Summer 2022

Design and Synthesis of Hoogsteen-Binding Peptide Nucleic Acid Monomers with Extended Linkers for Triple Helical U-A Recognition in RNA

Angelina B. Giglio-Tos

Tristan L. Mabee

James A. MacKay

Design and Synthesis of Hoogsteen-Binding Peptide Nucleic Acid Monomers with Extended Linkers for Triple Helical U-A Recognition in RNA

Angelina B. Giglio-Tos, Tristan L. Mabee, Dr. James A. MacKay

Department of Chemistry and Biochemistry, Elizabethtown College, 1 Alpha Drive Elizabethtown, PA 17022

ETOWN ELIZABETHTOWN COLLEGE

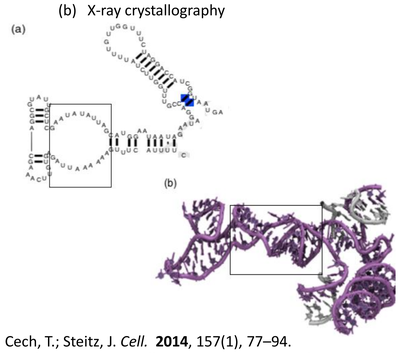
Abstract

RNA provides many functions within biological systems. For example, noncoding RNA (ncRNA), a form of RNA that is not part of transcription or translation, serves a variety of unique roles, such as catalysis or gene regulation. ncRNA generally forms double helical motifs that are ripe for molecular recognition. Sequence selective recognition of double helical RNA (dhRNA) can be achieved using Peptide Nucleic Acids (PNA) through triple helical formation by Hoogsteen hydrogen bonding of PNA nucleobases in the major groove of dhRNA. However, strong, and selective recognition is typically limited to polypurine strands and pyrimidine recognition remains an unsolved problem. A promising solution uses extended nucleobases to reach across the Hoogsteen face of the RNA base pair, bypassing the pyrimidine, and binding with the distal purine. Using this strategy, we designed and synthesized new extended nucleobases to help uncover the ideal linker length and heterocyclic substitution for optimal molecular recognition.

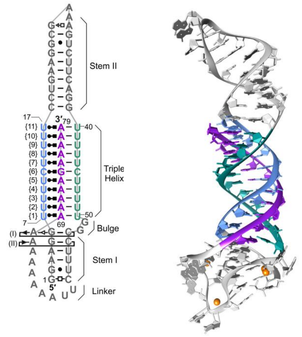
Importance of Noncoding RNA in Biological Systems

- ncRNA acts in biological catalysis, gene regulation, RNA processing, foreign nucleic acid digestion, and other roles within multicellular organisms
- Ribosomes, for example, use RNA to form glycosidic bonds in new proteins
- Double helical motifs in ncRNA are common and are ripe for molecular recognition
- Due to diverse tertiary structures in RNA, natural triple helices are also observed
- Sequence selective recognition of ncRNA would be useful for practical applications in biotechnology

Segment ES6 of the yeast ribosomal subunit
(a) Sequence
(b) X-ray crystallography



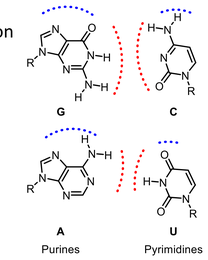
MALAT1 Sequence and Crystal Structure



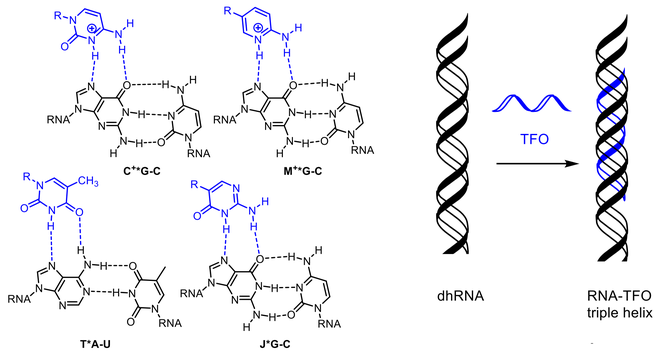
Cech, T.; Steitz, J. *Cell.* **2014**, 157(1), 77–94.
Novikova, I. V.; Hennelly, S. P.; Tung, C.-S.; Sanbonmatsu, K. Y. *J. Mol. Biol.* **2013**, 3731
Ruszkowska, A; Ruszkowski, M.; Hulewicz, J. P.; Duater, Z.; Brown, J. B. *Nucleic Acids Res.* **2020**, 3304.

Hoogsteen Recognition

- Hoogsteen hydrogen bonding is a type of base pairing that occurs on the top face of the Watson-Crick base pair
- Blue denotes Hoogsteen faces - red denotes Watson-Crick
- Pyrimidines have one Hoogsteen hydrogen bonding site, purines have two
- Pyrimidine recognition is more difficult due to fewer hydrogen bonding sites, weak and unselective binding is often observed
- Natural triplexes with U*A-U and C*G-C exist such as in MALAT1



TFO Recognition Strategy

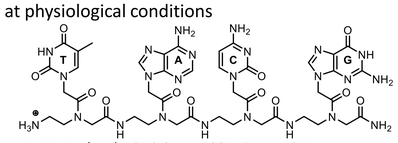


Challenges

- Anionic phosphate backbone results in electrostatic repulsion in RNA triplexes
  - Neutral Peptide Nucleic Acids avoid the electrostatic repulsion in the RNA backbone
- Cytosine (pka ~ 4.5) is not protonated at physiological conditions, poor binding & specificity
  - Rozners used M (pka ~6.7) as a protonated analog of C to obtain stronger binding
- Strong and selective recognition is limited to poly-purine regions of RNA
  - This challenge remains unsolved

Peptide Nucleic Acid (PNA)

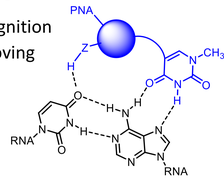
- Neutral amide PNA backbone mimics RNA/DNA phosphodiester backbone while avoiding electrostatic repulsion
- PNA backbone is resistant to nucleases
- Straightforward synthesis
- Work by Rozners and MacKay groups investigates nucleobase modifications that maximize selectivity and affinity at physiological conditions

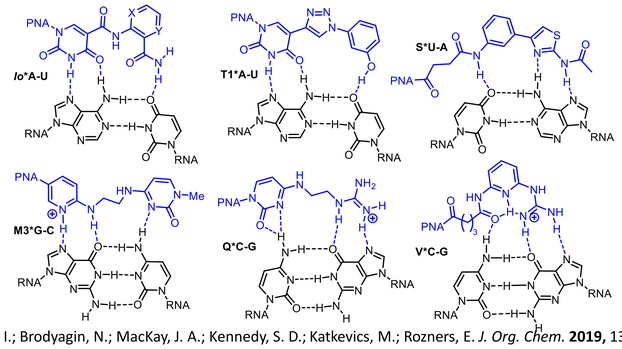


Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497.
Nielsen, P. E.; Egholm, M.; Buchardt, O. Bioconjugate Chem. **1994**, 5, 3–7.

Extended Nucleobases

- Extended nucleobases are a potential solution to pyrimidine recognition
- The Hoogsteen face allows for up to three hydrogen bonds, improving selectivity and affinity
- Extended conjugation may also improve π-stacking
- Figure at right shows PNA binding A for improved U-recognition
- Selected examples of extended nucleobases are below:





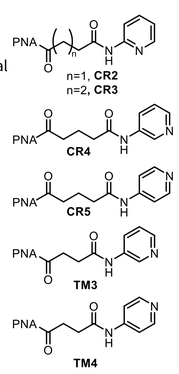
Kumpina, I.; Brodyagin, N.; MacKay, J. A.; Kennedy, S. D.; Katkevics, M.; Rozners, E. *J. Org. Chem.* **2019,** 13276.
Brodyagin, N.; Maryniak, A. L.; Kumpina, I.; Talbott, J. M.; Katkevics, M.; Rozners, E.; MacKay, J.A. *Chem. Eur. J.* **2021**, 27, 4332. Zhan, X.; Deng, L.; Chen, G. *Biopolymers* **2022**, 113, 23476.

Recognition by Aminopyridine Derivatives

- **CR2**-**5** were designed to selectively recognize C
- **CR2** binds A well and C poorly, unexpected result

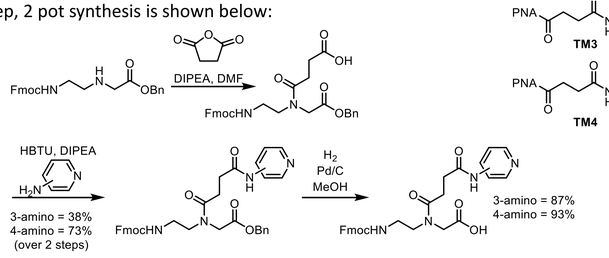
PNA	G	A	C	U
T	46.3 ±0.5	**69.6** ± 0.8	35.4 ± 0.4	34.6 ±0.2
CR2	NB	55.5 ± 0.3	40.3 ± 0.3	34.0 ± 0.4
CR3	30.3 ± 0.5	39.1 ± 0.3	**51.5 ± 0.3**	34.5 ± 0.2
CR4	31.8 ± 0.3	37.1 ± 0.5	**51.2 ±0.3**	34.6 ± 0.3
CR5	32.6 ±0.4	37.8 ± 0.1	**52.5 ± 0.3**	35.7 ± 0.3

Binding affinities determined via UV thermal melting (T_m, °C).
NB = no melting curve observed



Ryan, C.; Baskevics, V.; Katkevics, M.; Rozners, *E. Chem. Commun.* **2022**, 58, 7148.

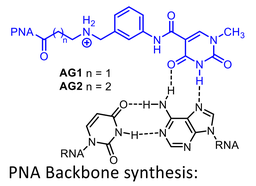
- To better understand binding, we have prepared 3- and 4-aminopyridine analogues of **CR2** (**TM1** & **TM2**)
- A 3 step, 2 pot synthesis is shown below:

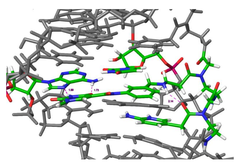


Acknowledgements

- Elizabethtown College Department of Chemistry & Biochemistry
- MacKay Research Group
  - John Talbott (’22)
  - Kyle Hess (’23)
  - Emily Harding (’23)
  - Hallie Bleam (’24)
- Rozners Research Group, Binghamton University
  - Nikita Brodyagin
  - Ilze Kumpina
  - Chris Ryan
  - Brandon Tessier
- Latvian Institute for Organic Synthesis – Martins Katkevics, Vladislavs Buškevičs
- Prestigious Scholarships & Fellowships Committee, Elizabethtown College
- Dr. Charles D. Schaeffer, Jr. Student & Faculty Research Fund provided by Dr. Joseph and Barbara Albanese
- NSF CHE-1708699 & CHE-2107911 (JAM)
- NSF CHE-1708761 & CHE-2107900 (ER)

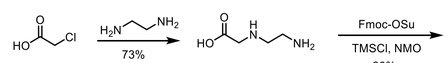
U Recognition Through A-binding

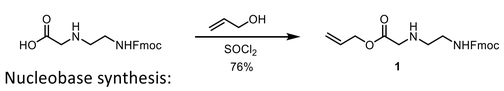




- Cationic bases designed to reach across U to bind with distal A
- Molecular mechanics (through Latvian collaborators) suggests that **AG1** should have ideal binding

PNA Backbone synthesis:





Extended Nucleobase synthesis: