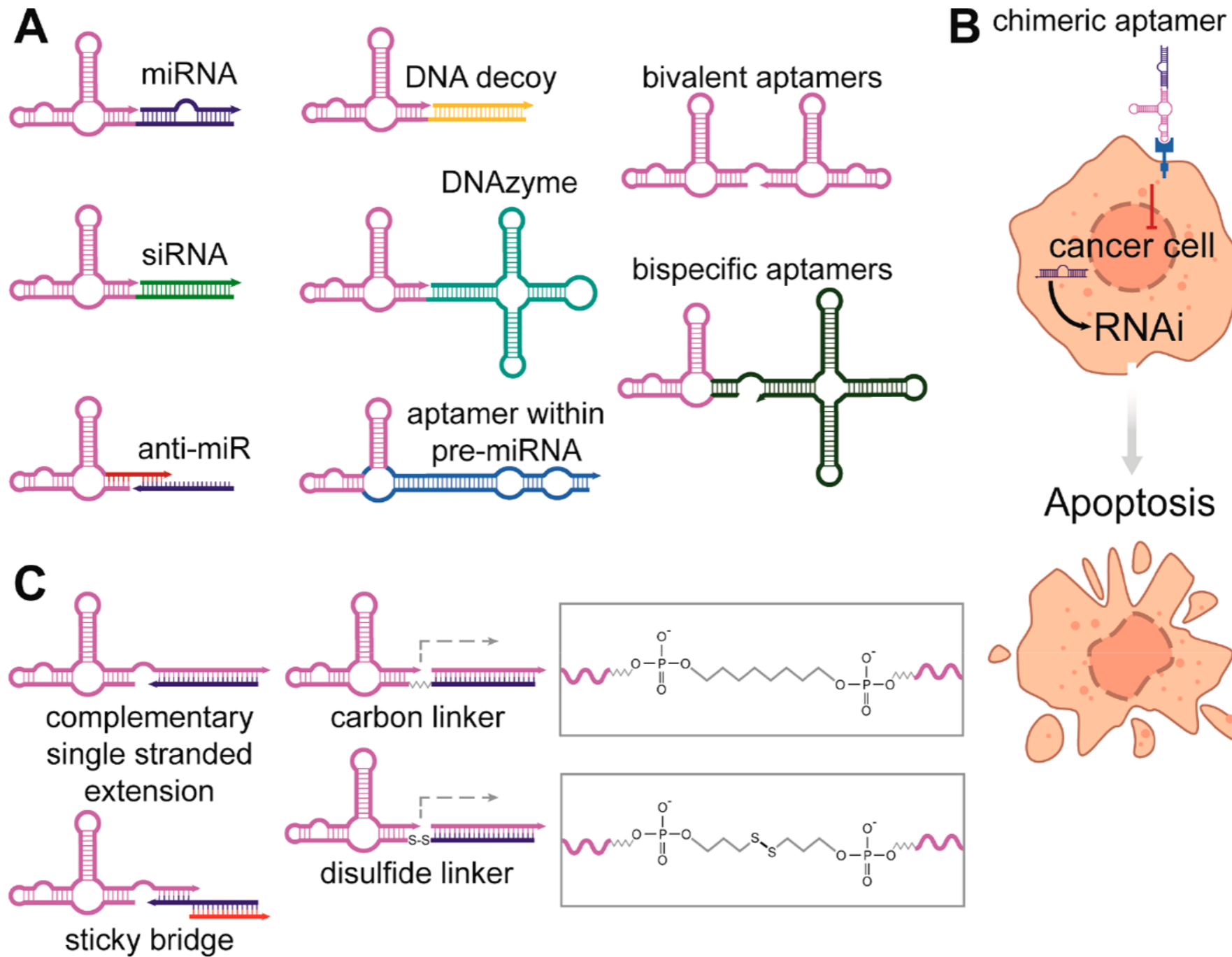
Live-animal-based SELEX



RNA as a drug & carrier

**sequence-dependent &
sequence-independent approaches**

Chimeric aptamers with miRNA, siRNA, etc



RNA as a drug

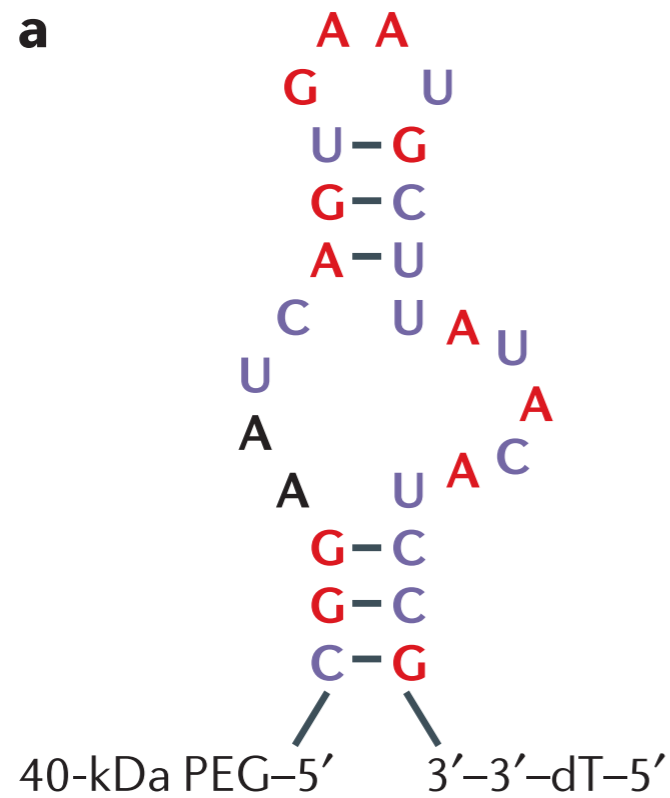
Approaches & Technologies

Aptamers in Therapeutics

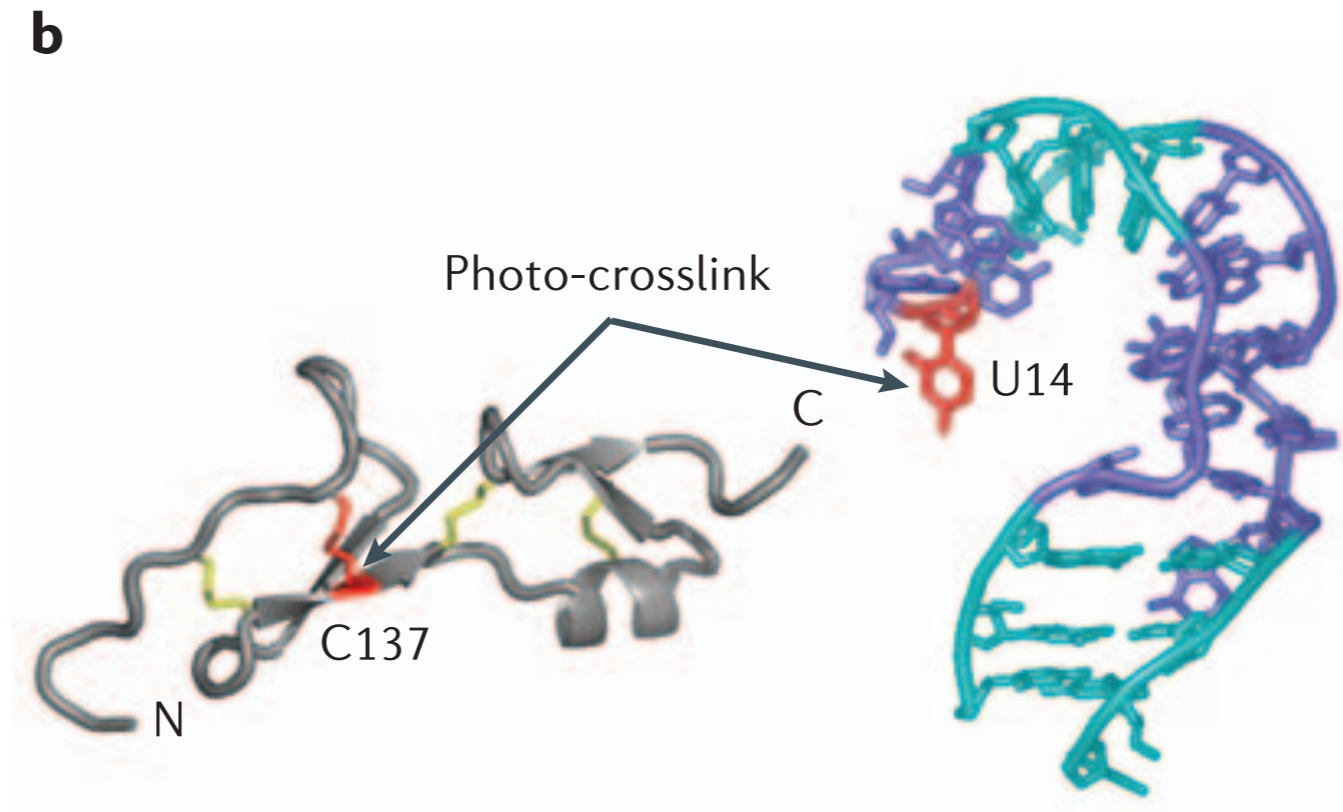
Name	Form	Target	Condition	Phase
Pegaptanib sodium (Macugen)	27-nt RNA	VEGF (Vascular Endothelial Growth Factor)	Age-related macular degeneration	Approved
E10030	29-nt DNA	PDGF (Platelet-Derived Growth Factor)	Age-related macular degeneration	Phase III
REG1 (RB006 and RB007)	37-nt RNA	Coagulation factor IXa	Coronary artery disease	Phase III
ARC1905	38-nt RNA	C5 (Complement component 5)	Age-related macular degeneration	Phase III
AS1411	26-nt DNA	Nucleolin	Acute myeloid leukemia	Phase II
ARC1779	39-nt DNA	A1 domain of von Willebrand factor	Von Willebrand disease/thrombotic thrombocytopenic/purpura	Phase II
NOX-E36	40-nt RNA	CCL2 (Chemokine C-C motif Ligand 2)	Chronic inflammatory diseases/type 2 diabetes mellitus/systemic lupus erythematosus	Phase II
NOX-A12	45-nt RNA	CXCL12 (Chemokine C-X-C motif Ligand 12)	Multiple myeloma and non-Hodgkin lymphoma/autologous or hematopoietic stem cell transplantation	Phase II
NU172	26-nt DNA	Thrombin	Heart disease	Phase II
NOX-H94	44-nt RNA	Hepcidin peptide hormone	Anemia/end-stage renal disease/inflammation	Phase II
ARC19499	32-nt RNA	TFPI (Tissue Factor Pathway Inhibitor)	Hemophilia	Phase I

FDA 2004

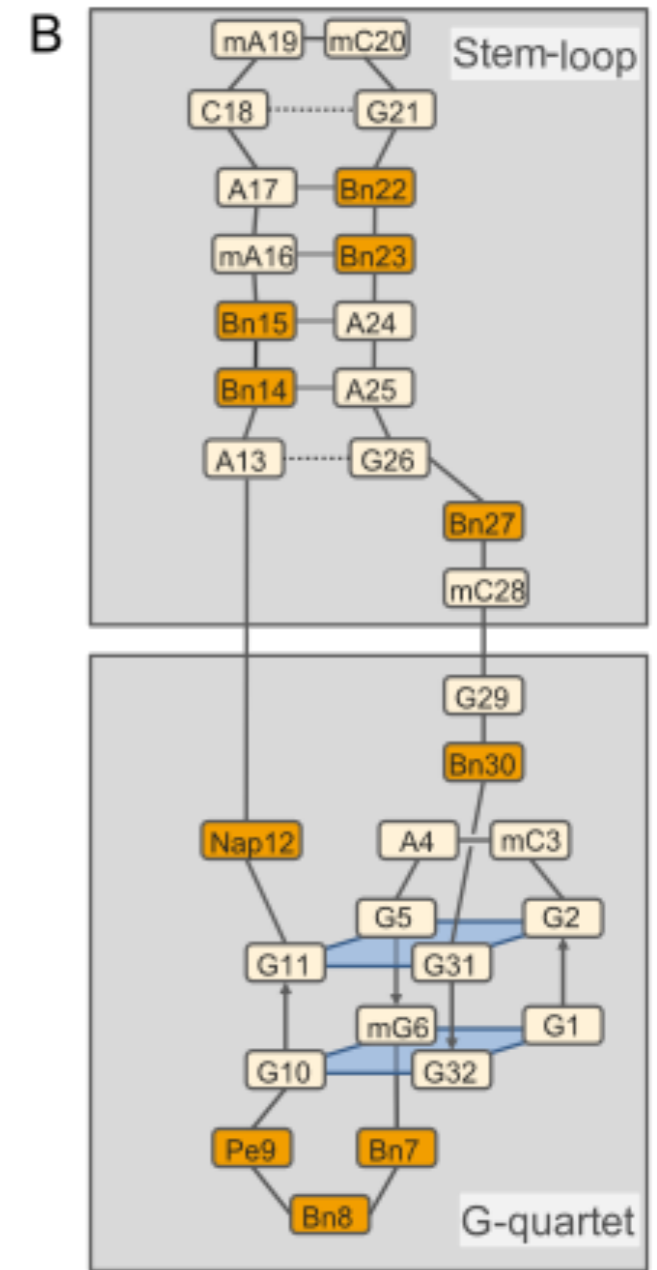
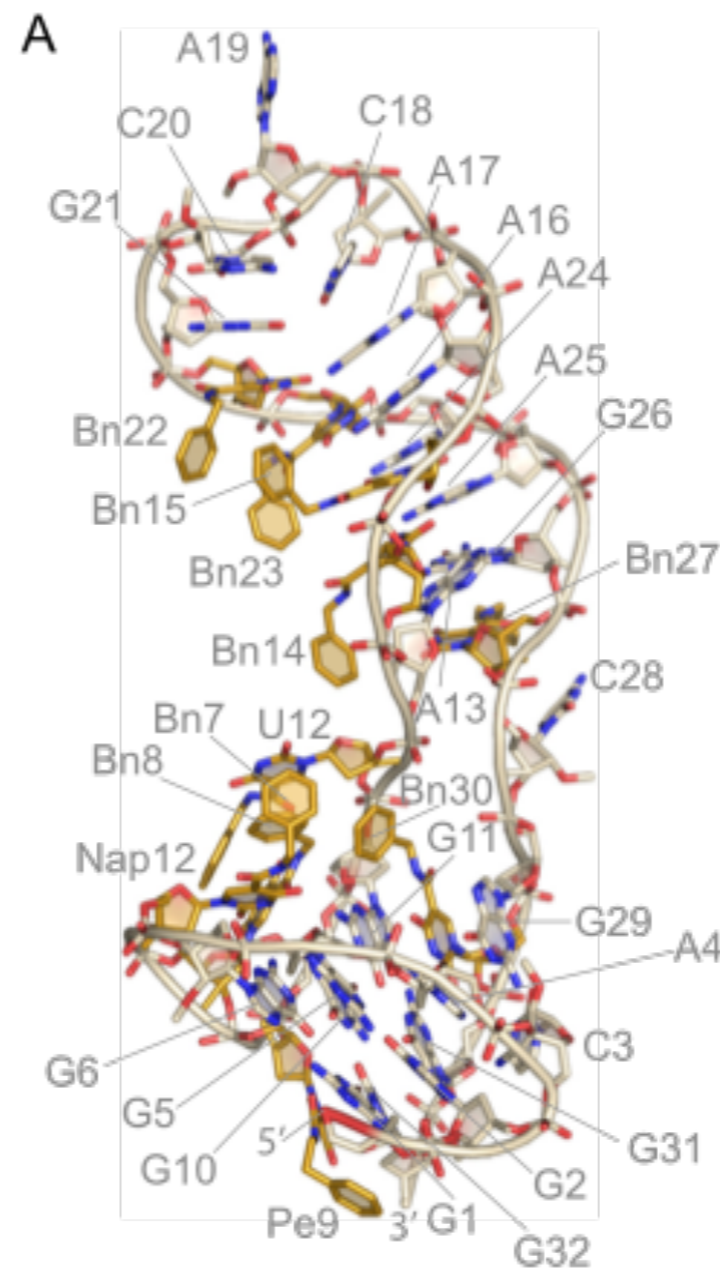
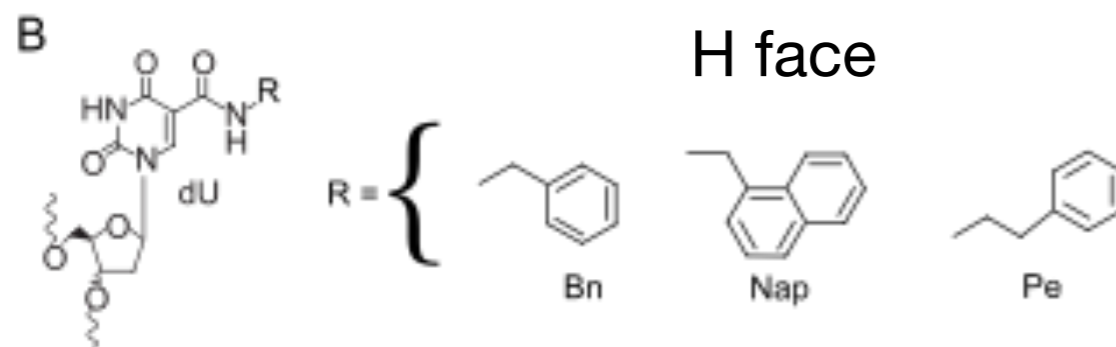
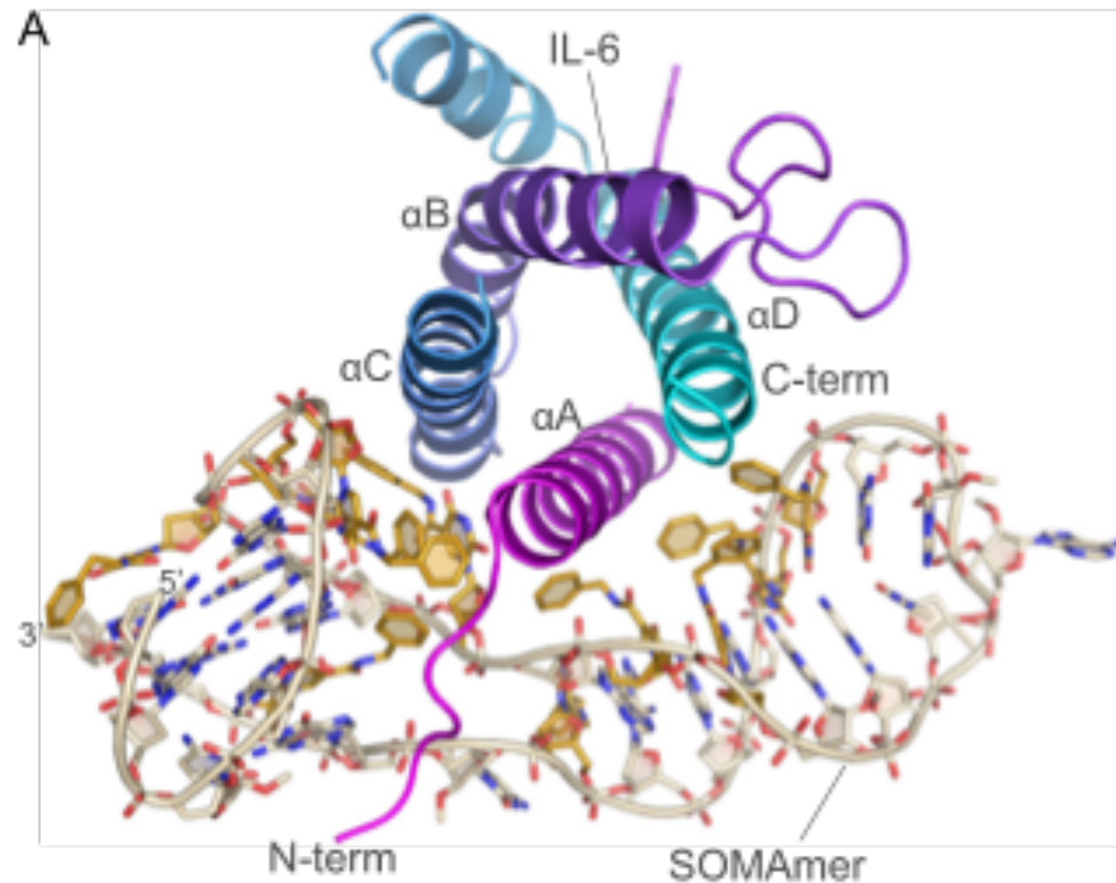
Pegaptanib (macugen): anti-VEGF aptamer



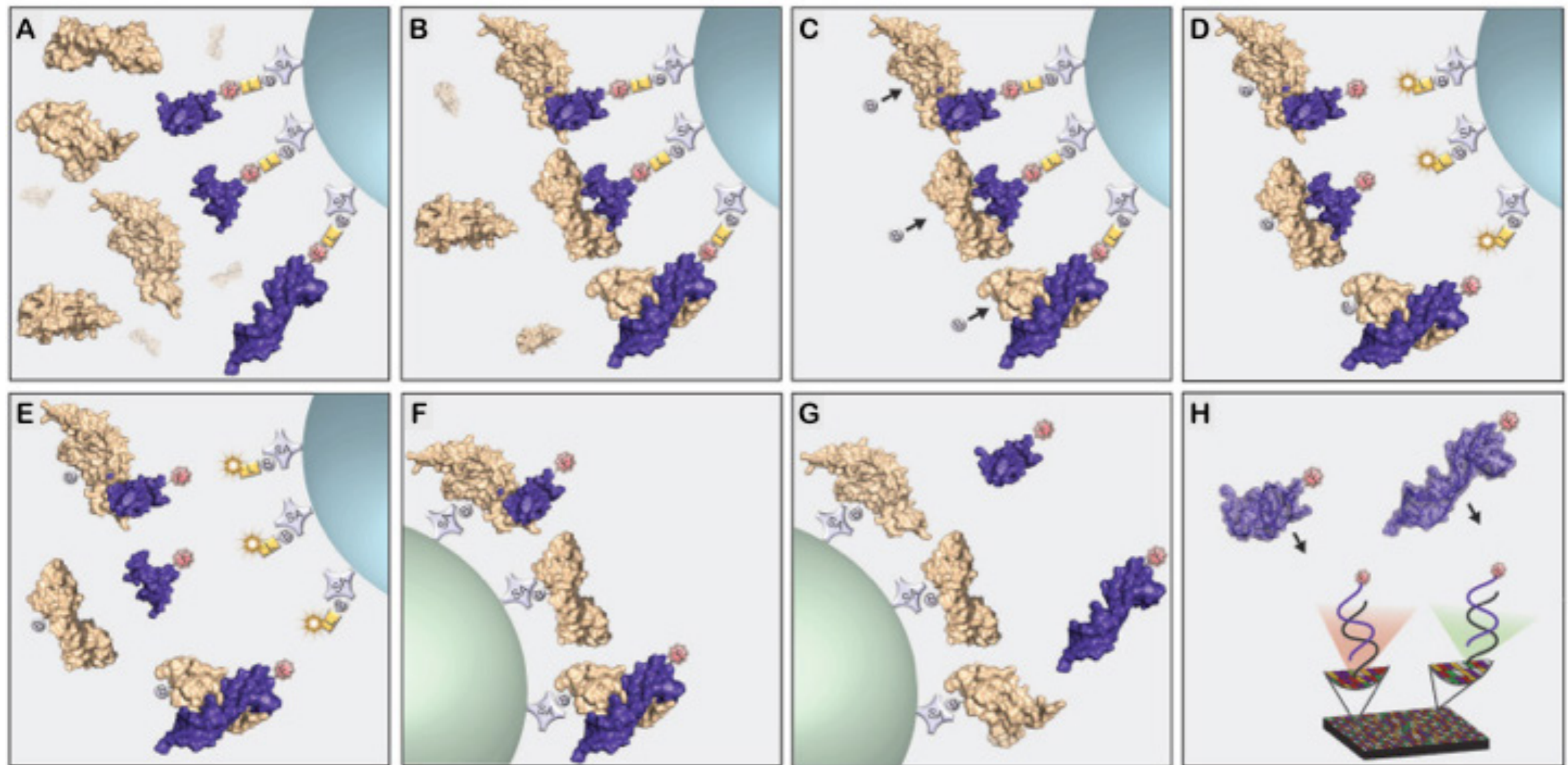
N: 2'O-Me



SOMAmers: Slow Off-rate Modified Aptamers

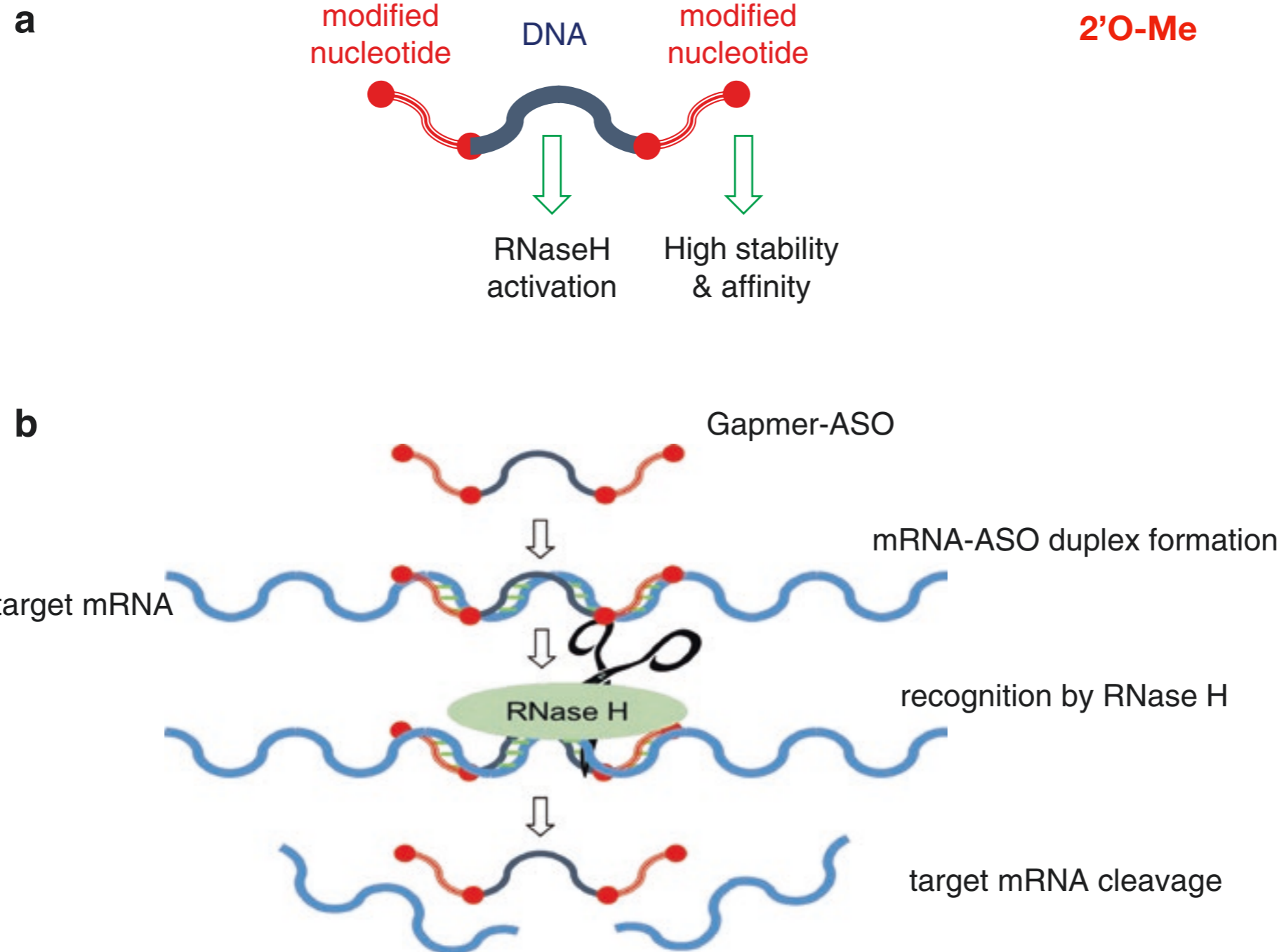


SOMAscan: SOMAmer applied technology in proteomic

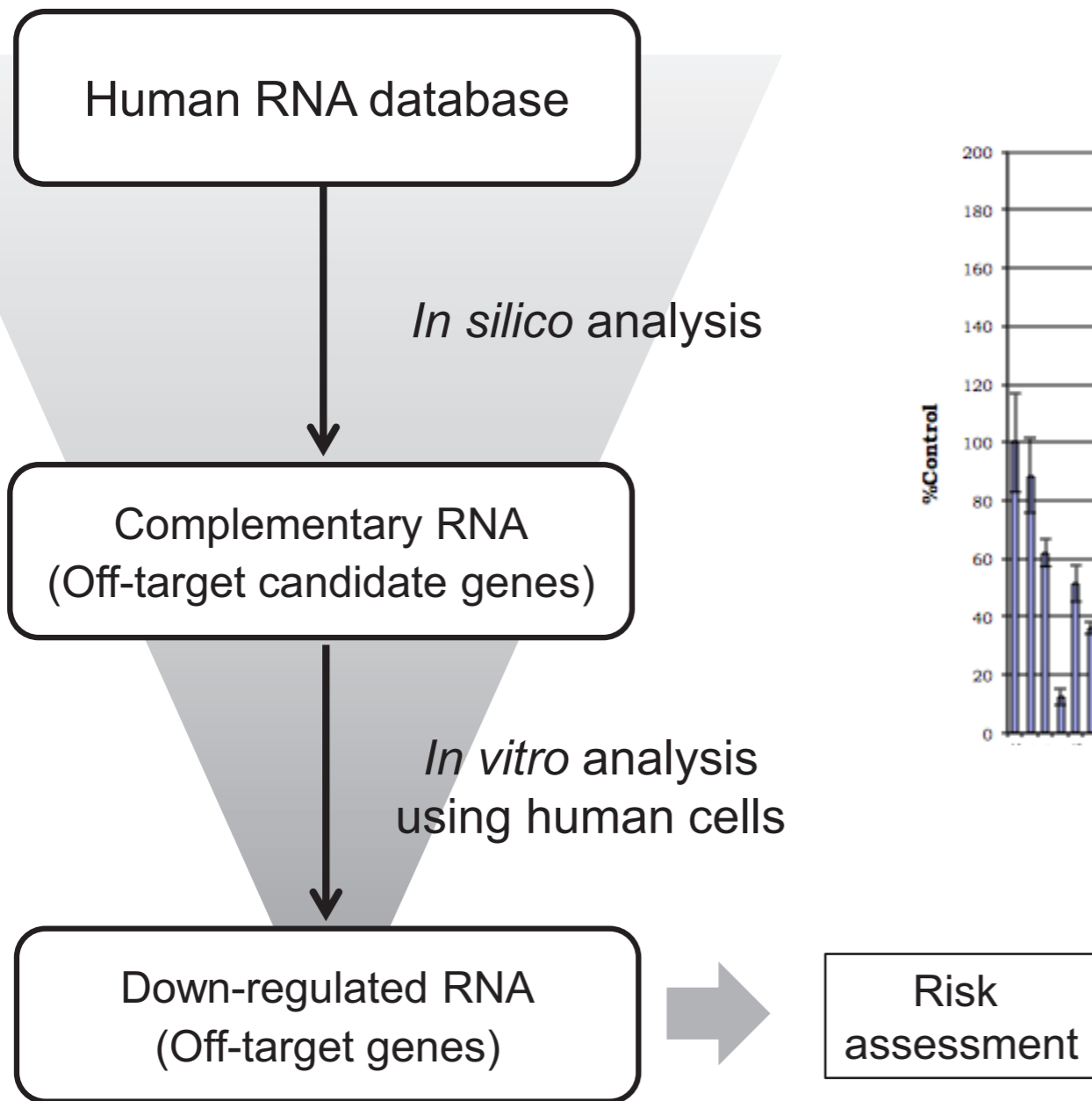


detection of 1300 human proteins

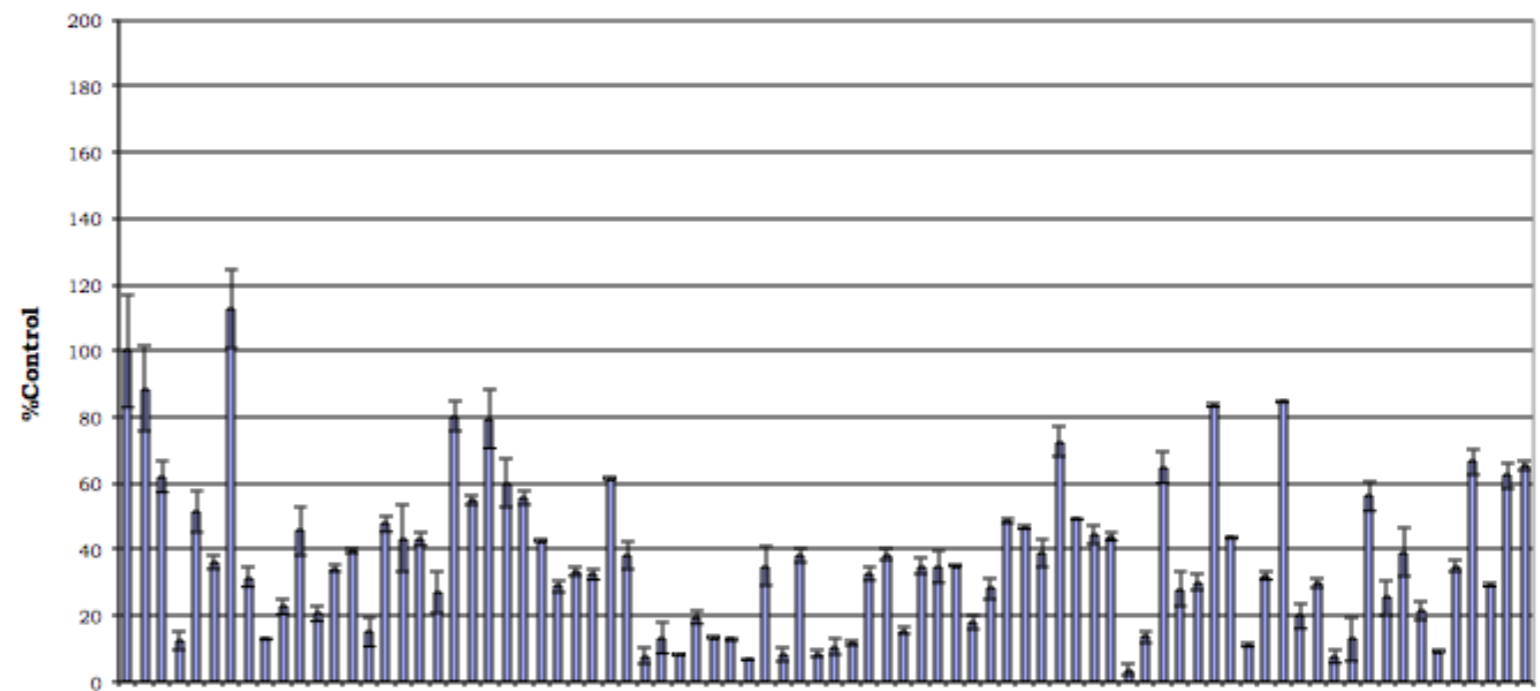
ASOs and gapmers



Design of ASOs-gapmers using *in silico* methods



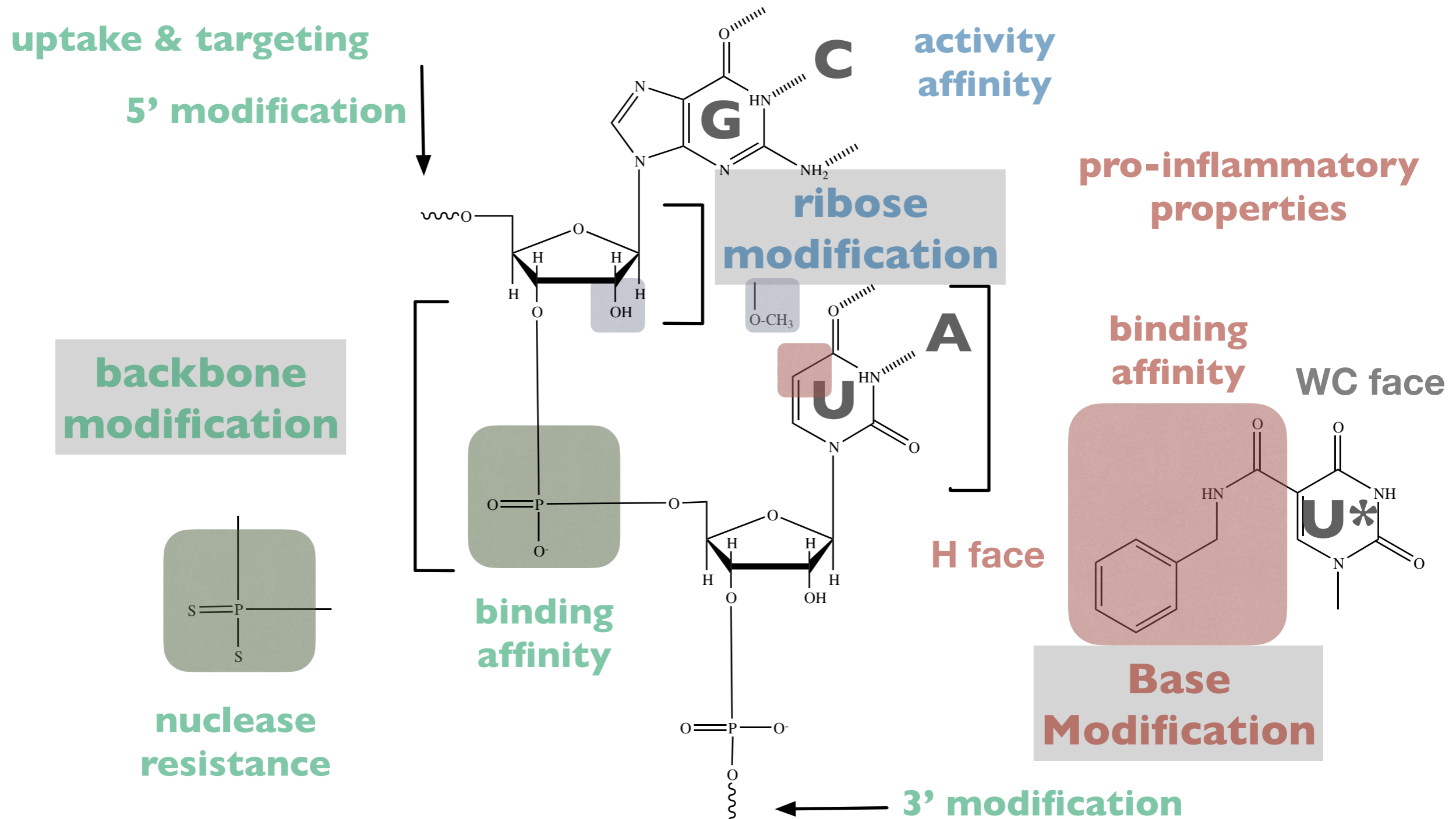
Activity for perfect match of ASOs



Limitations of “RNA as a drug”-based approaches

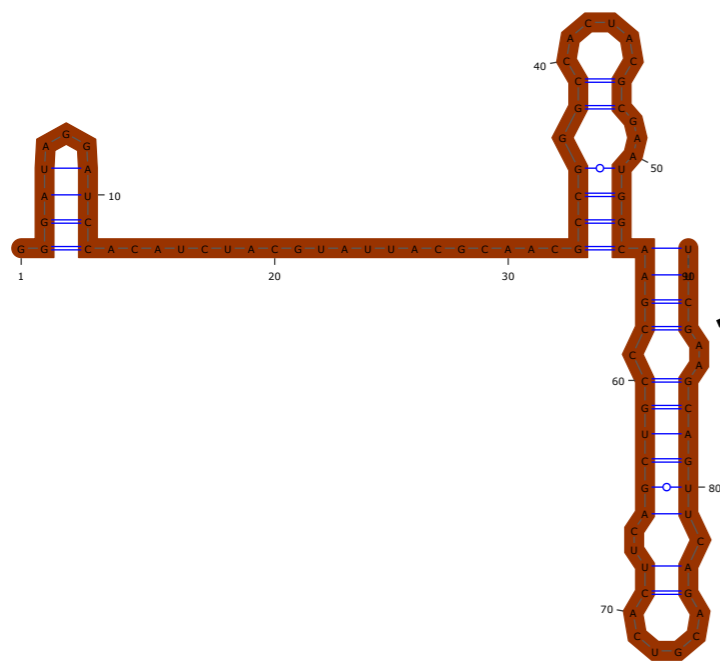
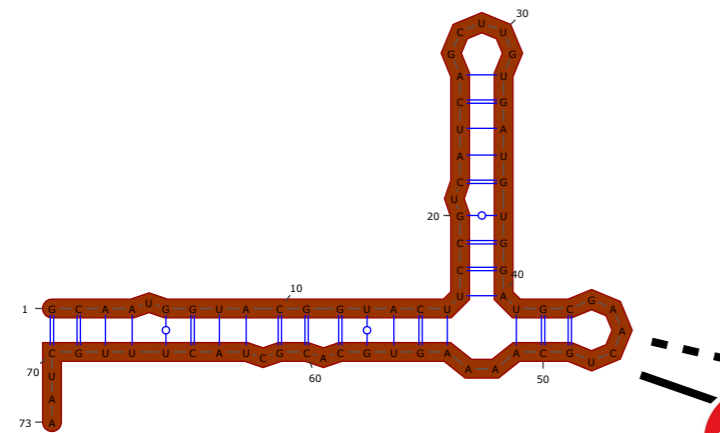
- general limitations:
 - safety, toxicity (immunogenicity), uptake, delivery, etc
- specific to RNAs: sequence & hybridisation dependent
 - activity/sequence match (structure, accessibility, etc)
 - uptake, toxicity
- specific to RNAs: sequence & hybridisation independent
 - inflammation (CpG motifs and TLR9 activation)
 - modification-related toxicities

Optimizations via chemical modifications

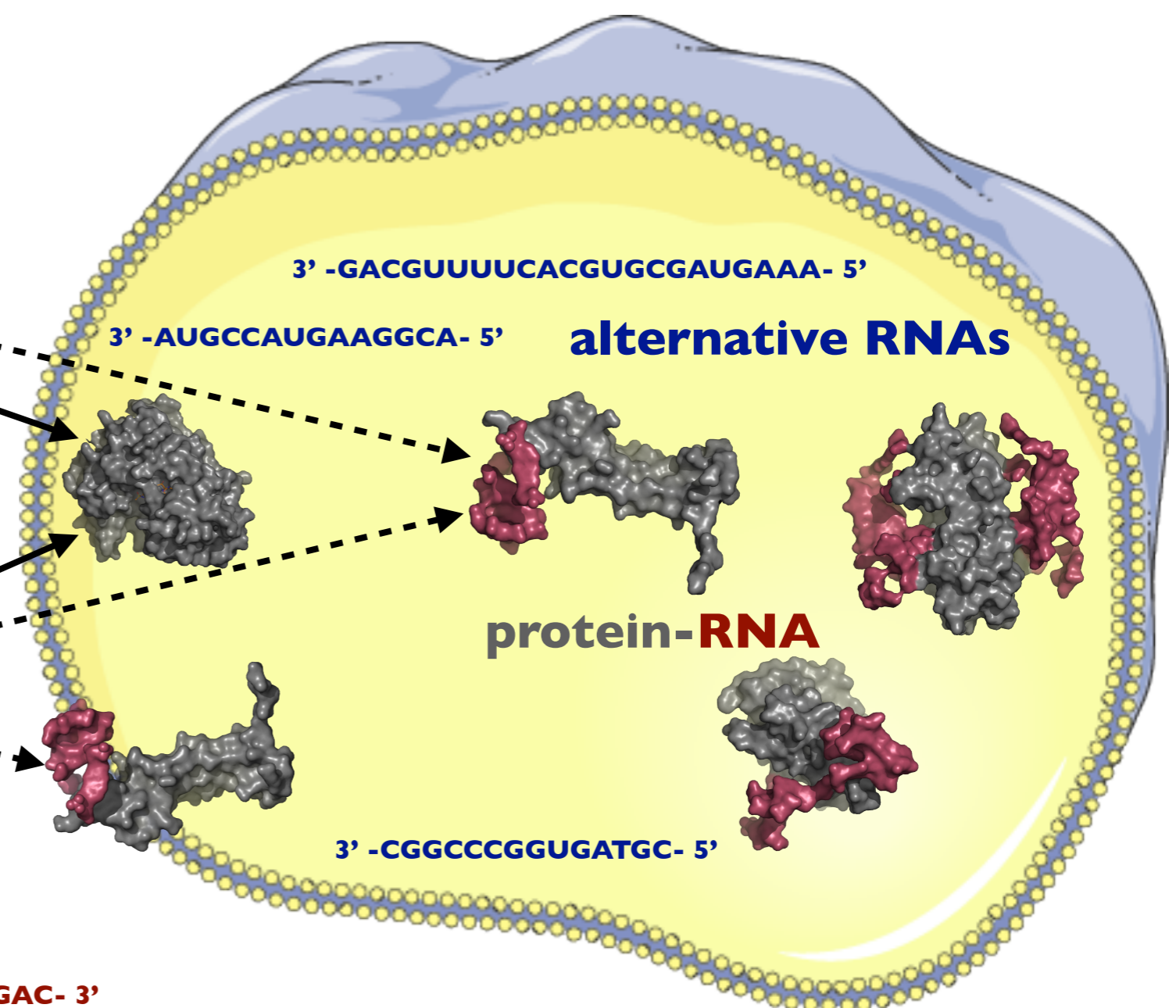
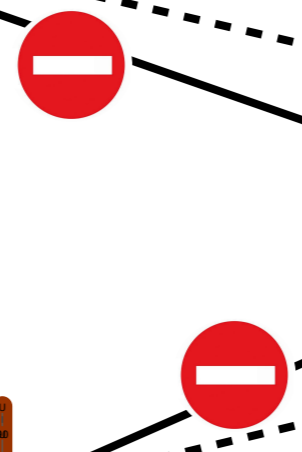


Hybridisation-dependent & independent OTEs

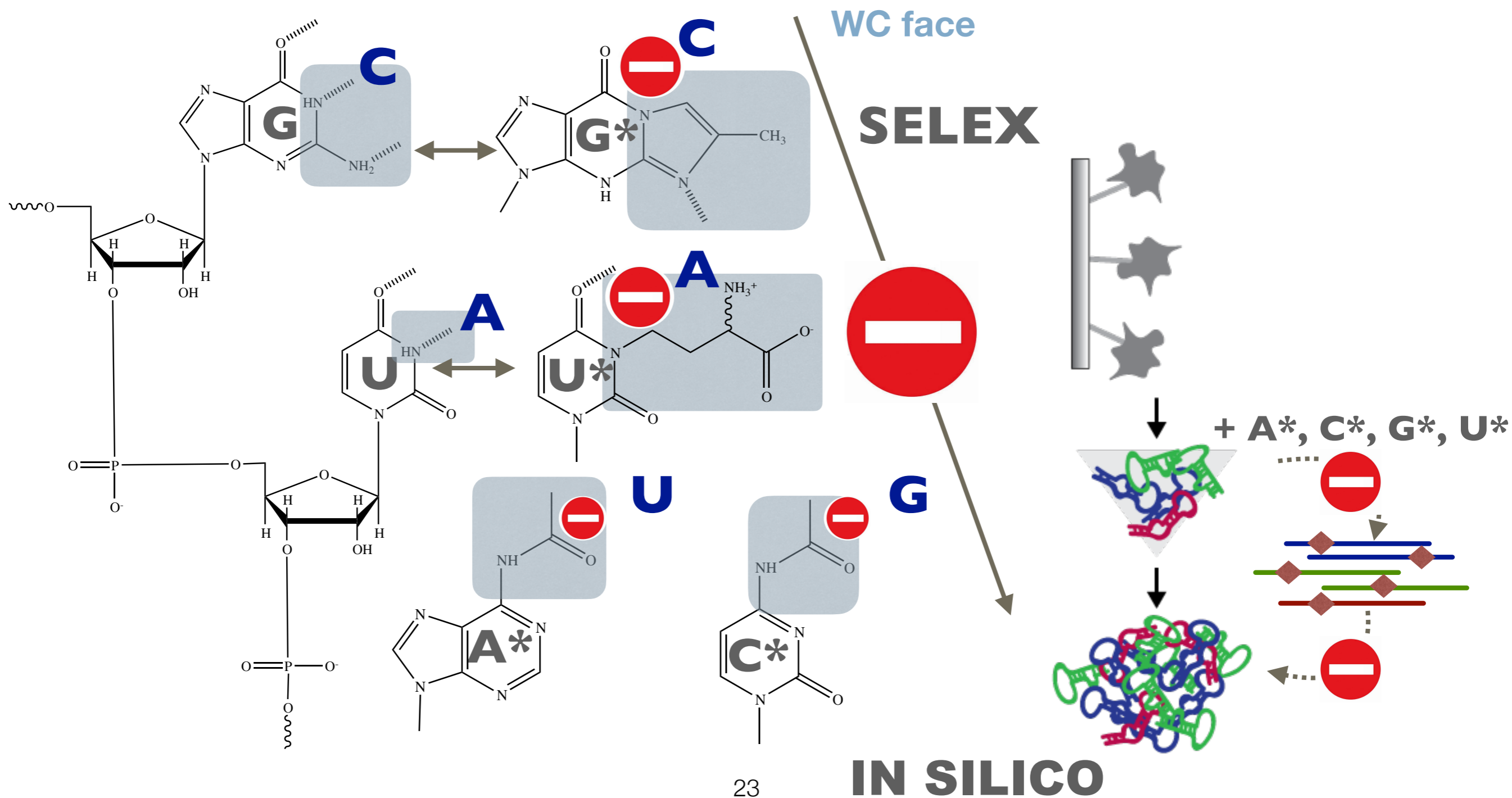
5' -GCAAUGGUACGGUACUUCGCUCAUCAGCUUGUGAUGUGGAUGCGAACUGCAAAGUGCACGCUACUUUGC UAA- 3'



5' -CGCAACGCCGGGCCACTACGCGAATGGCAAGCCCGTCGAC- 3'



How to avoid hybridisation-dependent OTEs ?

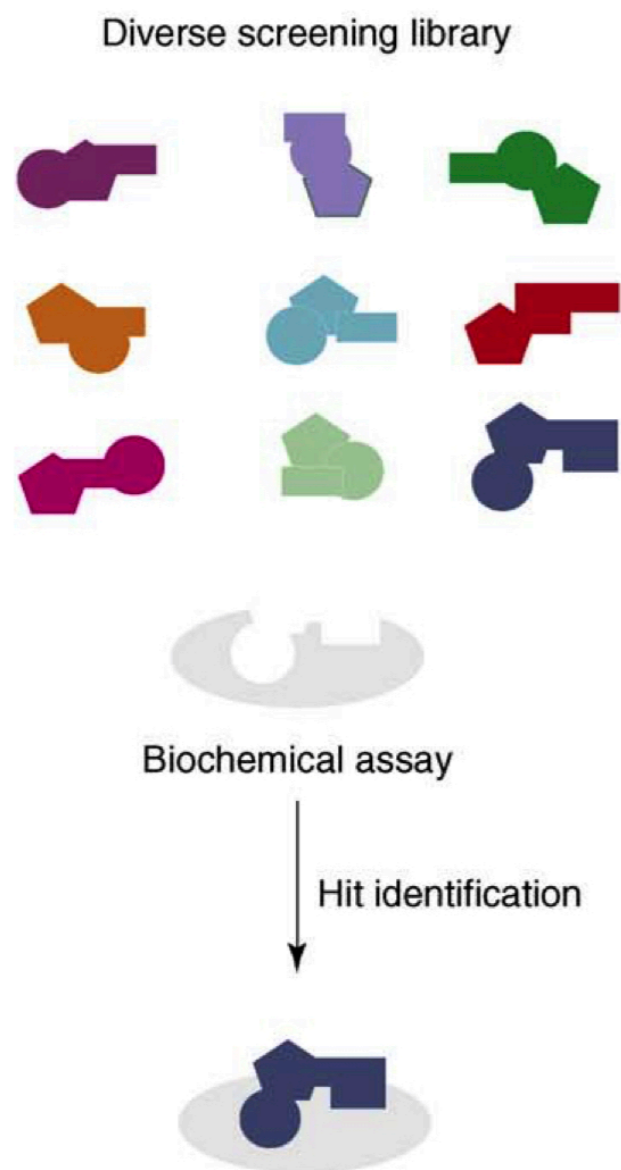


In silico design of modified oligonucleotides

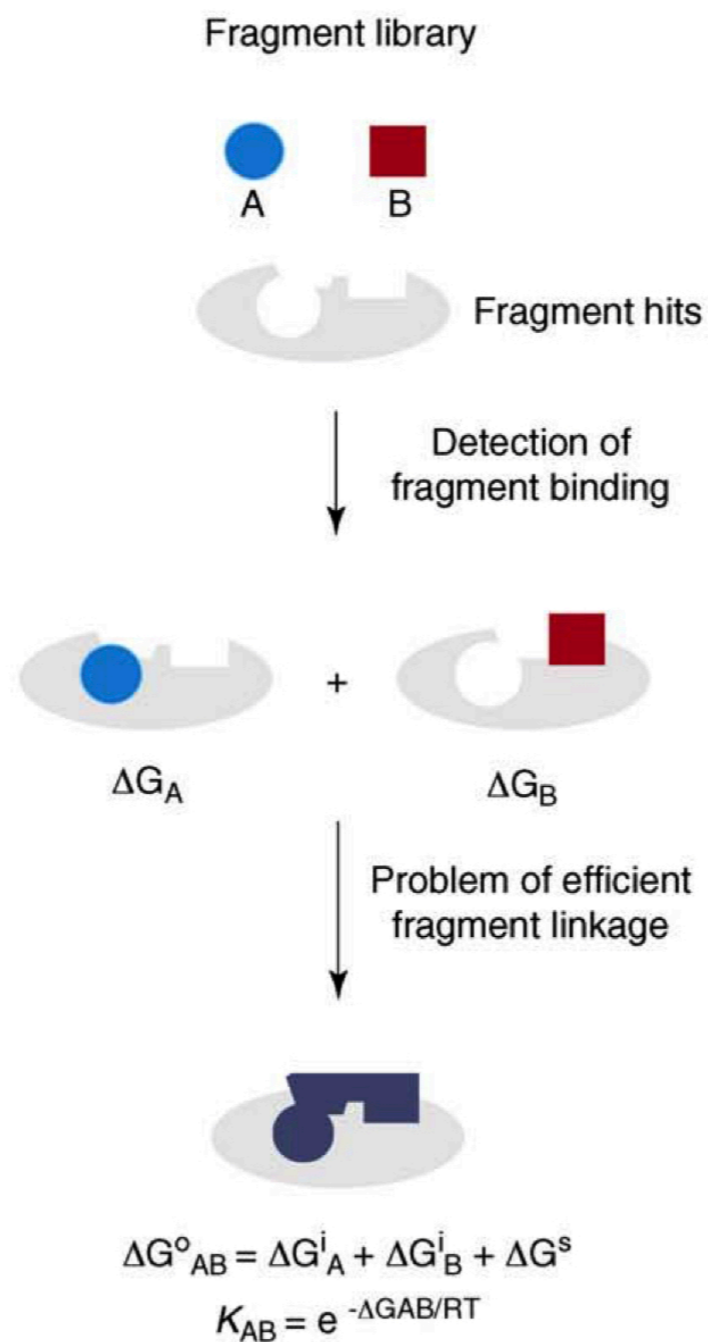
- selection of a protein target with a **known 3D structure** and a binding region
- **fragment-based approach** to identify single nucleotide binding sites (fragments: mapping, clustering, ranking)
- **(re)construction of short oligonucleotides** (selection of fragments, search for connectivity, building and optimisation, ranking)

Drug design/discovery

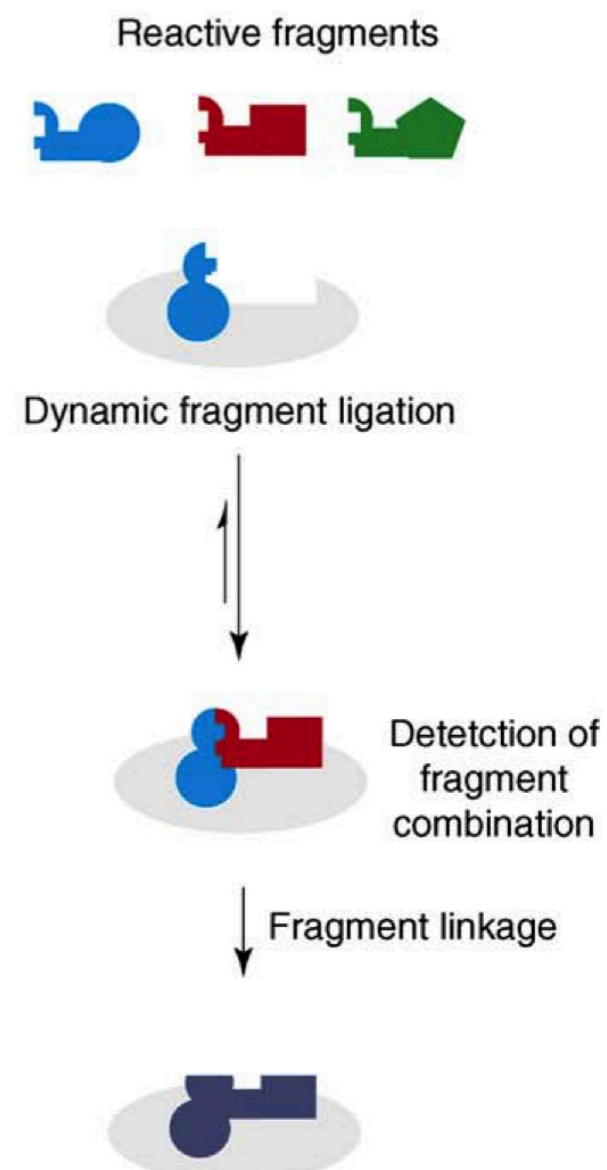
(a) High-throughput screening



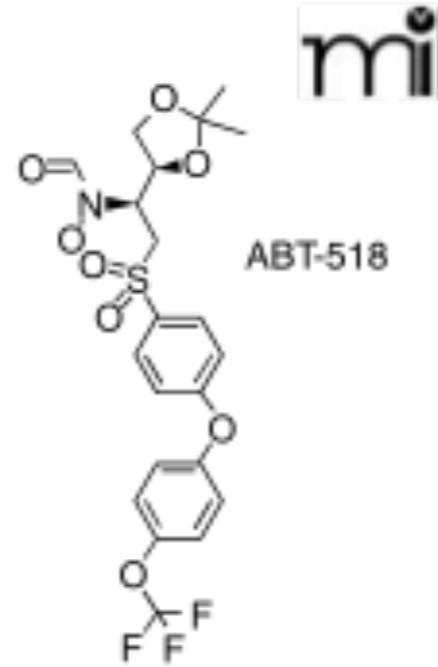
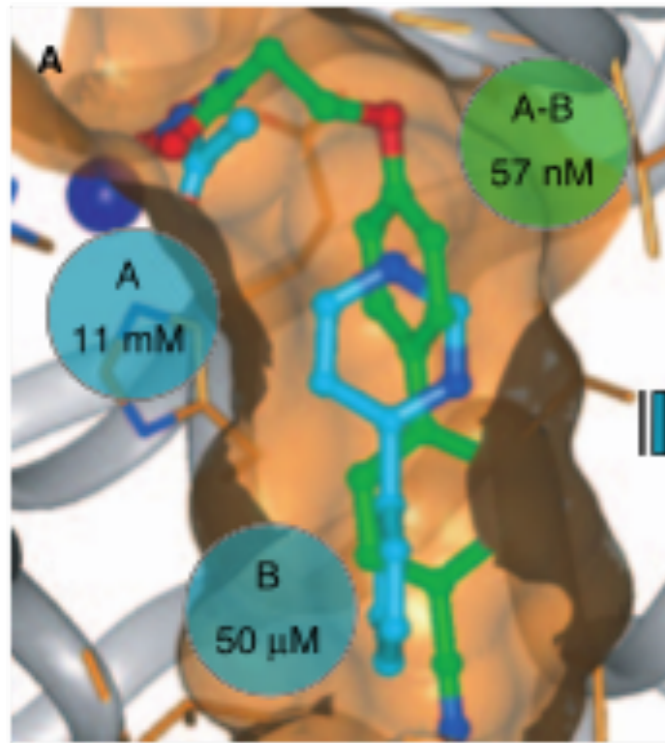
(b) Fragment-based drug discovery



(c) Dynamic, template-assisted drug discovery



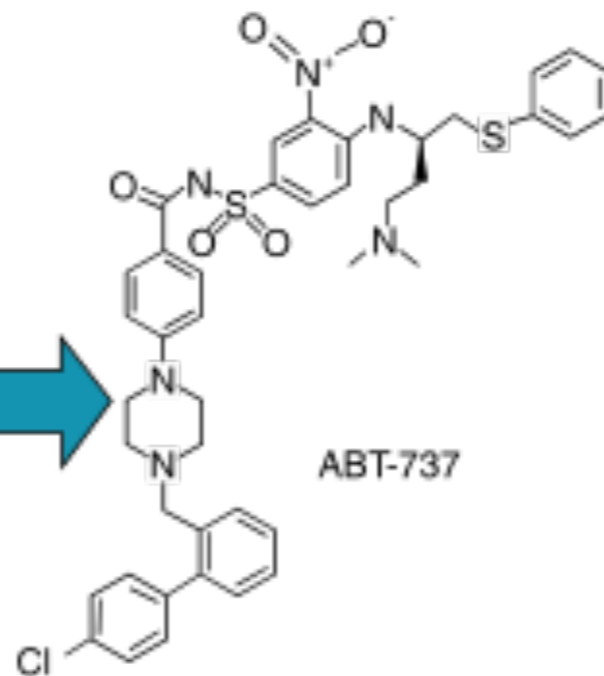
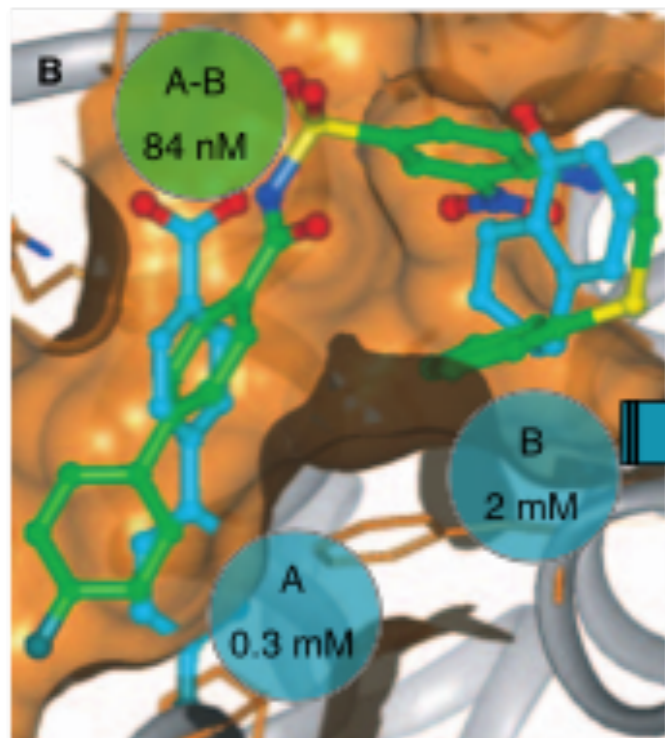
Success stories in FBD



“SAR by NMR”

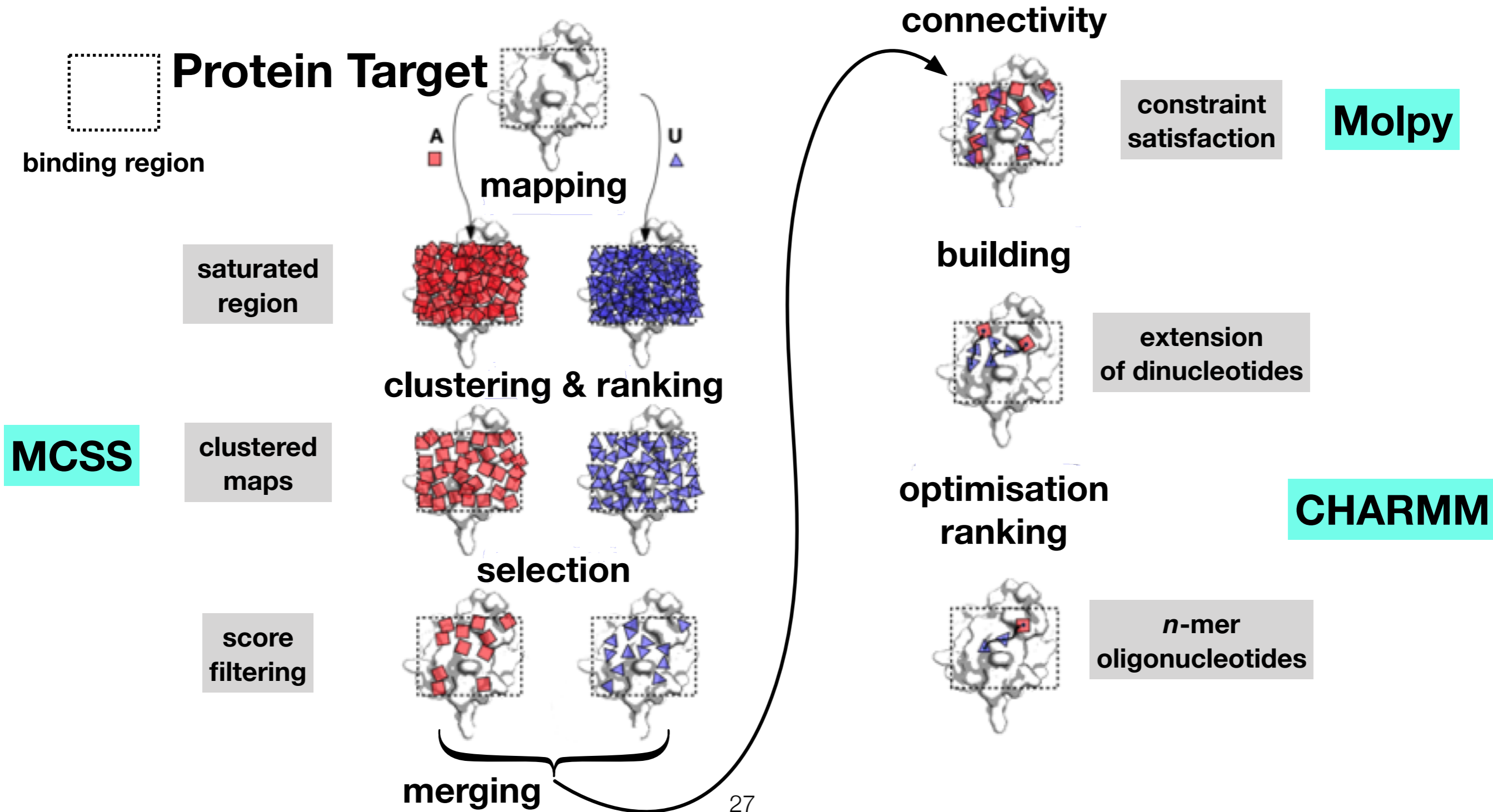
Shuker *et al.*, 1996

Metalloproteinases (MMP-3)
(arthritis and tumor metastasis)



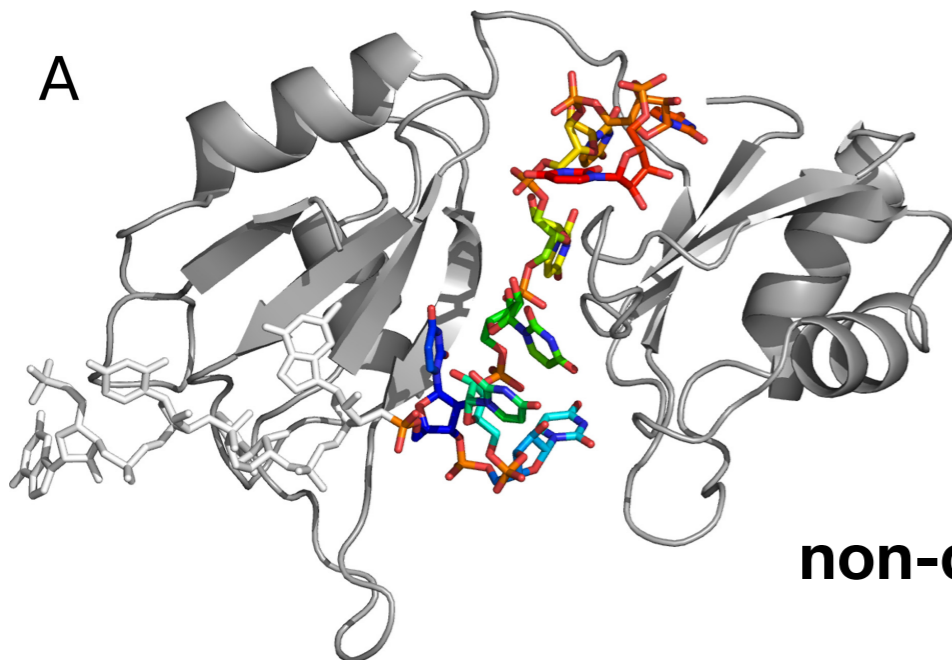
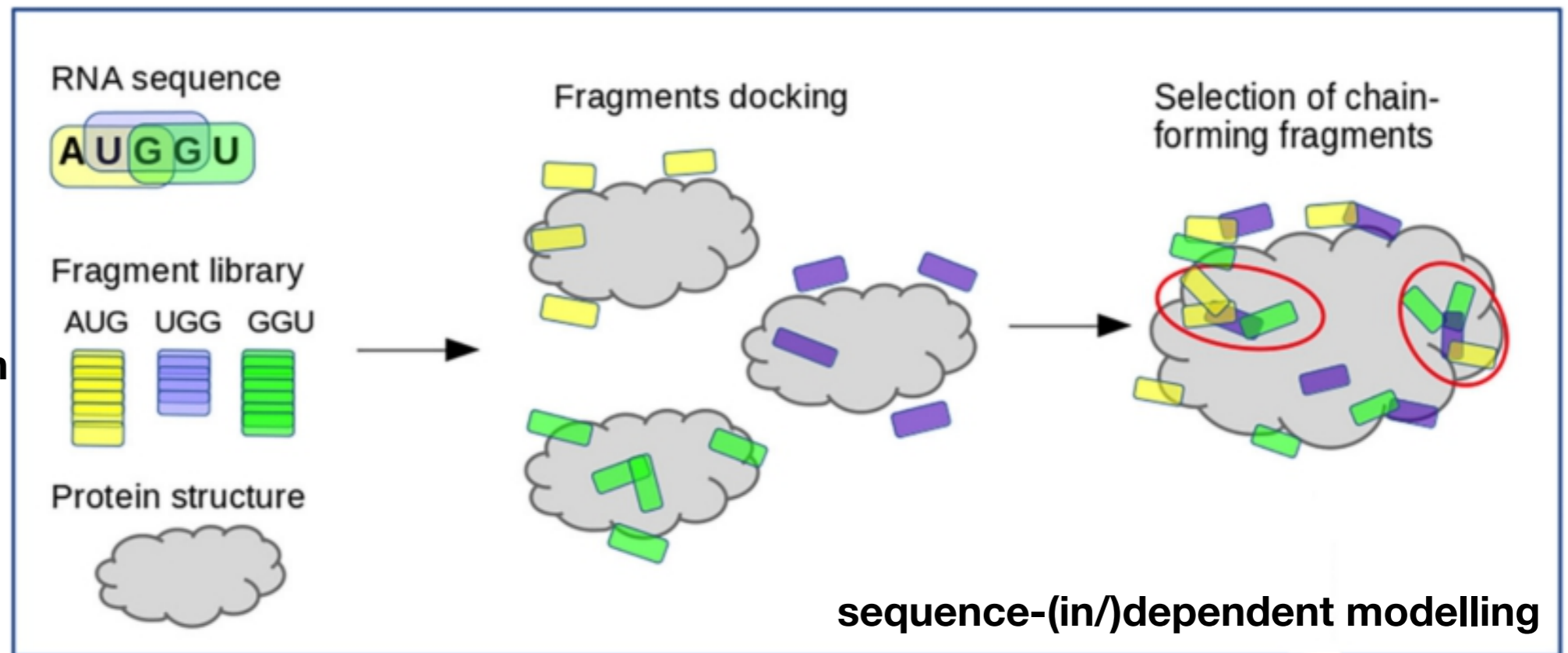
Bcl-2
(pro- and anti-apoptotic)

Fragment-Based Design of Oligonucleotides

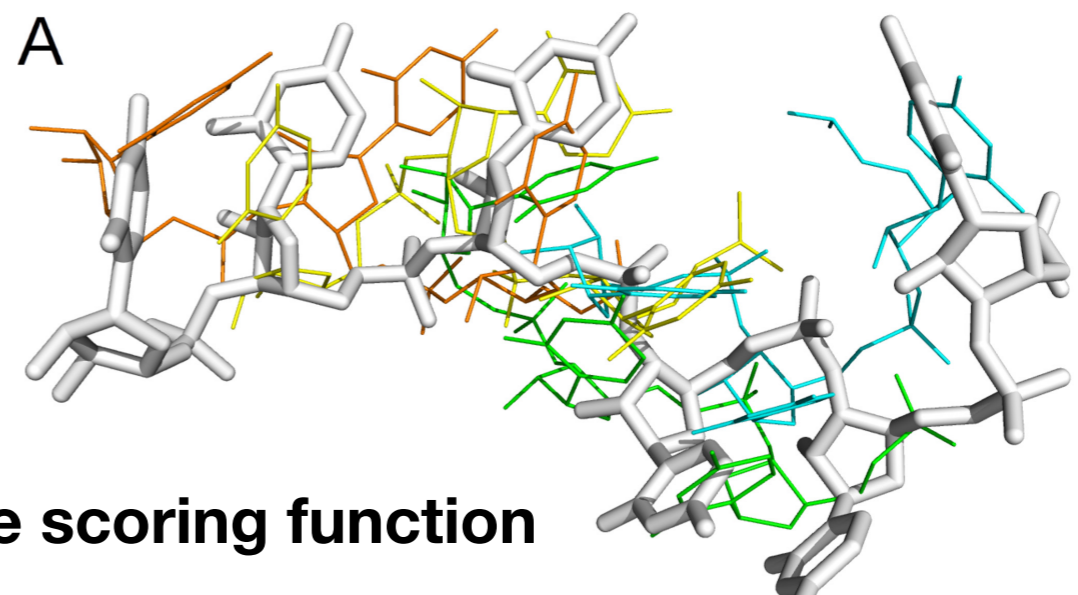


Fragment-based approach applied to RNAs: proof of concept

discrete representation of 3-mers



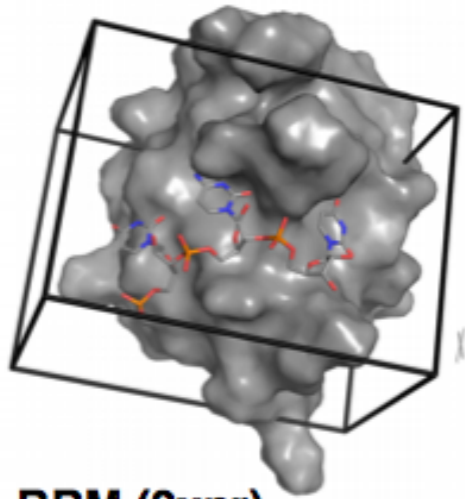
non-discriminative scoring function



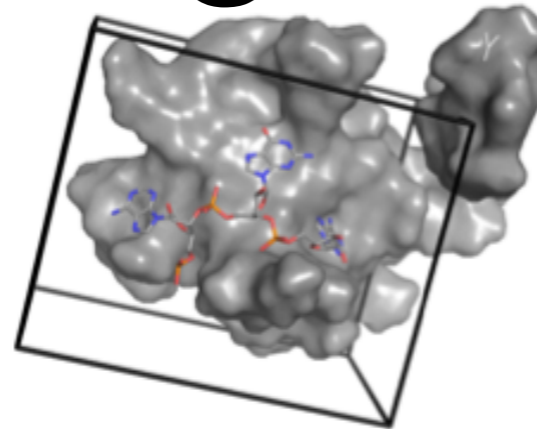
Preliminary results on three RBPs testcases

- **General objective:** validation by reproducing known ssRNA-protein interactions
- **Specific objectives:**
 - sequence constraints modeling: predict binding mode
 - sequence constraints free modeling: predict binding mode and sequence specificity
- **Perspectives in molecular design**

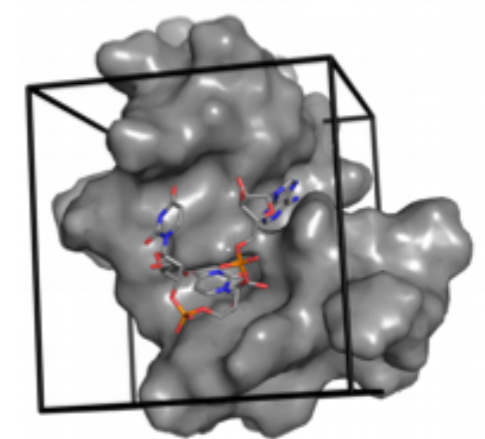
RBPs: optimal/sub-optimal binding sites



RRM (2xnr)
5'-UCU-3'



KH (5wwx)
5'-AGA-3'

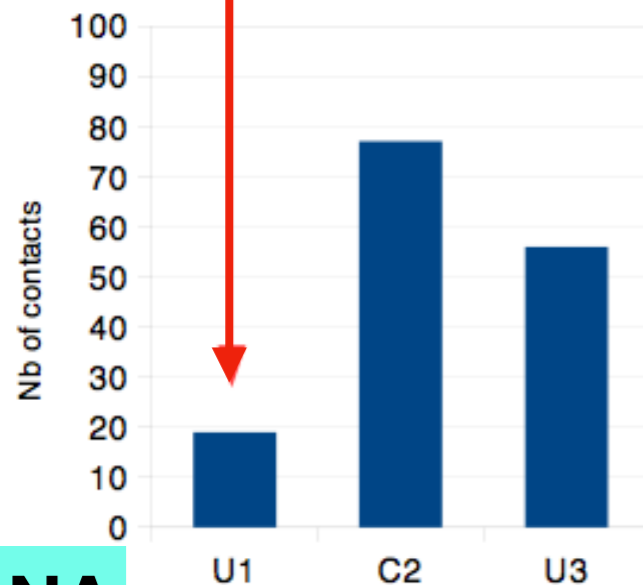


Zn-CCCH (5elh)
5'-UUA-3'

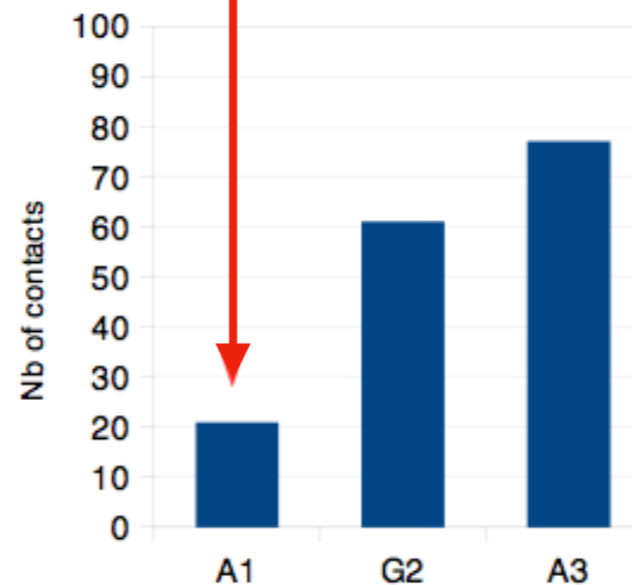
native poses identified (RMSD \leq 2.0Å)

MCSS
score index

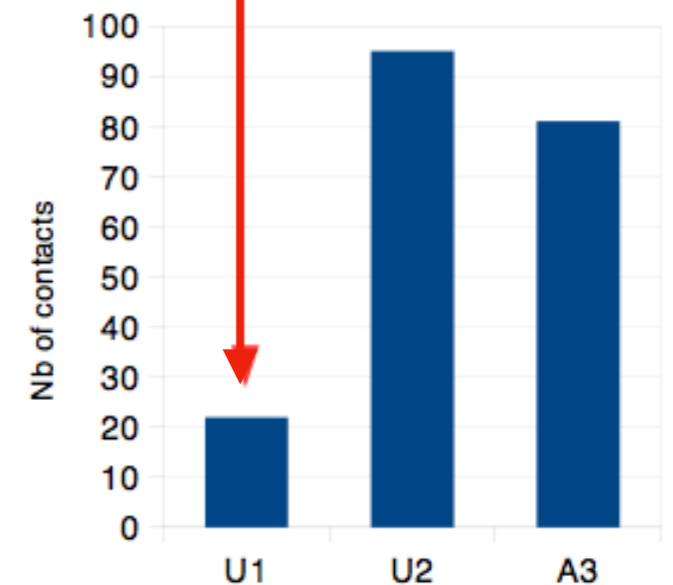
U1	C2	U3
155	1	9



A1	G2	A3
691	7	58



U1	U2	A3
490	3	1



Preliminary results on RBPs

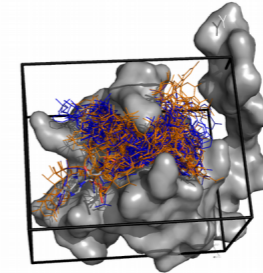
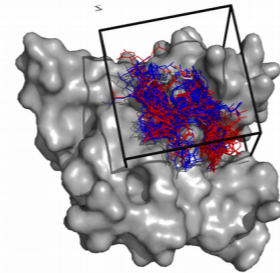
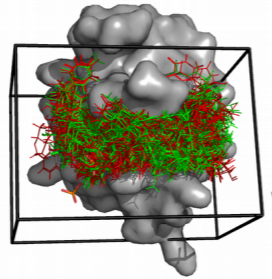
2 & 3-mers oligonucleotides

200 initial poses

200U & 200C
2XNR

200U & 200A
5ELH

200A & 200G
5WWX



Chain search
for CU

Chain search
for UA

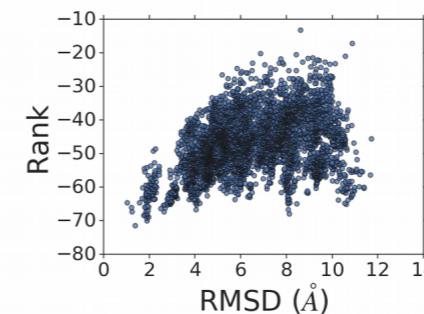
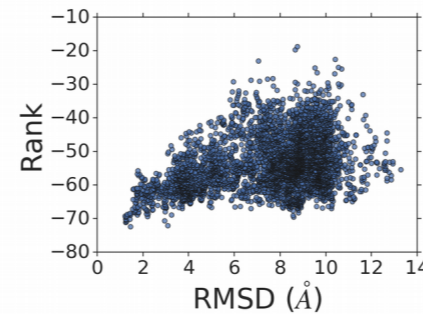
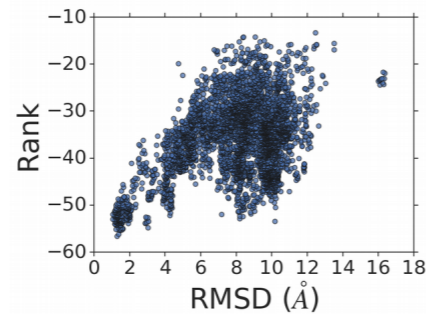
Chain search
for AG

4670 chains

5550 chains

5151 chains

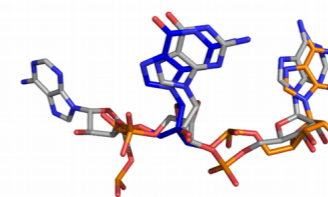
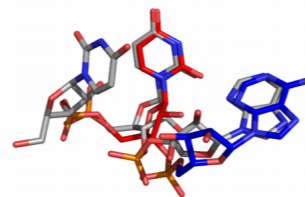
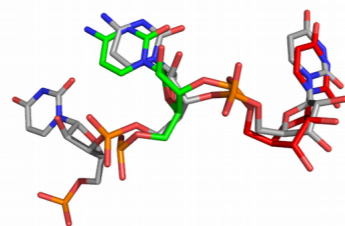
~ 5000 2-mers



Top 1 chain

Top 1 chain

Top 1 chain



Top1 for 2-mers

RMSD = 1,3 Å

RMSD = 1,36 Å
31

RMSD = 1,45 Å

MCSS

+

Molpy

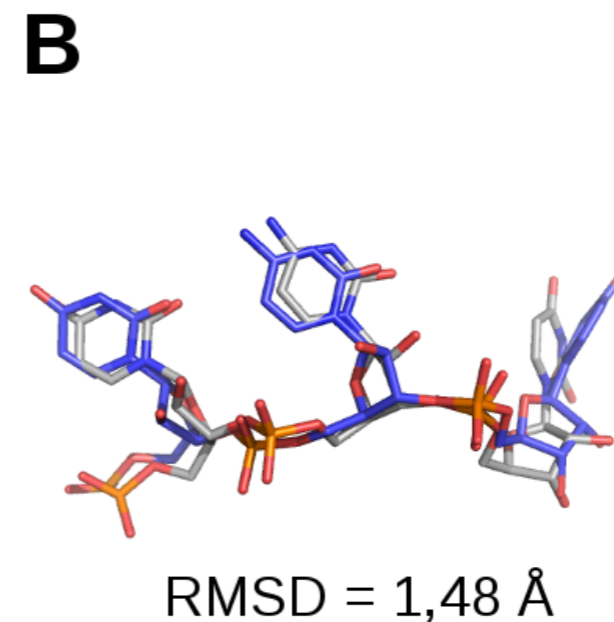
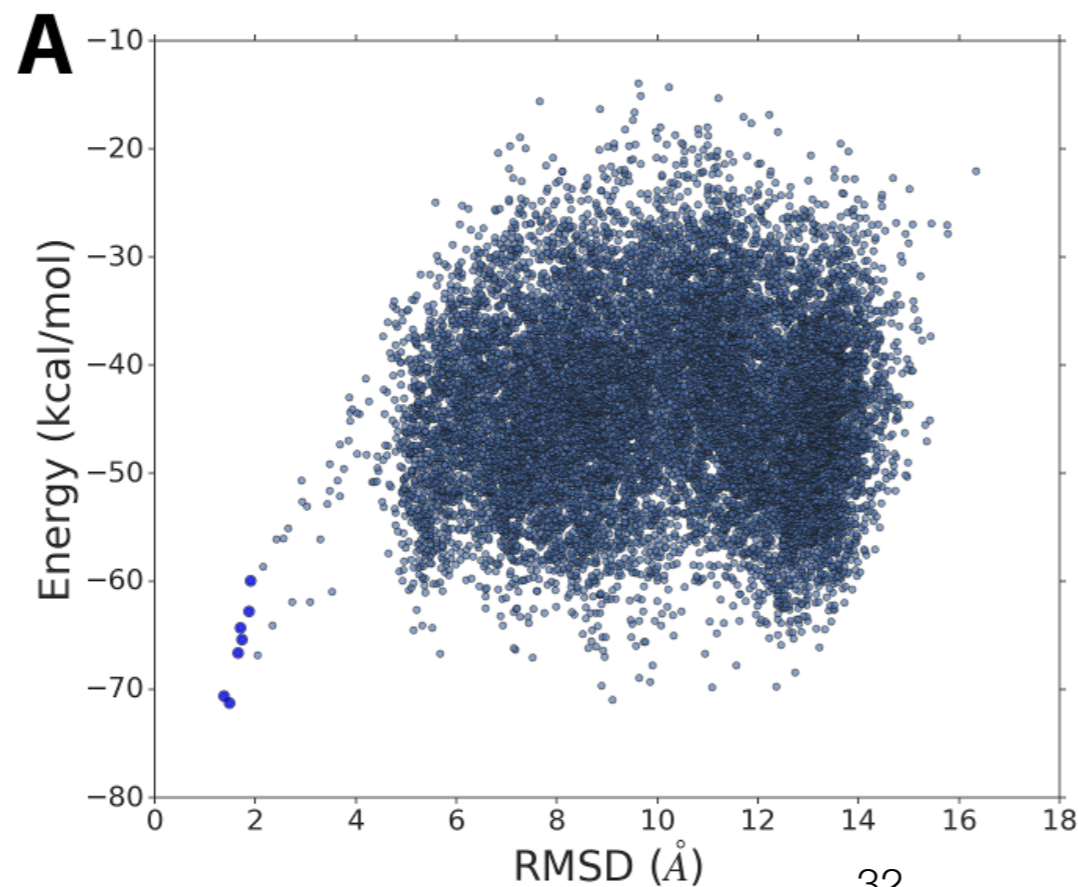
+

CHARMM

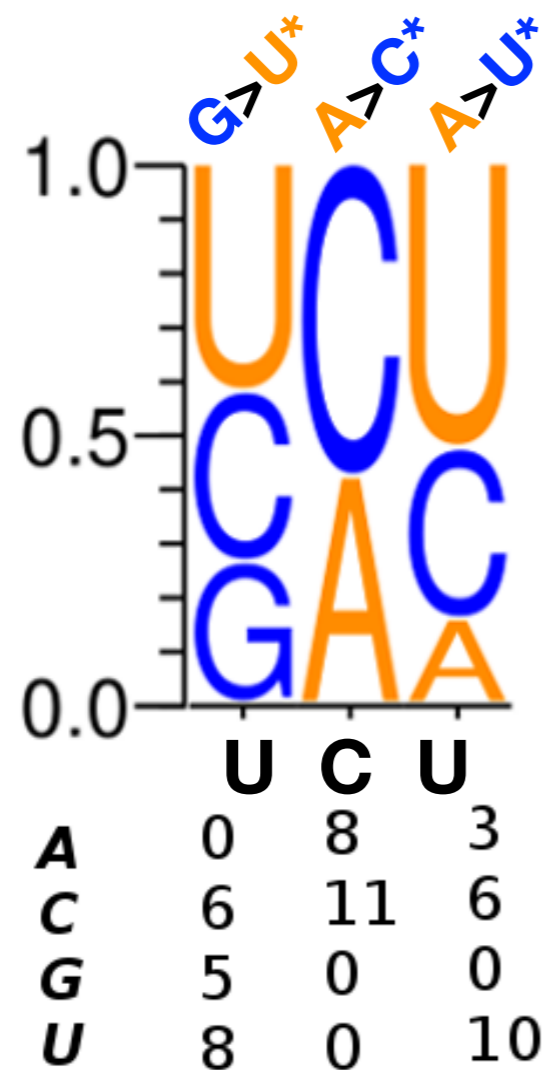
Preliminary results on RBPs

3-mers oligonucleotides

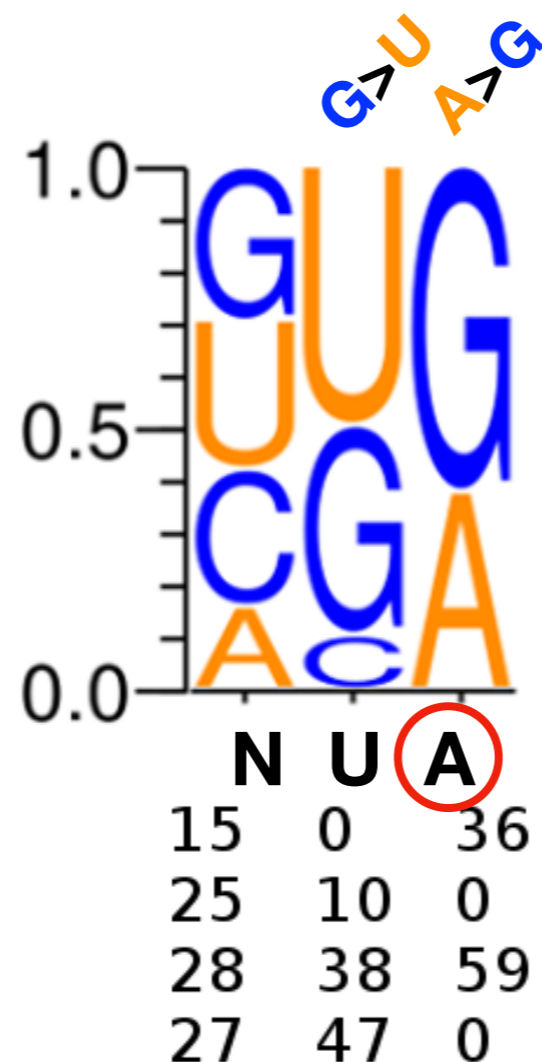
Code PDB	Nombre de poses sélectionnées				Séquence recherchée	Nombre de chaînes identifiées	Chaîne native ($\leq 2 \text{ \AA}$) de meilleure énergie d'interaction	
	A	C	G	U			Rang global	RMSD (\AA)
2XNR	/	155	/	155	UCU	13 782	1	1,48
5ELH	490	/	/	490	UUA	1 584 619	ND	ND
5WWX	691	/	691	/	AGA	2 433 270	ND	ND



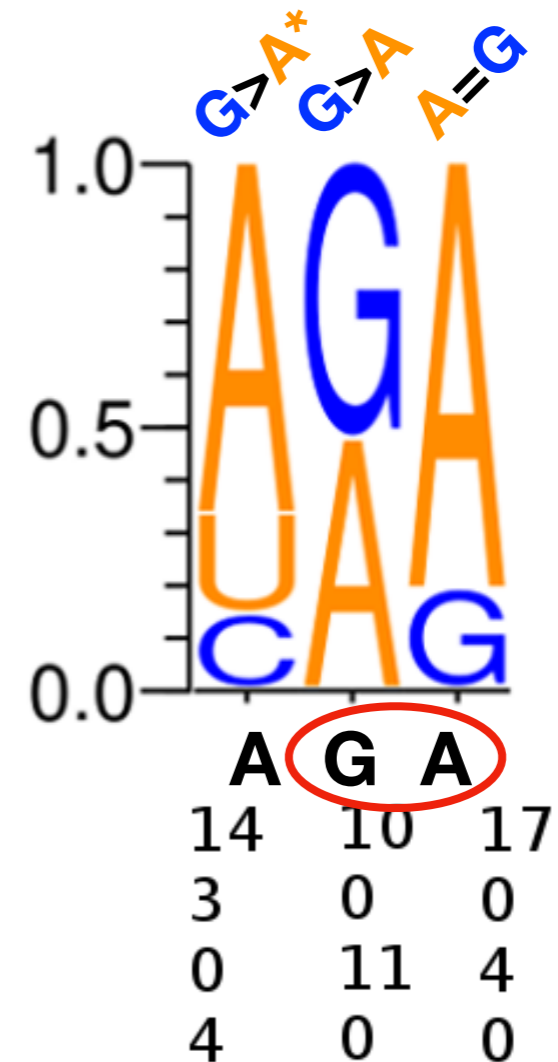
What about “sequence-free constraints” modeling ?



2XNR



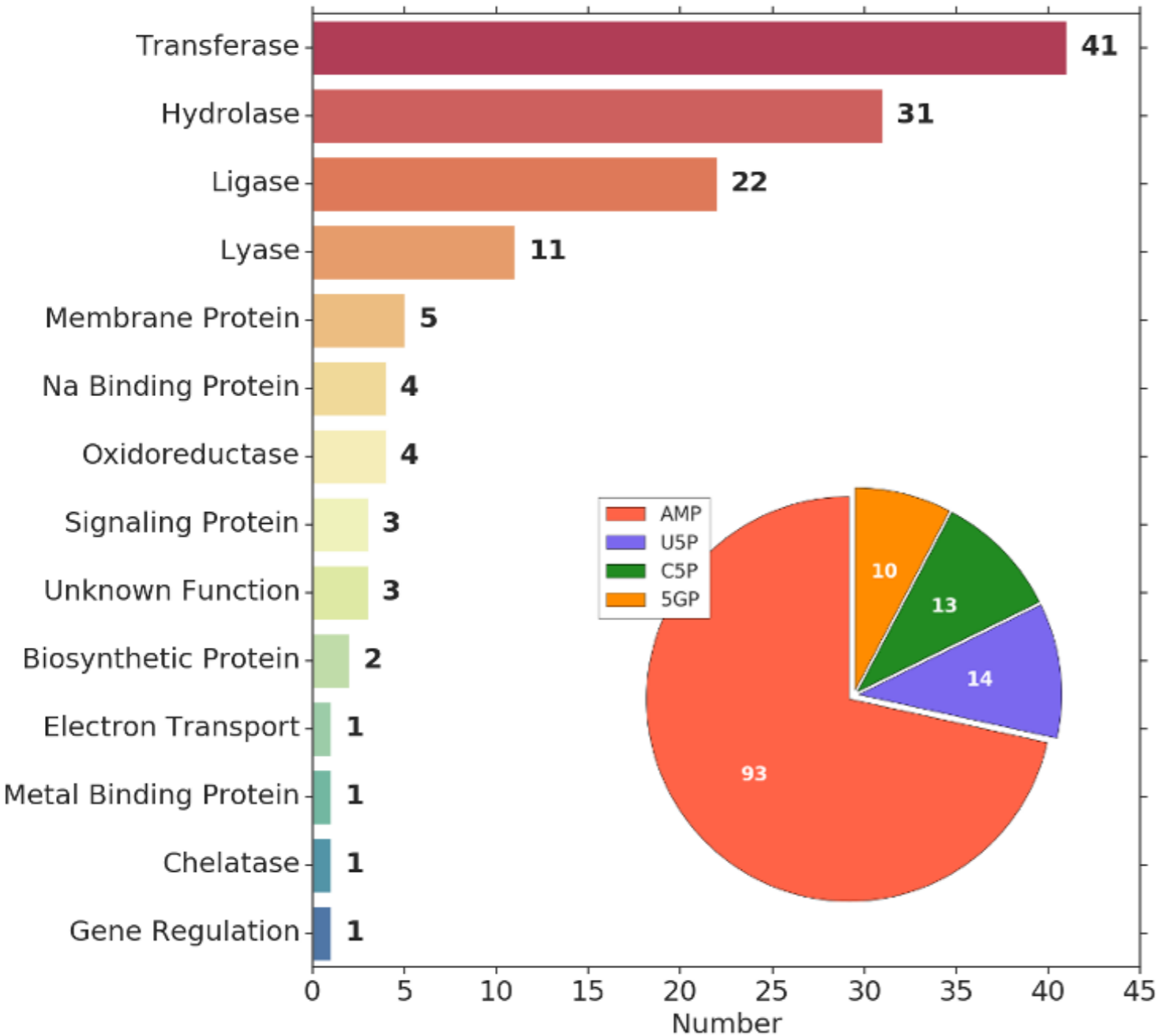
5ELH



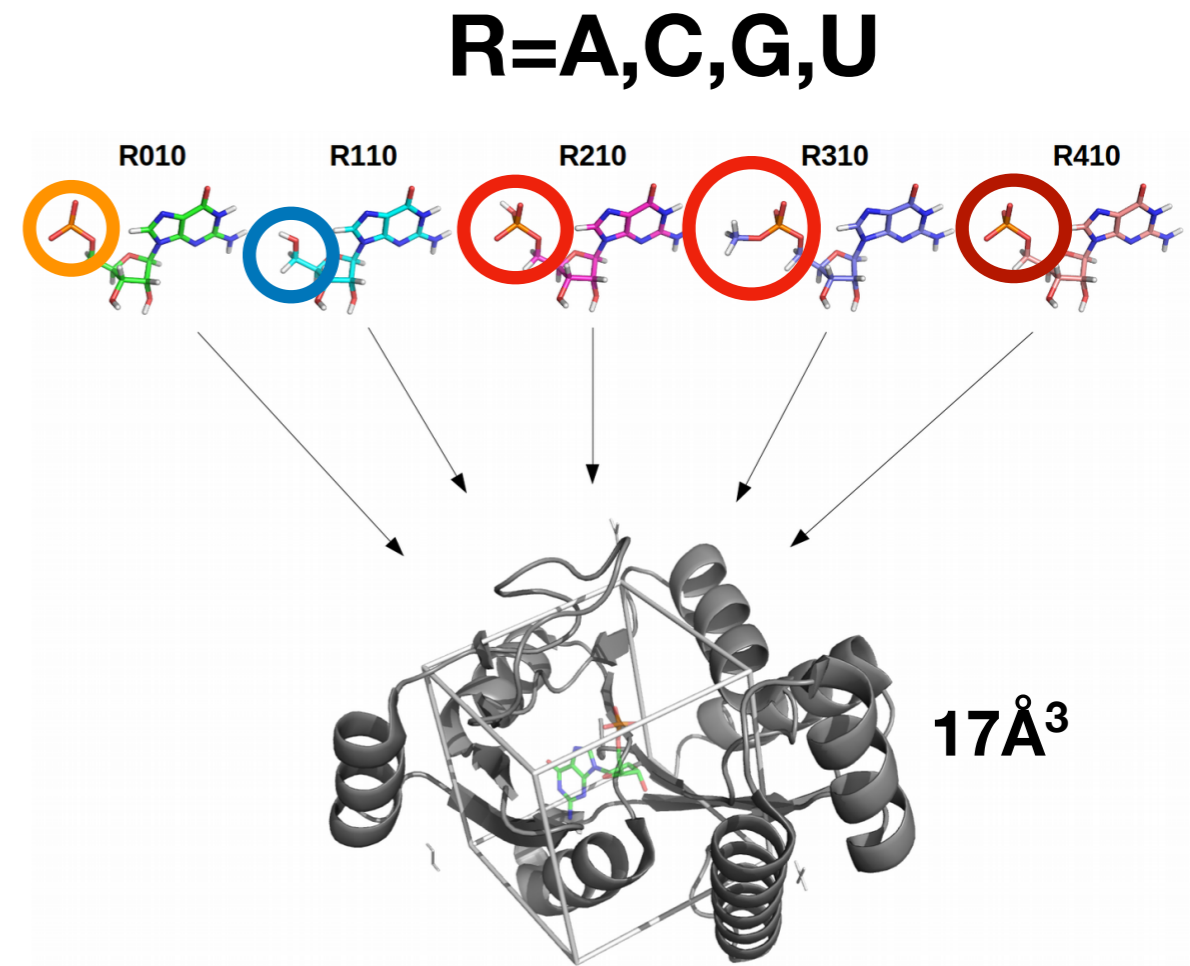
5WWX

Back to the basics ...

Nucleotide Binding Sites



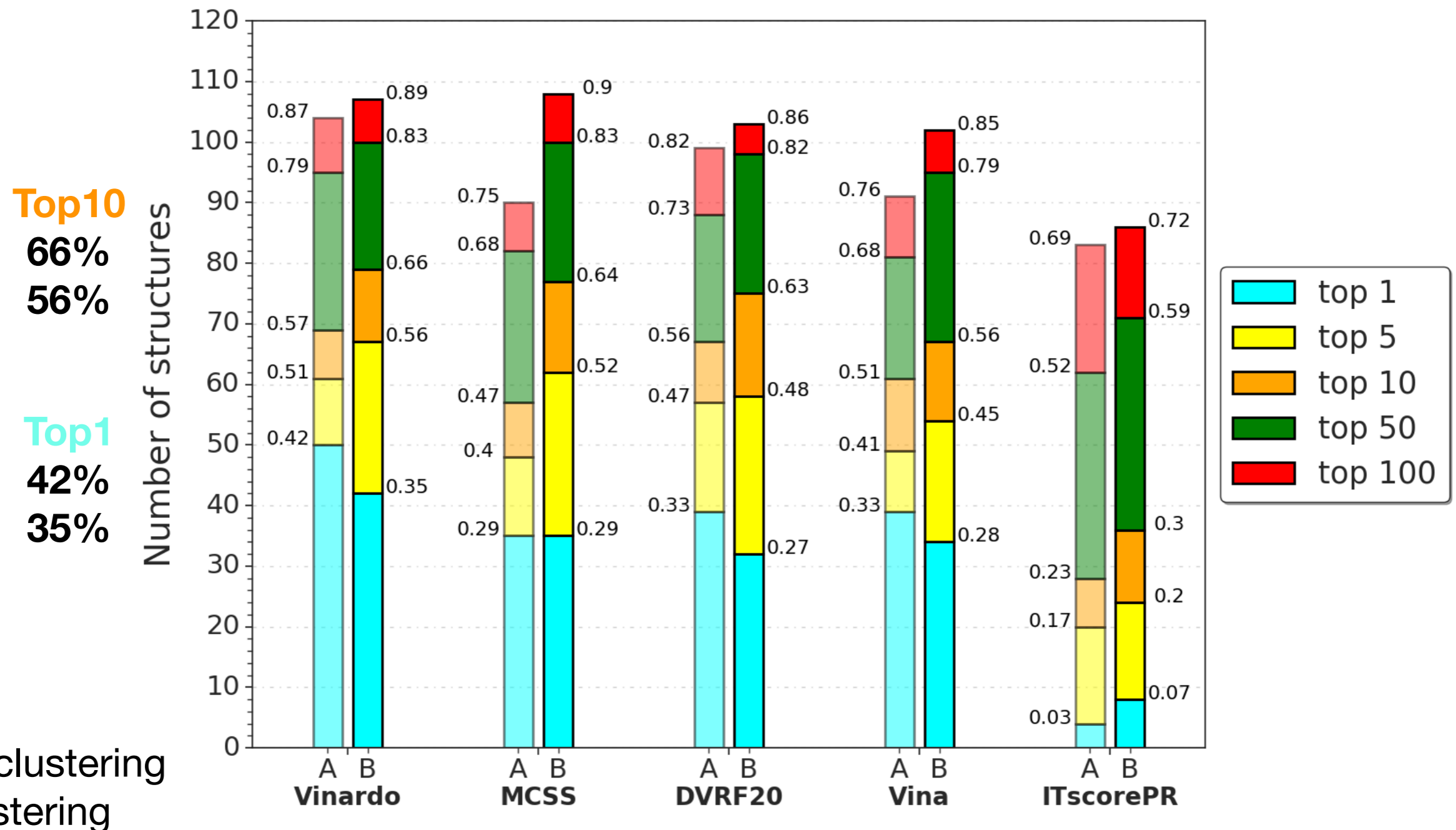
Benchmark120



**Minimized protein
without ligand**

native pose: $\text{RMSD}_{(\text{mcSS}/\text{exp})} \leq 2.0\text{\AA}$

MCSS Performance on nucleotide-binding sites

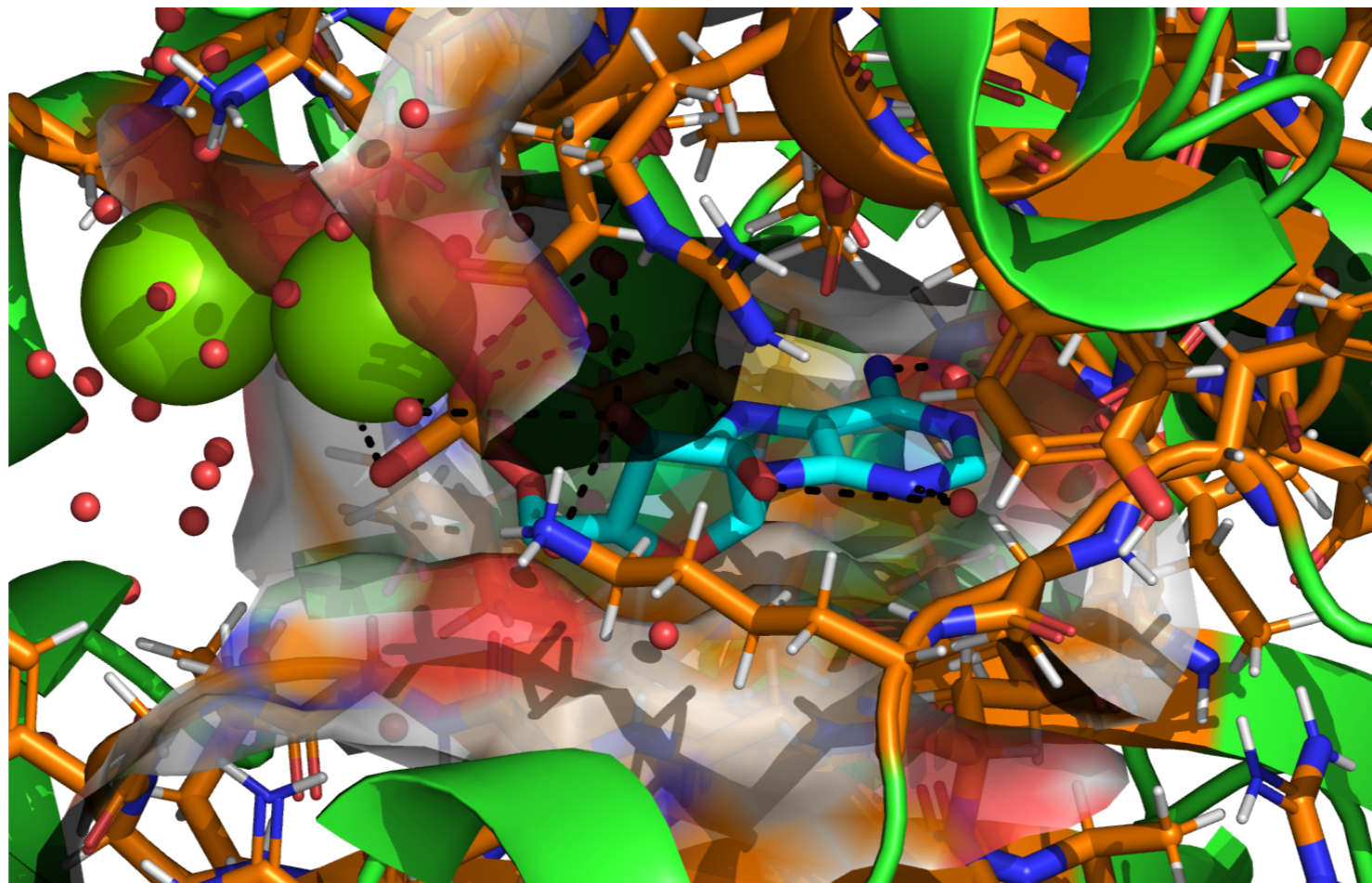


Docking, Scoring Functions and Solvent Models

- empirical scoring function: Vina, etc
- force-field based scoring function: CHARMM (MCSS) & implicit solvent model(s)

• ...

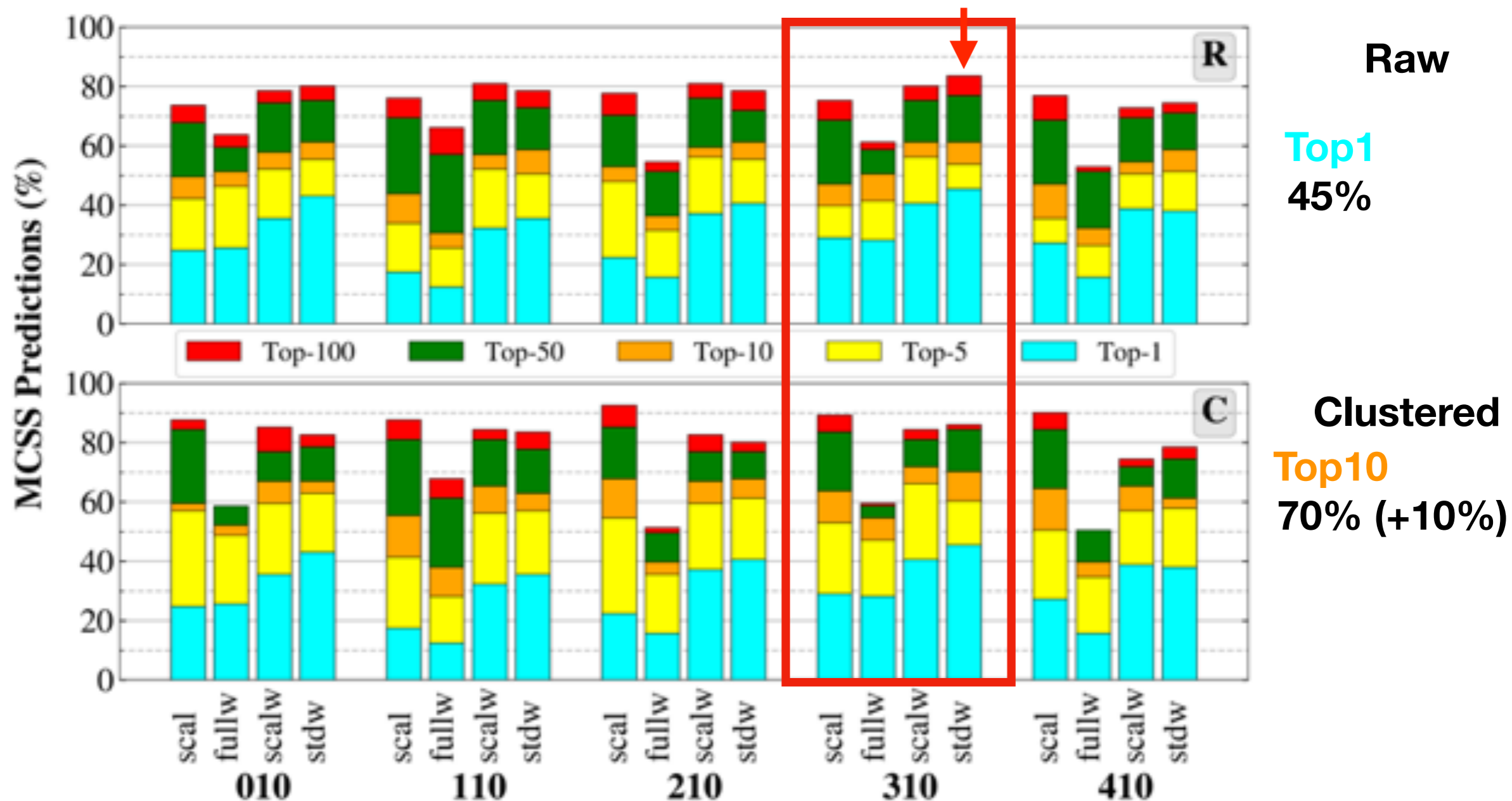
metals
crystallised
waters



Nucleotide Binding
Site

MCSS Performance with implicit/explicit solvent

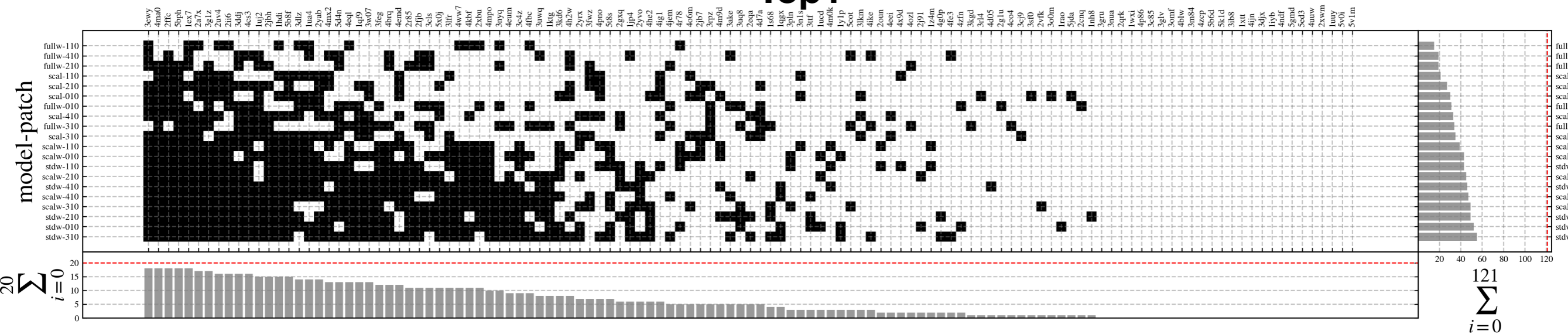
scal: no water / scalw, stdw, full: waters



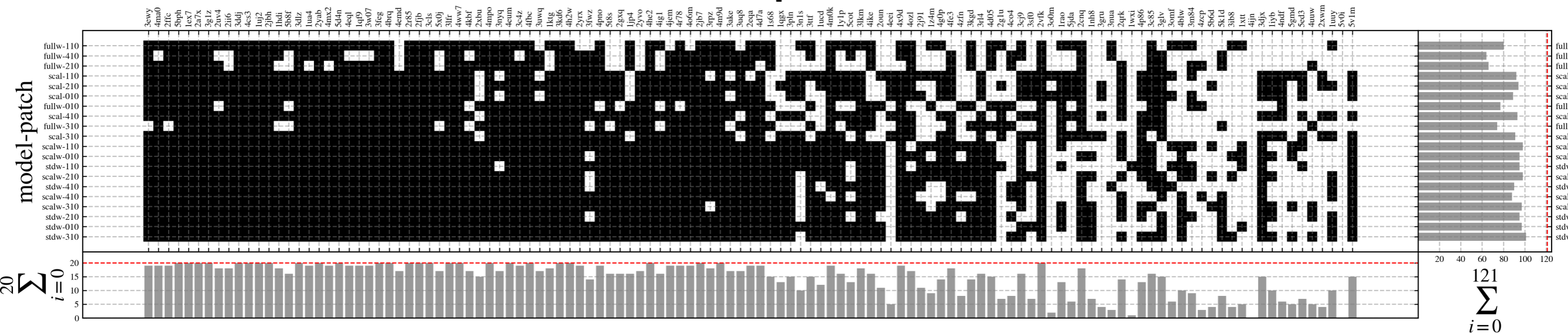
MCSS Performance with implicit/explicit solvent

scal: no water / scalw, stdw, full: waters

Top1



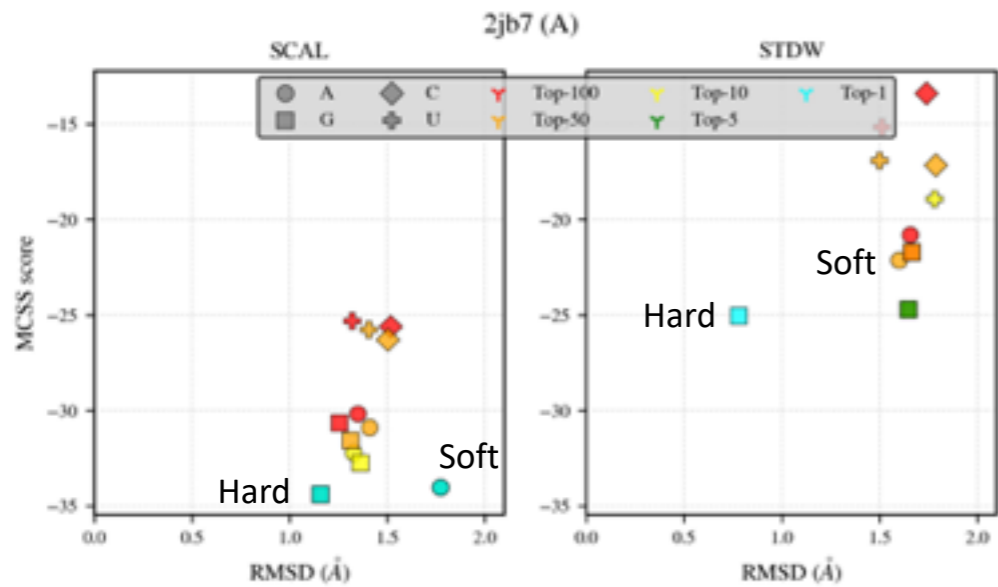
Top100



Unspecific/specific (A/C/G/U): Hard/Soft predictions

scal=implicit solvent

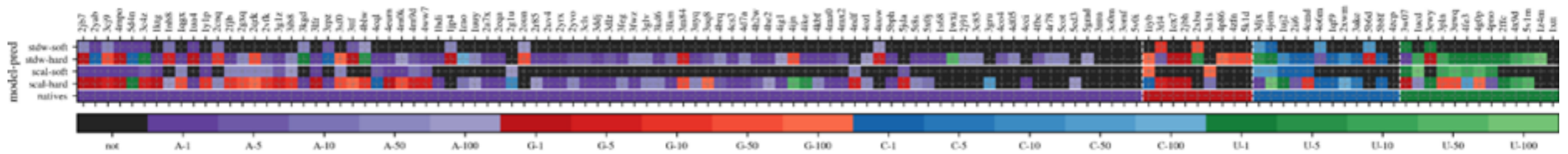
stdw=implicit+explicit solvent



Hard: native pose for specific nucleotide in top n score
Soft: idem but with score up to +2kcal/mol from top n

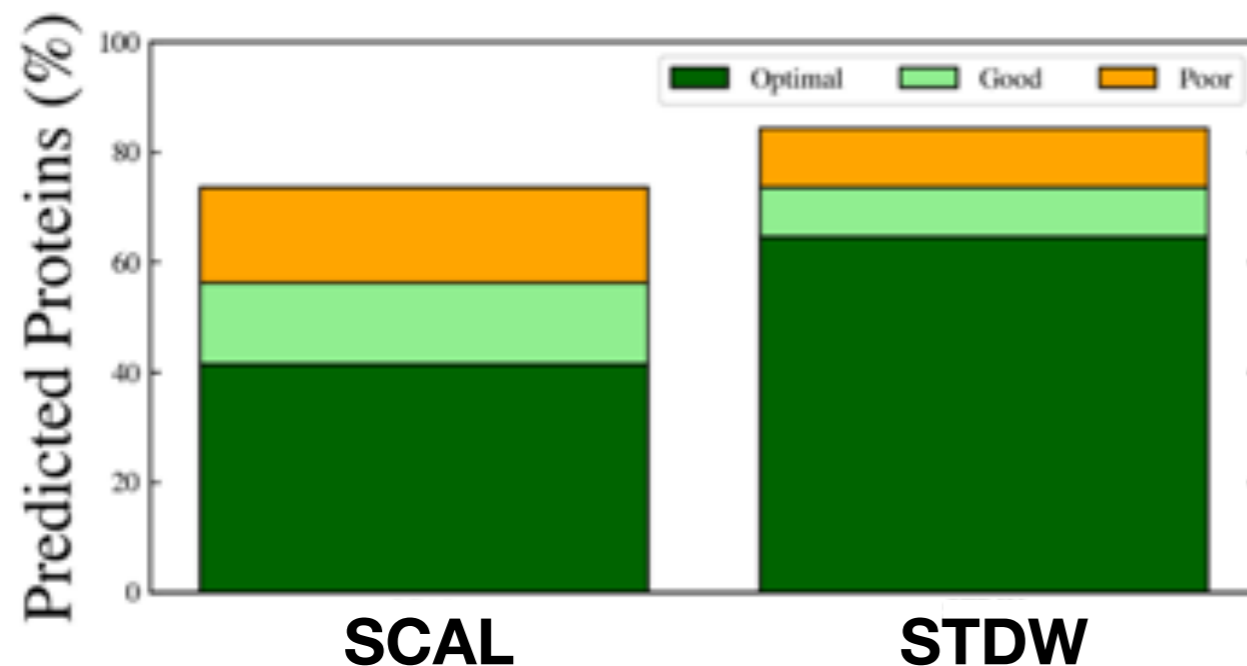
Proteins-Nuc (121)
 A --> 87 G --> 9 C --> 12 U --> 13

Benchmarked Protein

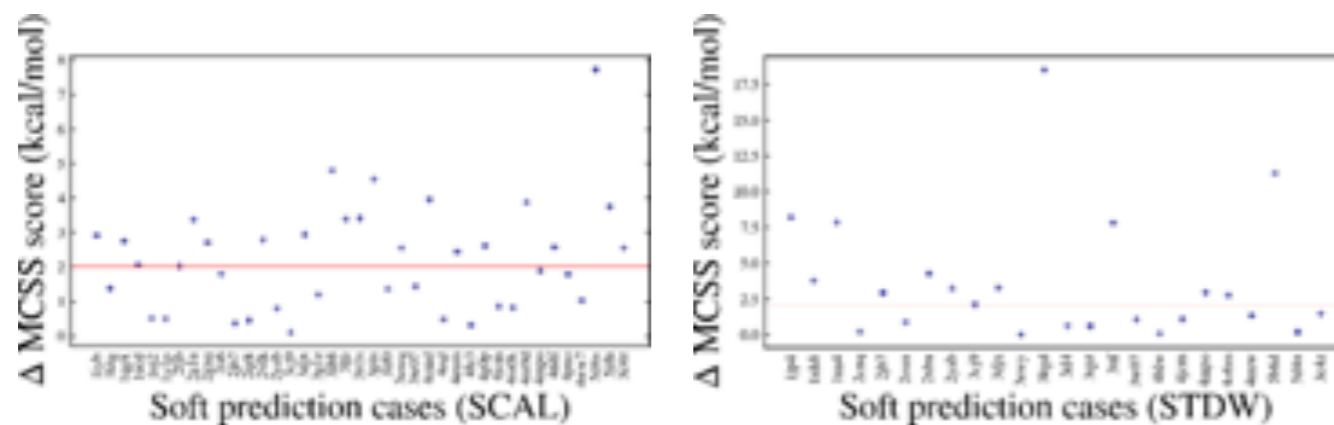


Unspecific/specific (A/C/G/U): Hard/Soft predictions

Top100



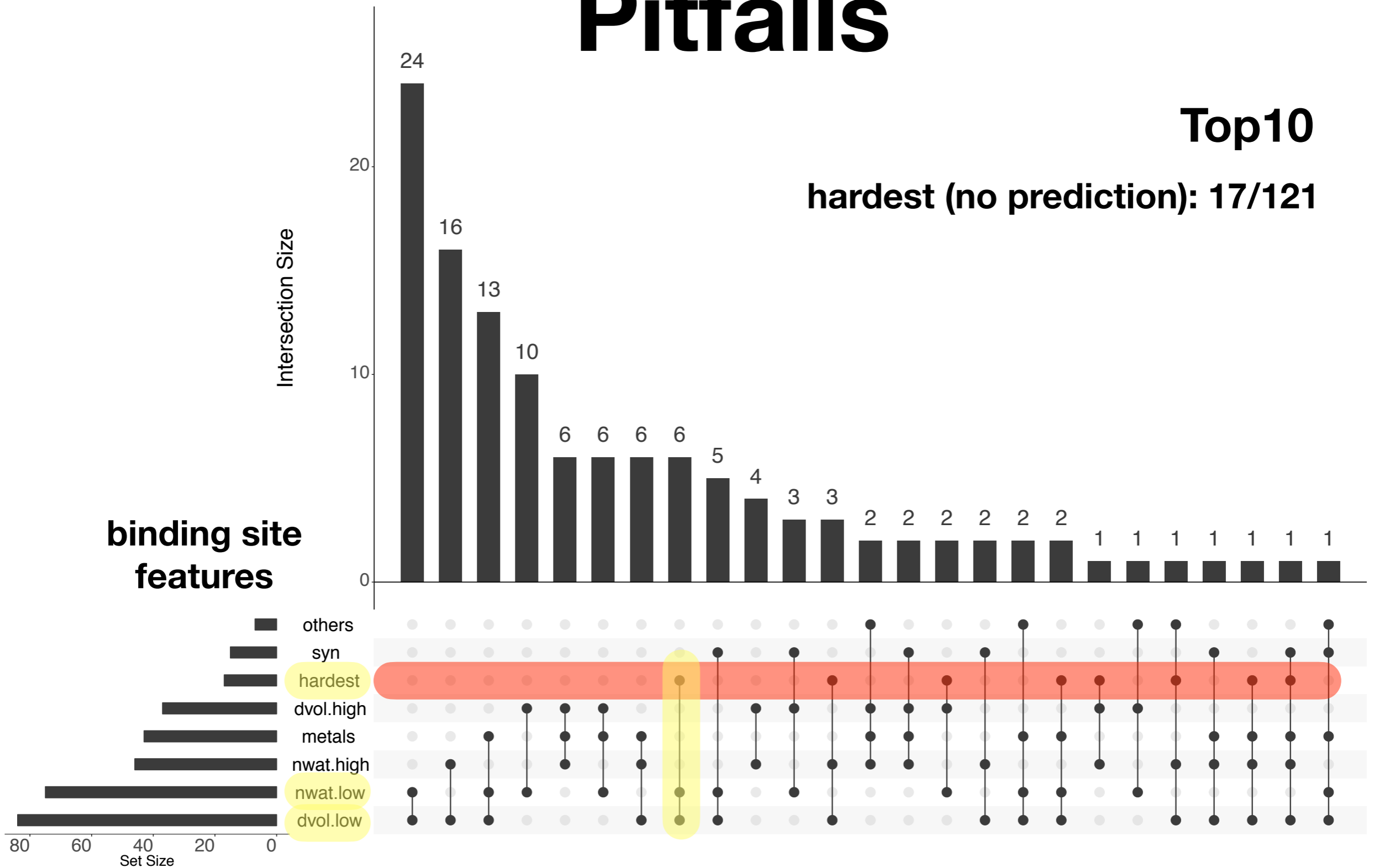
- **Optimal:** hard predictions
- **Good:** soft predictions
- **Poor:** predictions with score over the 2kcal/mol cutoff value from top1 score
- **No-prediction**



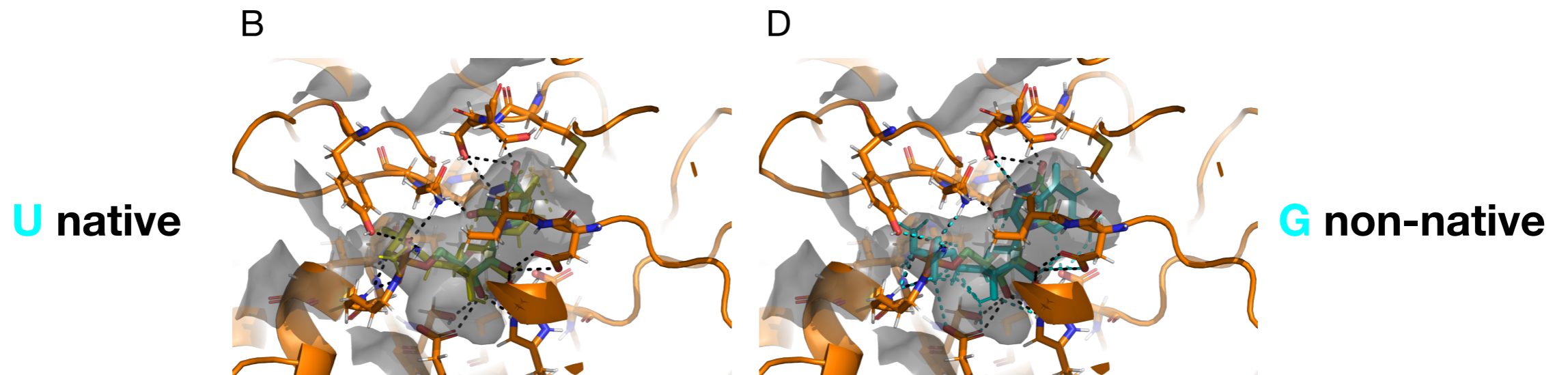
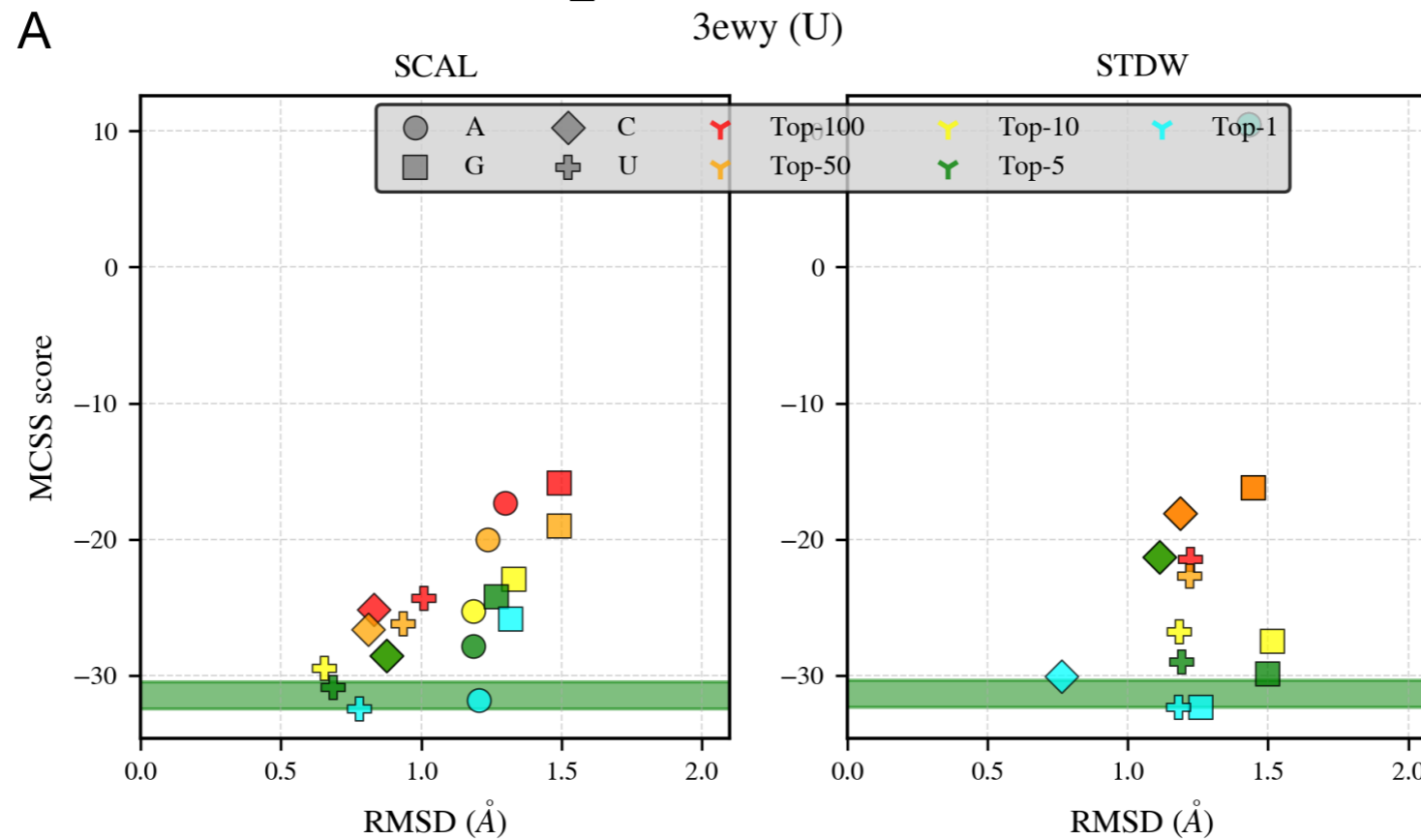
Poor-No Prediction(s) & Pitfalls

Top10

hardest (no prediction): 17/121

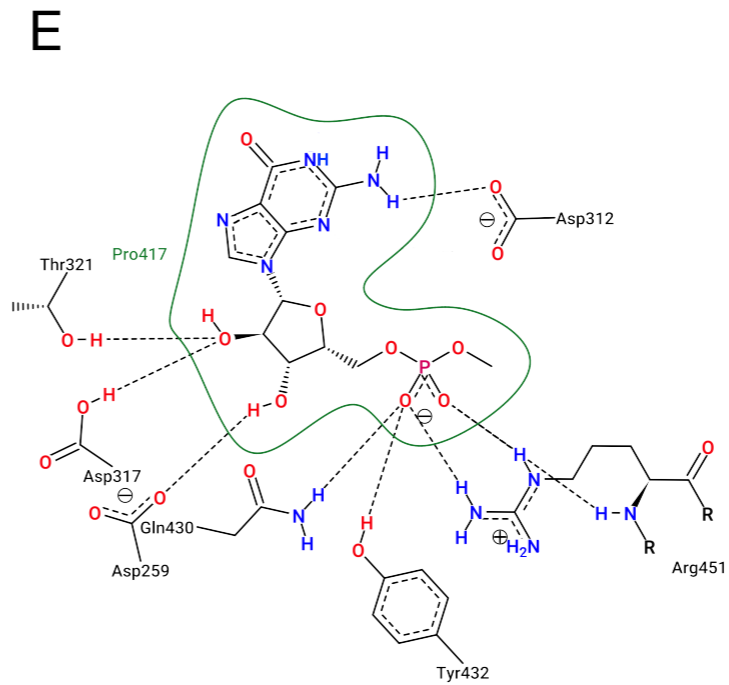
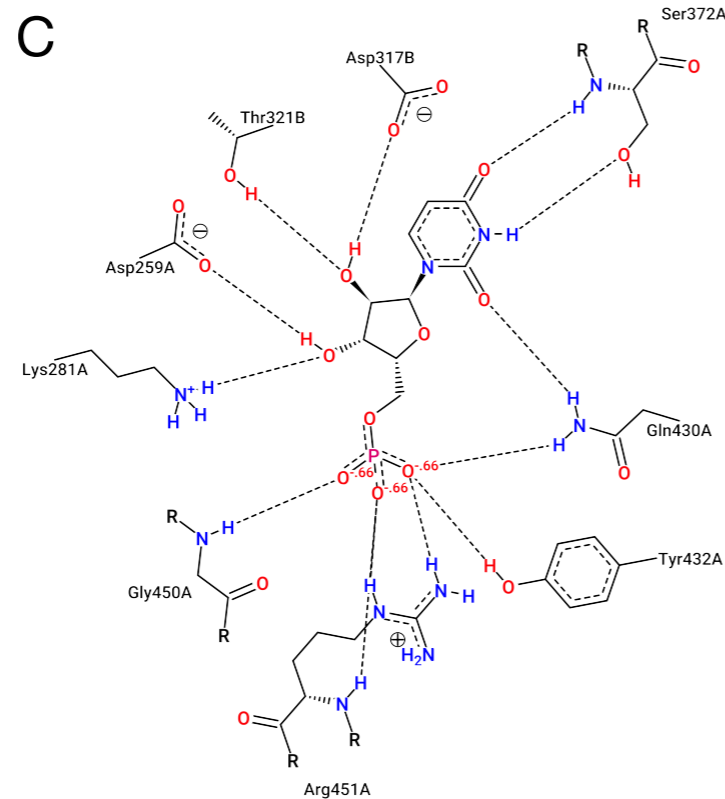


Learning from the analysis of soft predictions



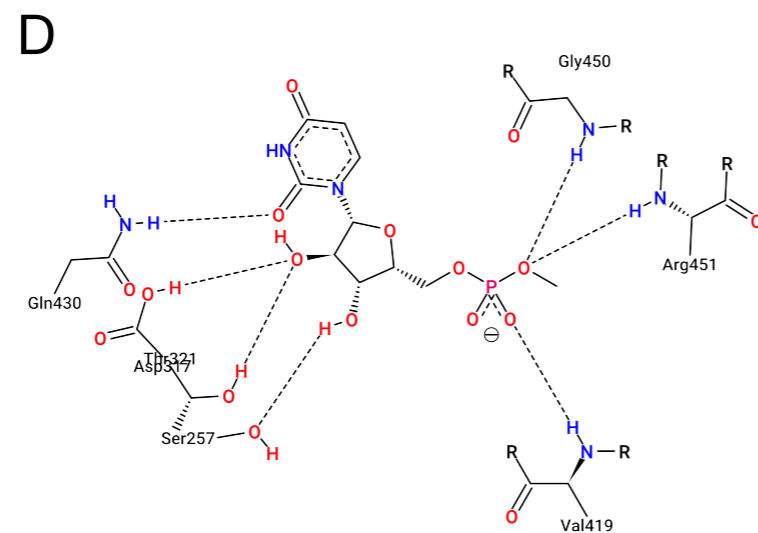
Learning from the analysis of soft predictions

X-ray



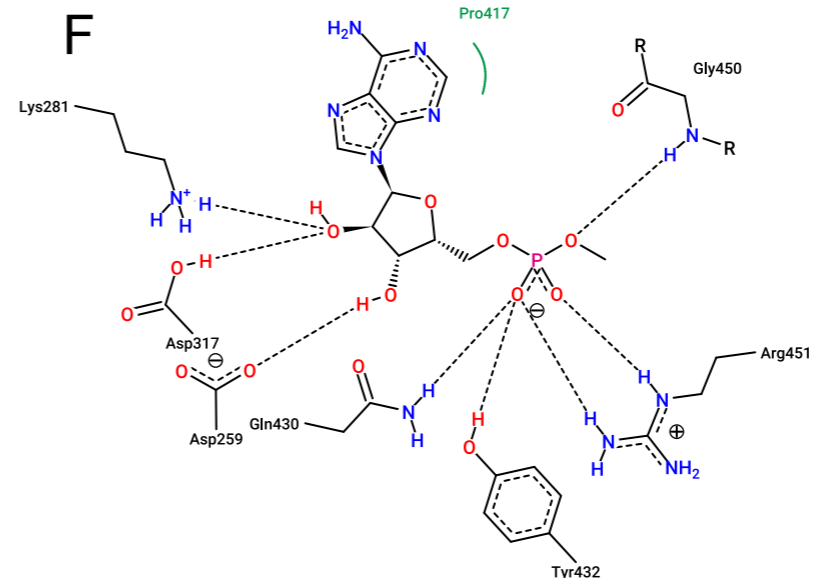
MCSS

G non-native (stdw)



MCSS

U native

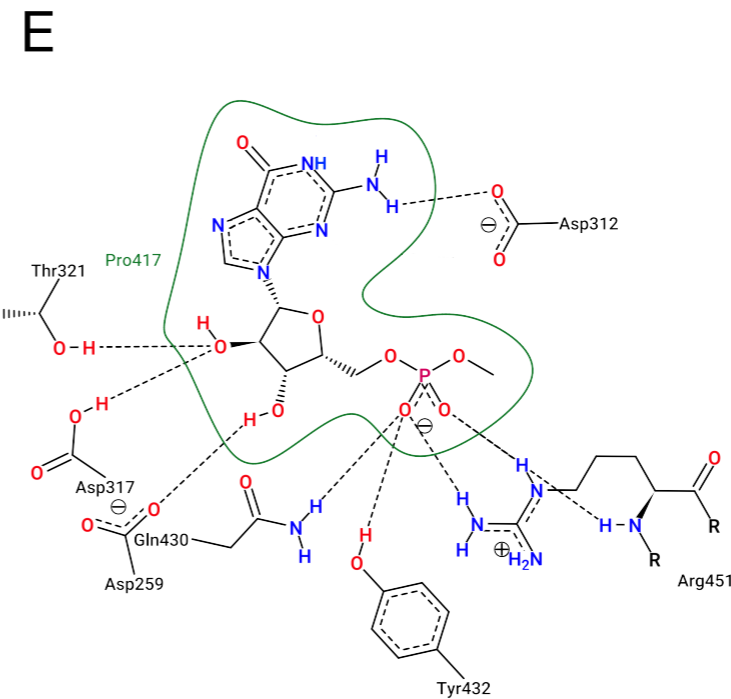
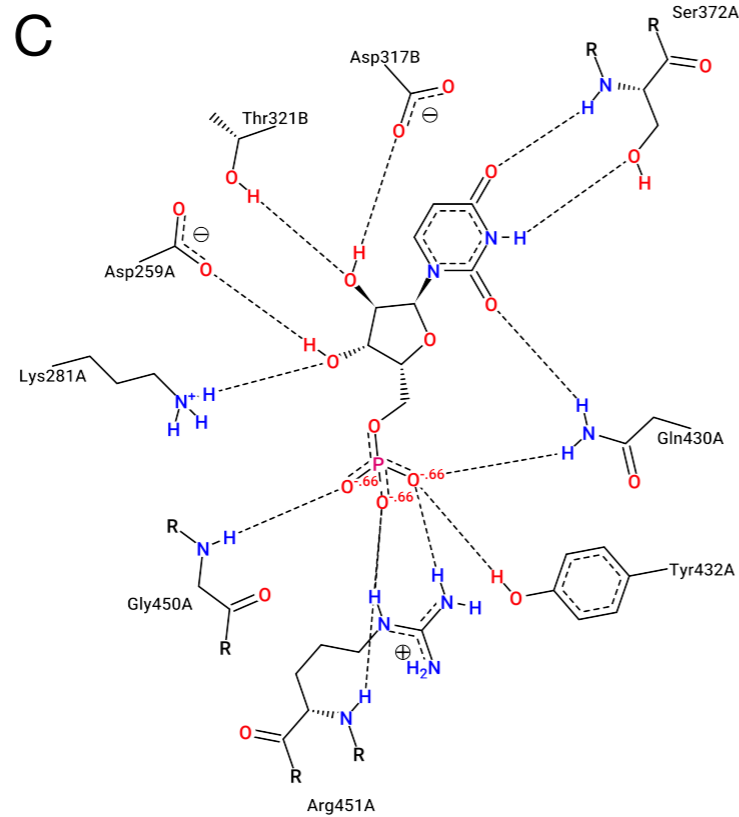


MCSS

A non-native (scal)

Learning from the analysis of soft predictions

X-ray

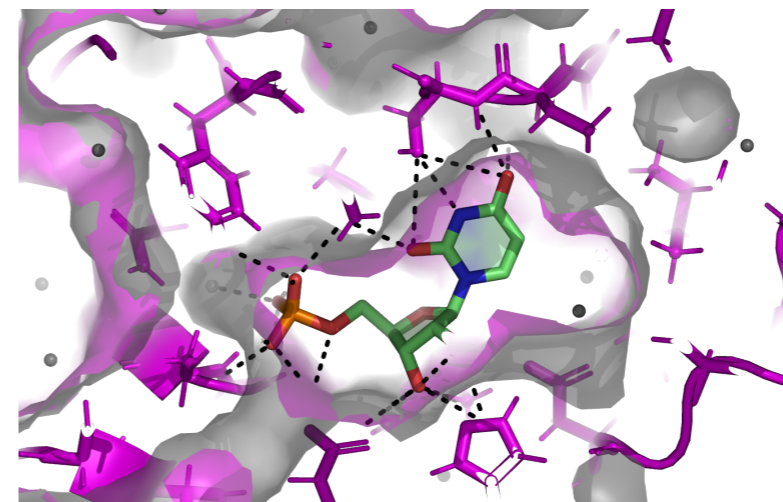
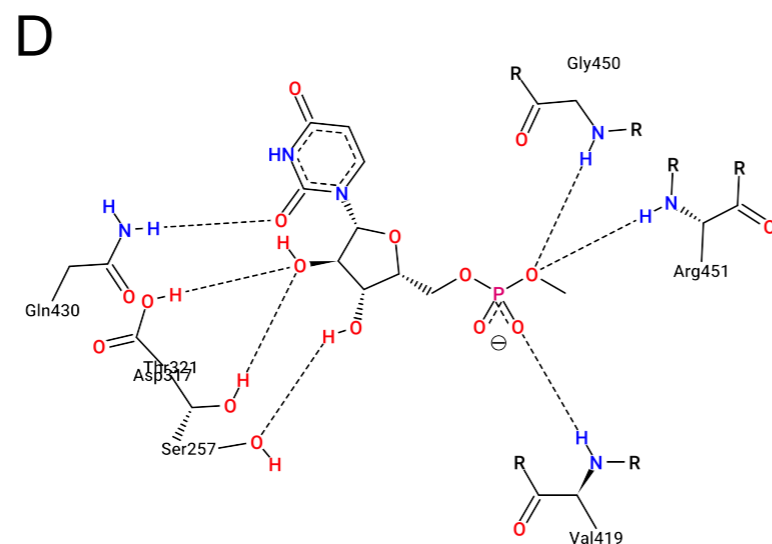


MCSS

G non-native

MCSS

U native



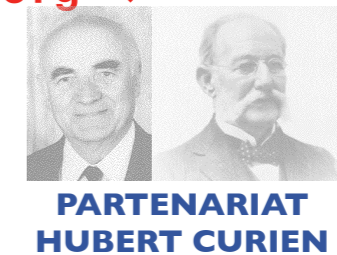
X-ray/minimized

Conclusions

- fragment-based approach applied to **single nucleotide binding**
- **explicit solvent** improve predictions of nucleotide binding
 - performance: Top1: 45%; Top5: 60%; Top10: 70%
 - discrimination of **native/non-native** nucleotide
- pitfalls:
 - **flexibility** of protein binding site, binding thermodynamics, binding kinetics, etc

Acknowledgments

- Roy González-Alemán, Ph.D. student (2019-)
- Nicolas Chevrollier, Ph.D. (2019)
- Coralie Rohmer, MSc (2017).
- Manuel Simoes
- Martin Karplus



Useful Links

- **RNA Collaborative Seminar Series** (Univ. Michigan) - youtube channel

- **The RNA Institute**

HMS Initiative for RNA Medicine Virtual Seminar

June 16 @ 3:00 pm - 4:00 pm

“Epitranscriptome: The Role of RNA Methylation in Stress and Viral Defense”

- **iRNA COSI: Integrative RNA Biology**

Next journal club **June 23 @11:00am EDT**

Johannes Linder, Georg Seelig’s group, University of Washington:

“Engineering Alternative Polyadenylation with Deep Generative Neural Networks”

- **OTS: Oligonucleotide Therapeutics Society:** webinars & Journal Club

- **GDR RNA: RNA as a tool and a target for medicinal chemistry and chemical biology**

