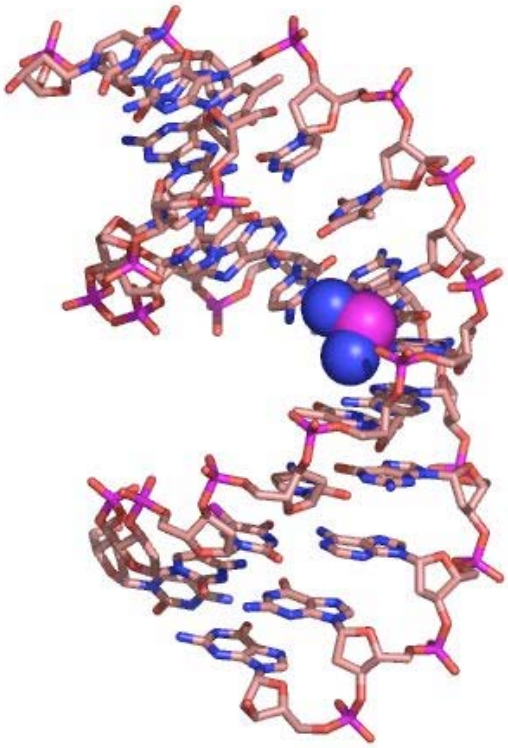
Small Molecule-Nucleic Acid interactions

- Drugs, cytotoxins, probes, regulation
- Forces
  - H-bonding
  - Electrostatics
  - van der Waals
  - Stacking
  - Shape complementarity
- Major & minor grove, intercalators, 3° structures
- Specificity (actually, selectivity)
  - Differential binding affinity to one partner over another

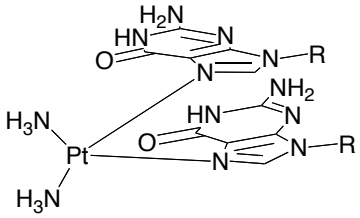
DNA-binding ligands

- Most DNA-binding ligands are not very sequence-specific
  - pyrolle-imidazole compounds (distamycin/netropsin analogs) are the most promising exceptions
- Binding:
  - Intercalators (mostly via major groove)
  - Non-intercalators (mostly via minor groove)
- Covalent
  - Become attached to DNA/RNA
  - Result in single-strand cleavage
  - Result in double-strand cleavage
- Can alter the nucleic acid structure
  - Local
  - Global
- Usually toxic/carcinogenic (utility as drugs is limited)

Cisplatin/DNA adduct

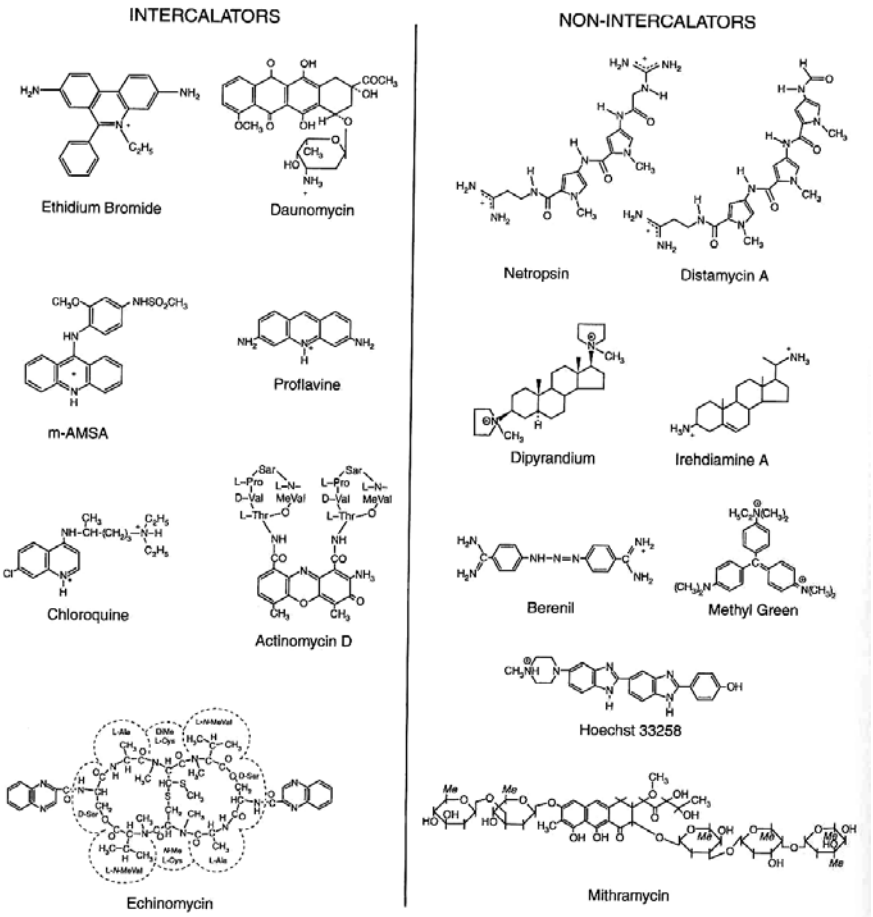


Anticancer agent
Preferential binding to GpG
Leads to DNA bending
Interferes with transscription
General toxin

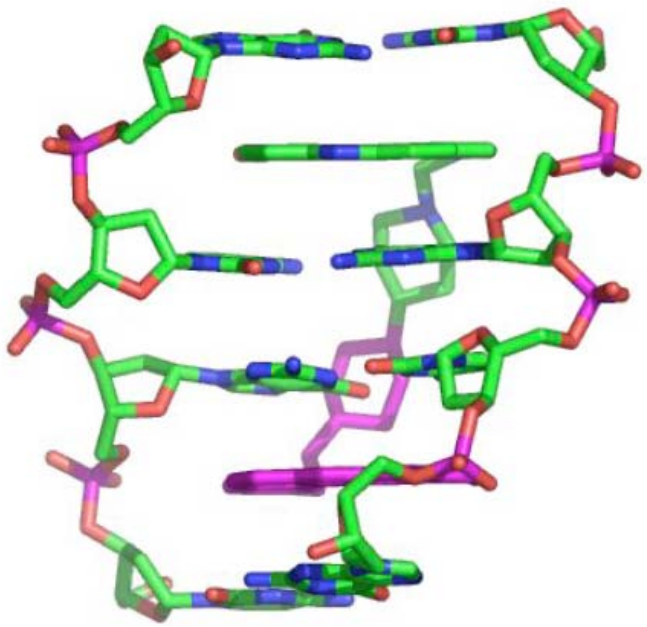


1AIO.pdb

DNA-binding small molecules



Intercalation by Ditercalinium



1D32.pdb

DNA cleaving agents - mitomycin C

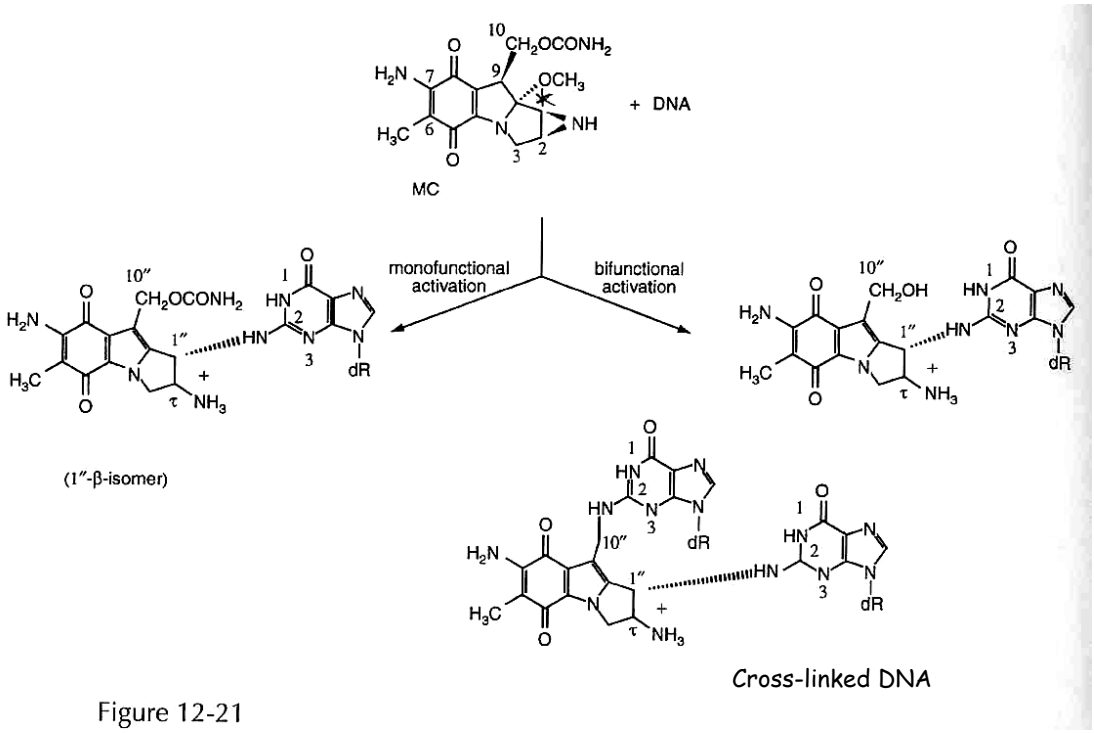
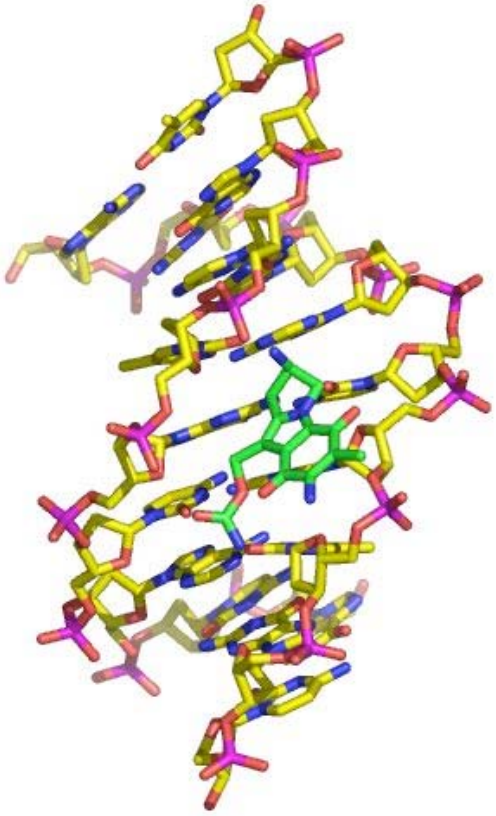
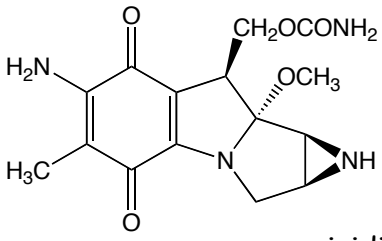


Figure 12-21

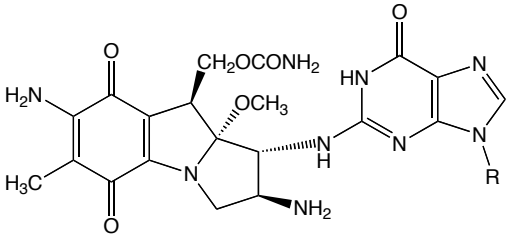
Mitomycin C-DNA covalent complex



carbamate



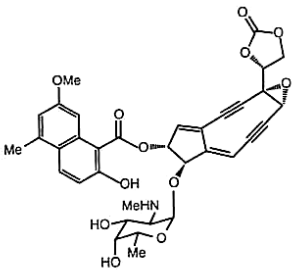
aziridine



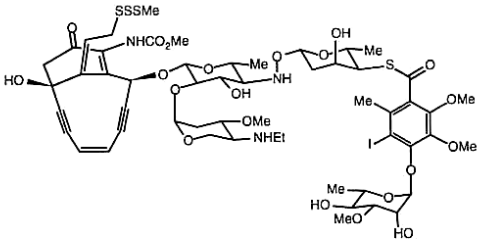
Guanine adduct

199D.pdb

Enediyne antibiotics

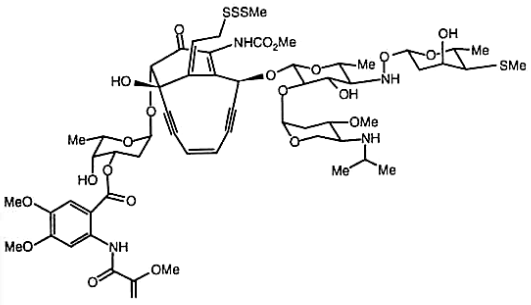


1

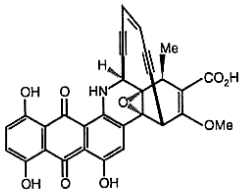


2

Very reactive
Low sequence selectivity
Toxic



3



4

Figure 12-26
Structures of the enediyne antibiotics. Neocarzinostatin chromophore (**1**), calicheamicin γ_1^I (**2**), esperamicin A_1 (**3**), and dynemicin A (**4**). [Reprinted with permission from Nicolau and Dai, 1991.]

DNA cleavage by enediyne antibiotics

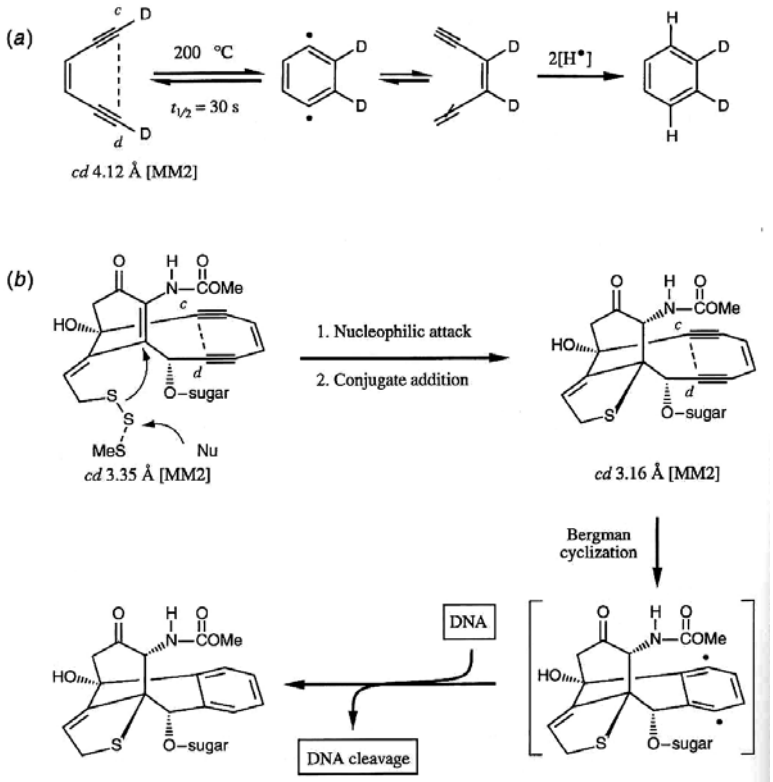
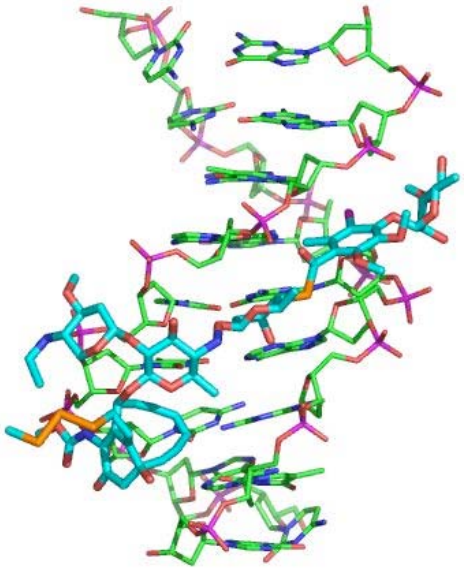
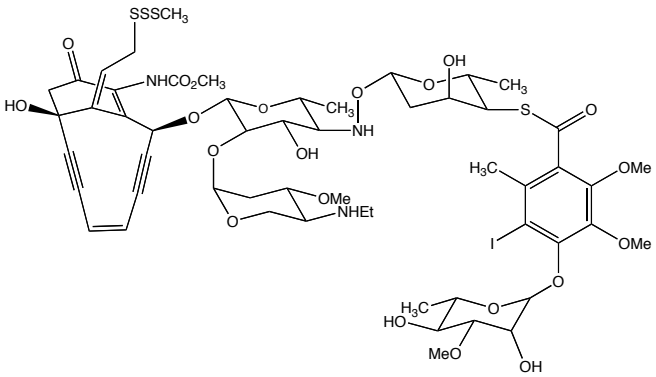


Figure 12-27
(*a*) The Bergman cycloaromatizion reaction. The distance between the centers *c* and *d*, critical for the rate of the process, is calculated to be 4.12 Å. (*b*) The internally triggered Bergman cyclization reaction in the calicheamicin chromophore. [Reprinted with permission from Nicolau and Dai 1991.]

Calicheamicin-DNA complex

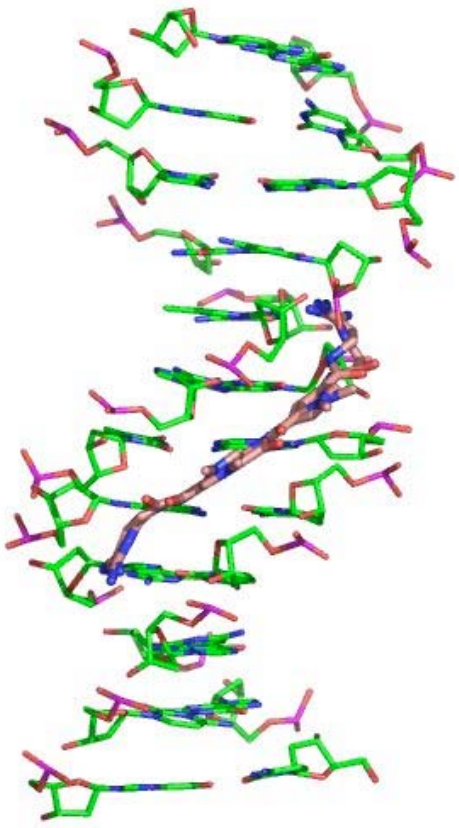




(non-covalent)

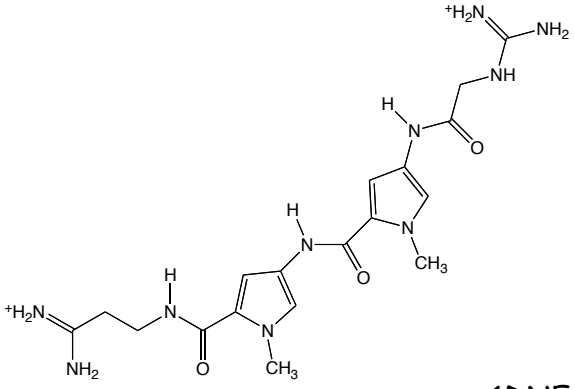
2PIK.pdb

Minor groove binding by netropsin



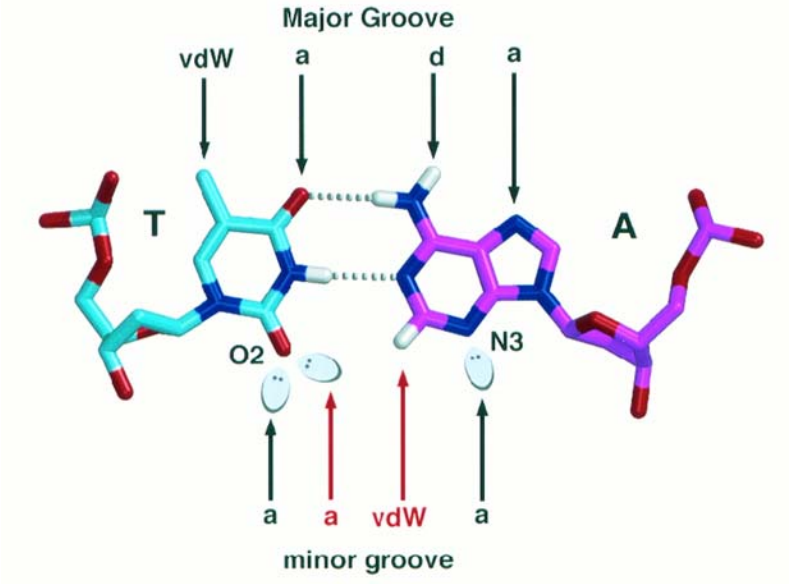
Crystal structure of the netropsin-DNA complex
Netropsin binding in minor groove
Two orientations

Artifact: in solution, two netropsin molecules bind side-by side.

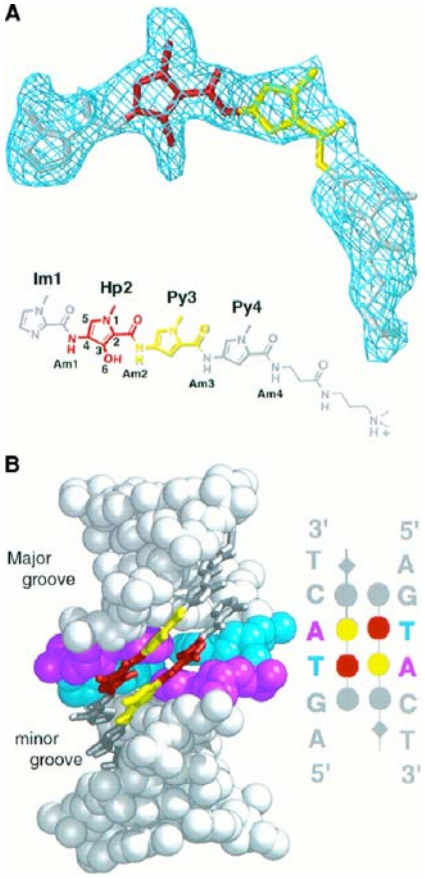


1DNE.pdb

A structural basis for recognition of A.T and T.A base pairs in the minor groove of B-DNA.



Kielkopf CL, White S, Szewczyk JW, Turner JM, Baird EE, Dervan PB, Rees DC. Science. 1998 Oct 2;282(5386):111-5.



- (A) Omit $|F_o| - |F_c|$ electron density map for one of the ImHpPyPy polyamide molecules, contoured at 1.5, showing the position of the 3-hydroxyl group. The numbering of the atoms used in the text is indicated below on the chemical structure. The Hp is red and the Py that would be paired with it is yellow. The Im, the other Py, , and Dp are silver.
- (B) Space-filling model of (ImHpPyPy)2 · 5'-CCAGTACTGG-3'. Adenosine is purple and thymidine cyan; polyamide is colored as above. A schematic is shown to the right, with the aromatic residues of the polyamide indicated by filled circles and by the diamonds. The overall structure of (ImPyPyPy)2 · 5'-CCAGTACTGG-3' is similar.

Kielkopf CL, White S, Szewczyk JW, Turner JM, Baird EE, Dervan PB, Rees DC. Science. 1998 Oct 2;282(5386):111-5.

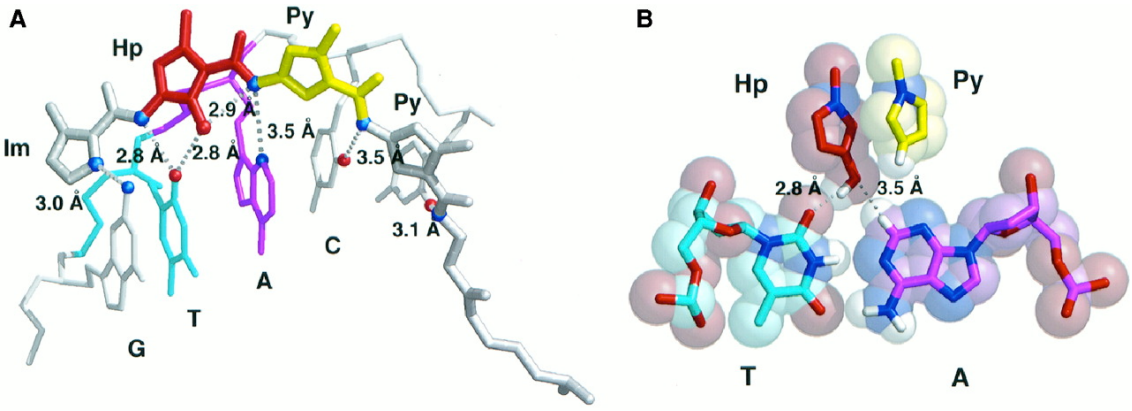
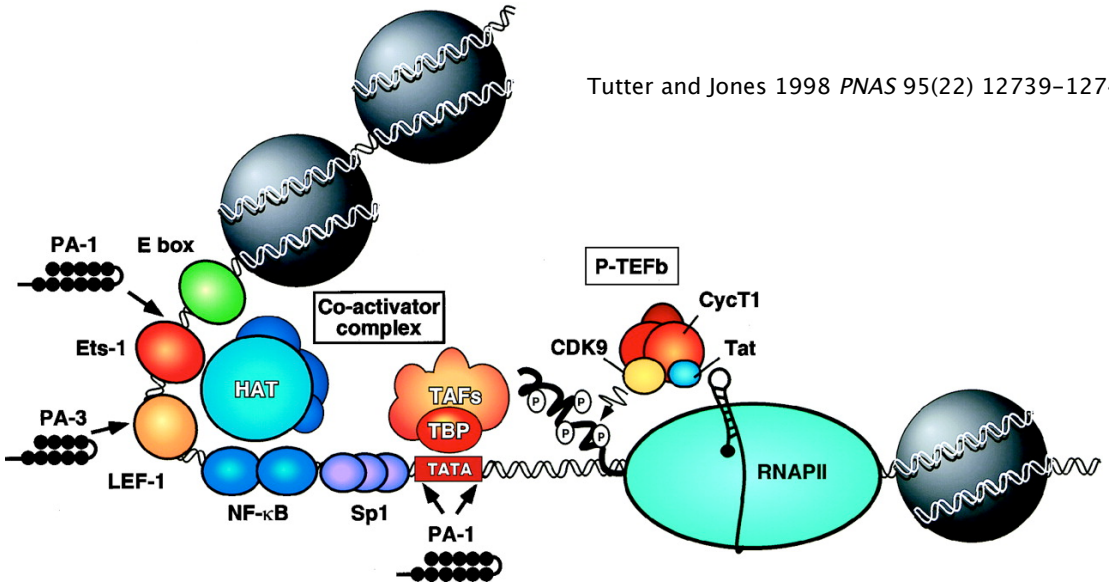


Figure 3. (A) The hydrogen bonds between ImHpPyPy and one strand of DNA, indicated by dashed lines. (B) Space-filling model of the Hp/Py pair interacting with the T·A base pair shows that the Hp-OH tightly fits the cleft formed by the adenine-C2H.

Kielkopf CL, White S, Szewczyk JW, Turner JM, Baird EE, Dervan PB, Rees DC. Science. 1998 Oct 2;282(5386):111-5.



View of the integrated HIV-1 proviral transcription control region. Shown are the recognition sites for the HIV-1 enhancer-binding factors. The enhancer promotes initiation of RNAPII transcription through the recruitment of a coactivator complex(es) that contains associated histone acetyltransferase (HAT) activities. Transcription initiation also requires the binding of Sp1 to the promoter, as well as the TATA-binding protein, TBP, and associated factors (TAFs). The HIV-1-encoded transcription factor Tat interacts with the cyclin T1 subunit of P-TEFb to direct the P-TEFb complex to nascent TAR RNA and enhance elongation of RNAPII transcription. The binding sites for polyamides used by Dickinson et al. (1) to target HIV-1 promoter and enhancer sequences are shown: Polyamide 1 (PA-1) blocks binding of TBP and ETS-1 whereas Polyamide 3 (PA-3) blocks the binding of the lymphoid enhancer-binding factor LEF-1.

Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands

Dickinson *et al.* 1998
PNAS 95(22):12890-5

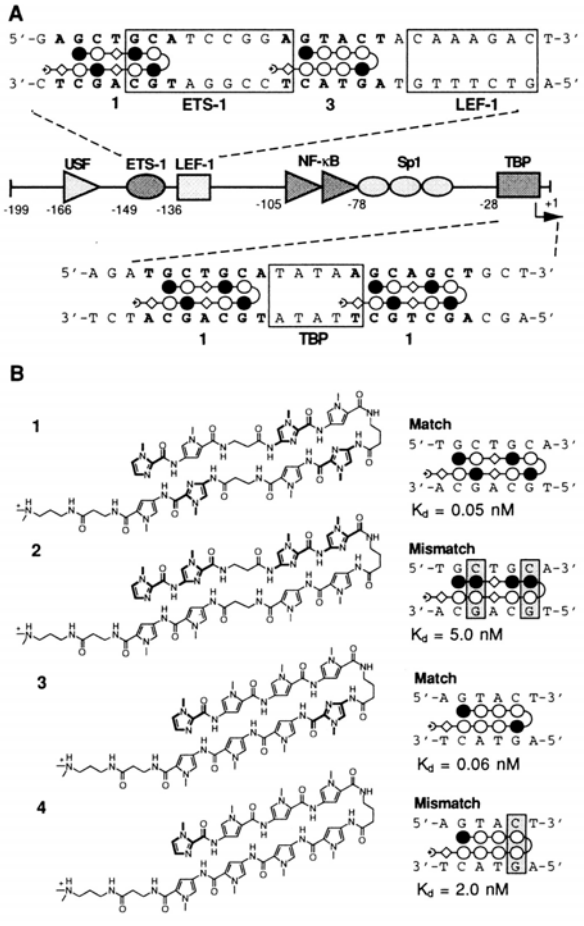
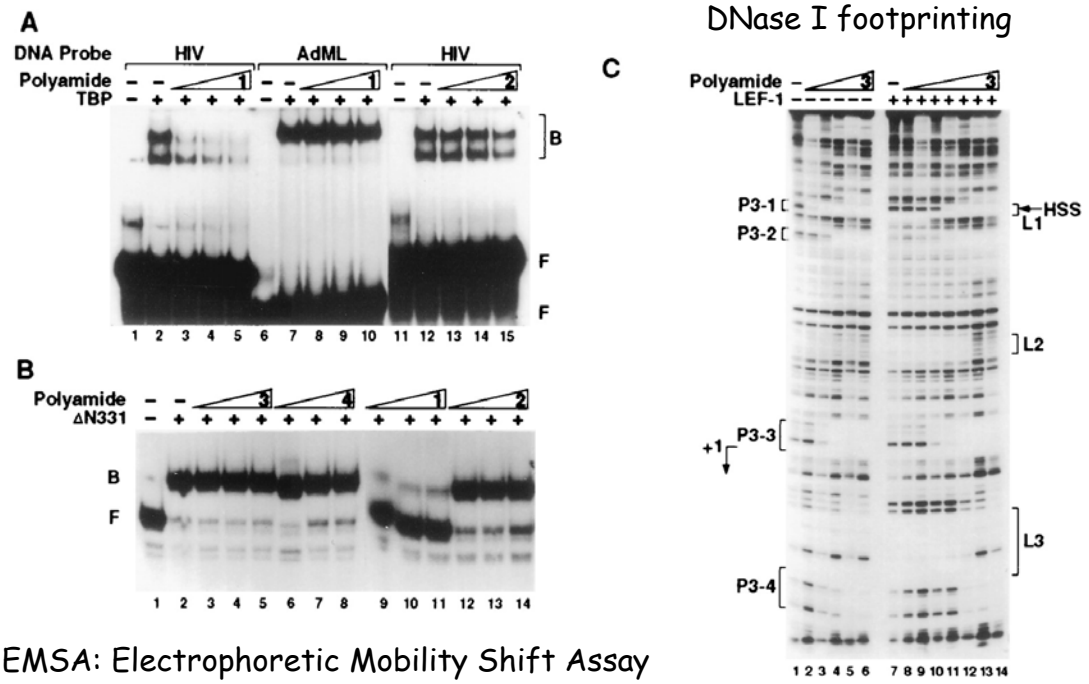


Fig. 1. Polyamide and transcription factor binding sites. (*A*) Schematic of the HIV-1 enhancer and promoter (nucleotide positions 199 to +1) showing binding sites for polyamides **1** and **3** and the transcription factors upstream stimulatory factor, Ets-1 (39, 40), LEF-1, NF-kB, Sp1, and TFIID (TBP). For polyamide binding models: shaded and unshaded circles, Im and Py rings, respectively. Binding models and measured dissociation constants are shown. Mismatches are highlighted.

Inhibition of transcription factor binding to the HIV-1 promoter and enhancer

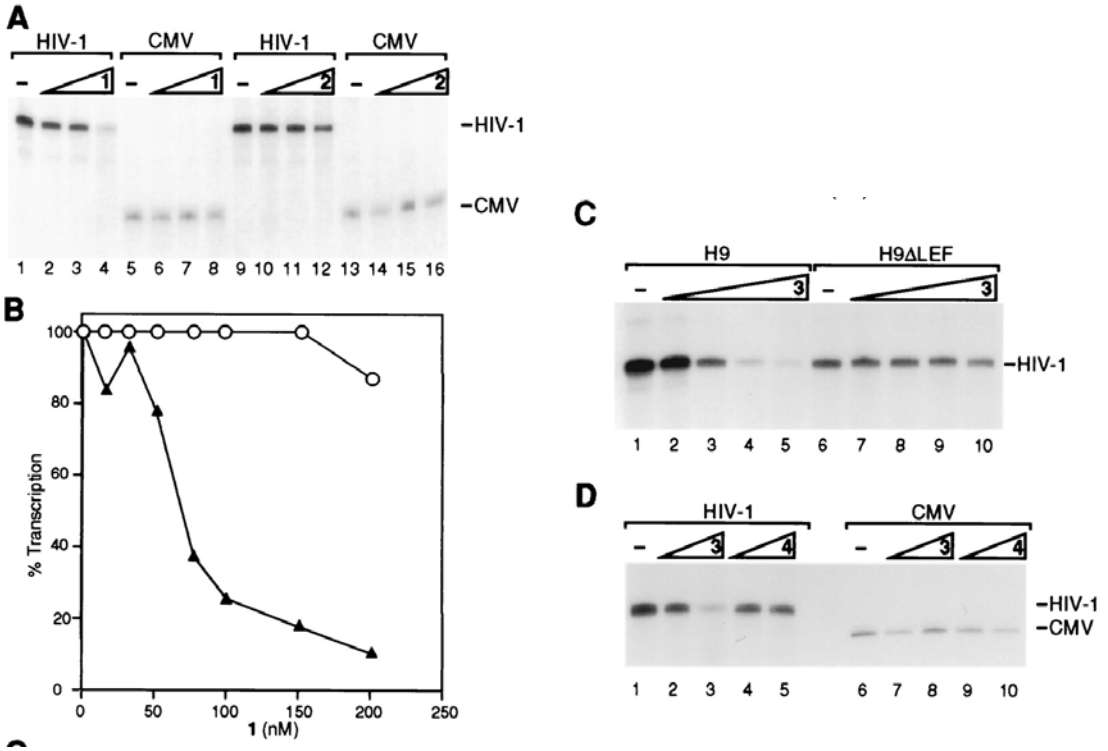


EMSA: Electrophoretic Mobility Shift Assay

DNase I footprinting

Dickinson *et al.* 1998 *PNAS* 95(22) 12890–12895

Polyamide inhibition of HIV-1 transcription *in vitro*

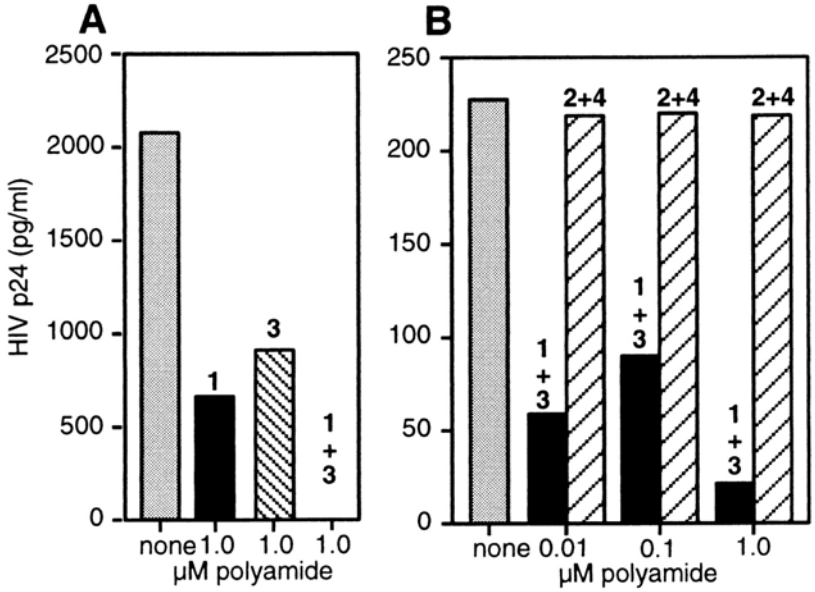


Dickinson *et al.* 1998 *PNAS* 95(22) 12890–12895

Polyamide inhibition of HIV-1 replication (in vivo)

In cell culture, polyamides nearly eliminate HIV-1 replication

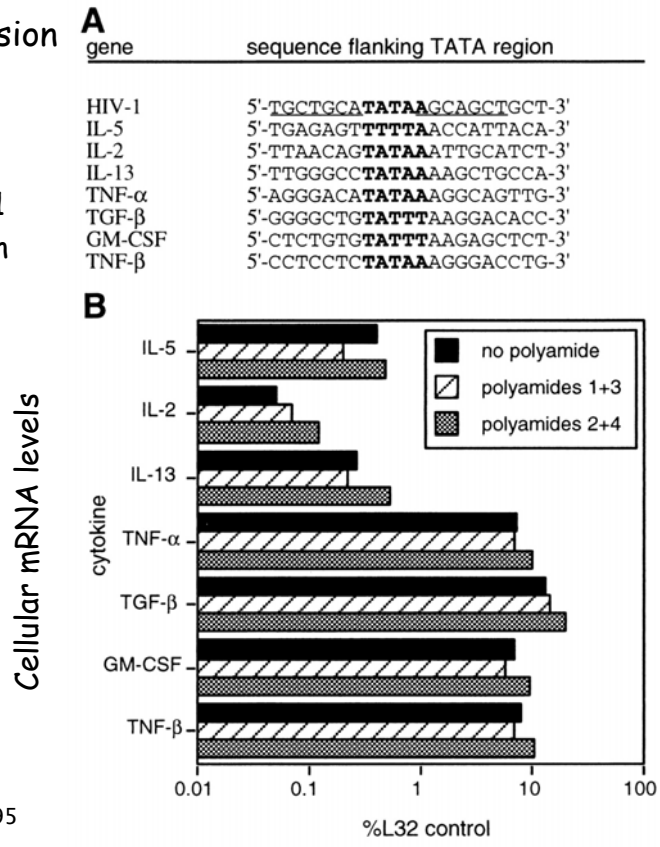
Only polyamides 1+3 inhibit HIV-1 replication



Dickinson *et al.* 1998 *PNAS* 95(22) 12890–12895

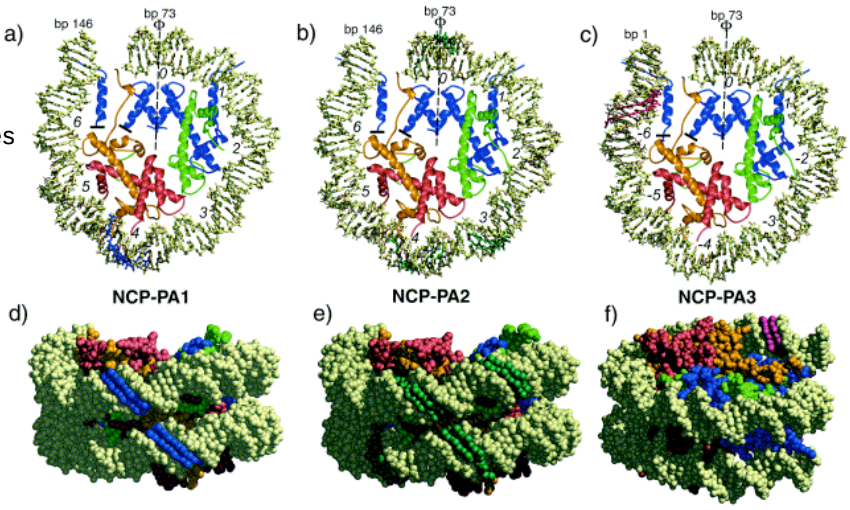
In vivo control of gene expression

Polyamides directed towards HIV-1 promoters do not inhibit expression of genes with different promoter sequences -- specificity

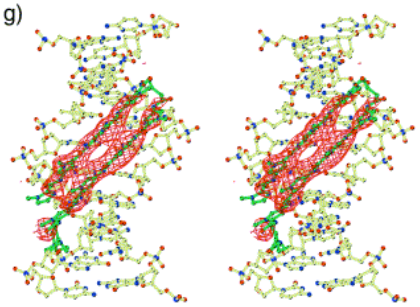


Dickinson *et al.* 1998 *PNAS* 95(22) 12890–12895

Crystal Structures of Nucleosome Core Particles in Complex with Minor Groove DNA-binding Ligands



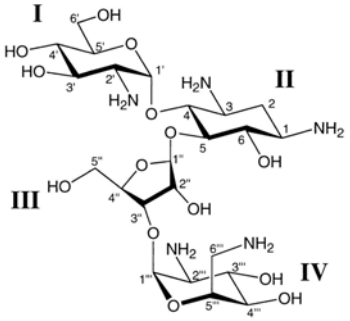
In vitro efficacy on model oligonucleotides is fine, but in vivo promoters are sequestered in chromosomes. How well do polyamides binding their target sites in nucleosomes?



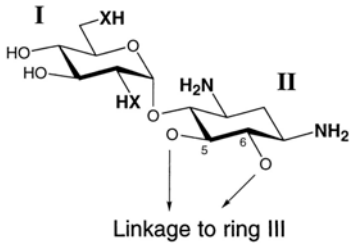
Suto *et al*. 2003
JMB 326(2):371-80

Aminoglycoside antibiotics

- E.g., Paromomycin, neomycin, gentamycin
- Bind to the ribosomal A-site (aminocyl-tRNA binding site) and inhibit protein synthesis by cause mis-translation
- Selective for bacterial ribosomes
- But, extended use leads to ototoxicity
- Post-transcriptional modifications of the A-site have led to clinical resistance



Paromomycin

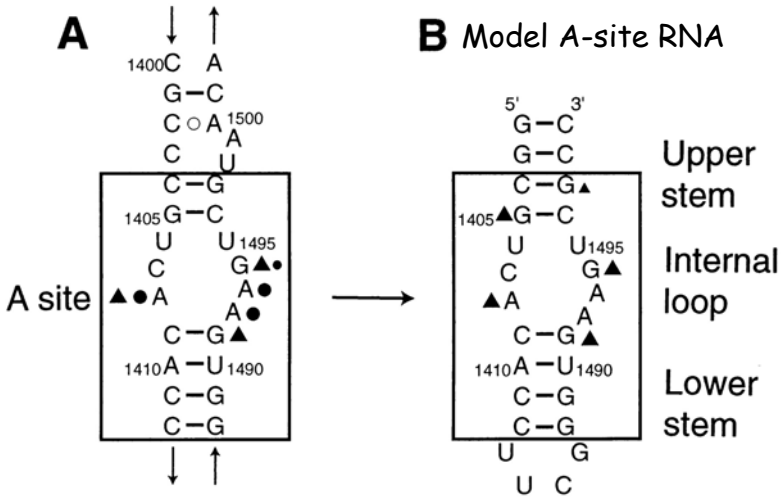


Conserved in A-site binders
X=NH or O

Unique arrangement of features and H-bonding potential allow aminoglycosides to target RNA w/ defined tertiary structures

Aminoglycoside-RNA interactions

E. coli A-site ribosomal RNA



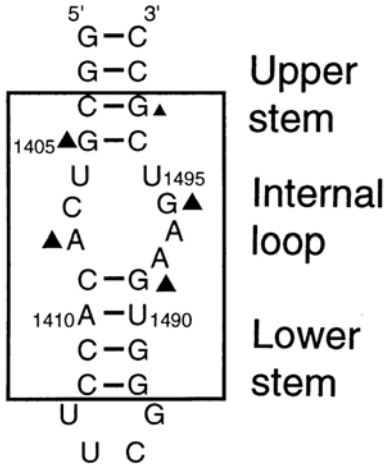
● Nucleotides protected from chemical probes by tRNA binding to ribosome

▲ Nucleotides protected by aminoglycoside binding to ribosome

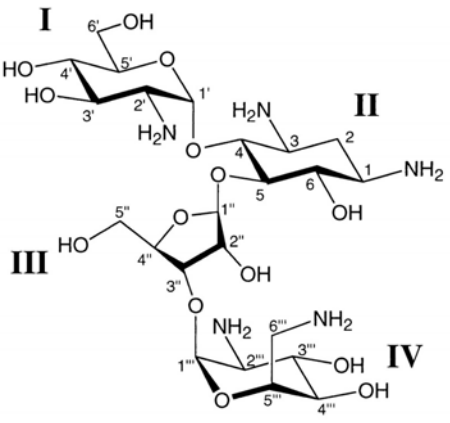
Because the protection pattern is similar in the model and intact A-site it serves as a tractable target for in vitro studies (NMR)

Fourmy *et al.* 1996 *Science* **274** (5291), 1367-1371

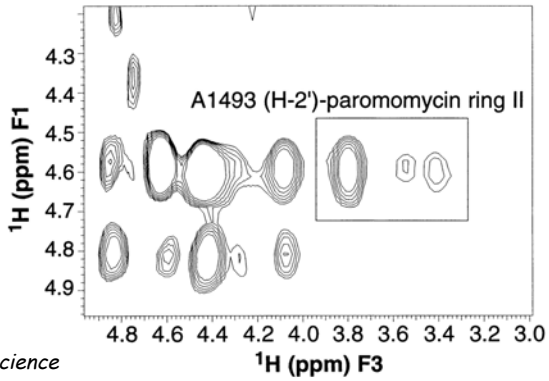
NMR of aminoglycoside antibiotic bound to mini A-site RNA



Intermolecular NOEs

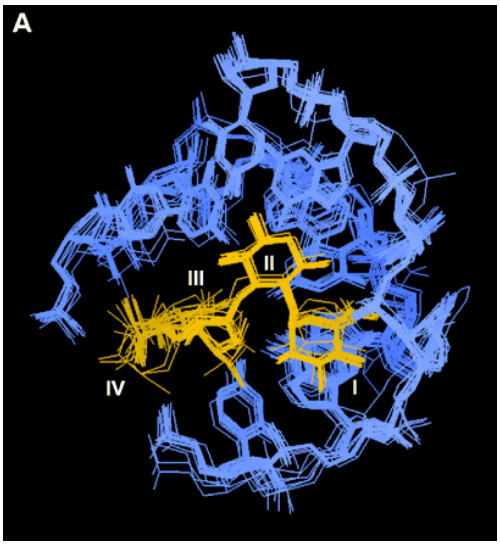


Paromomycin

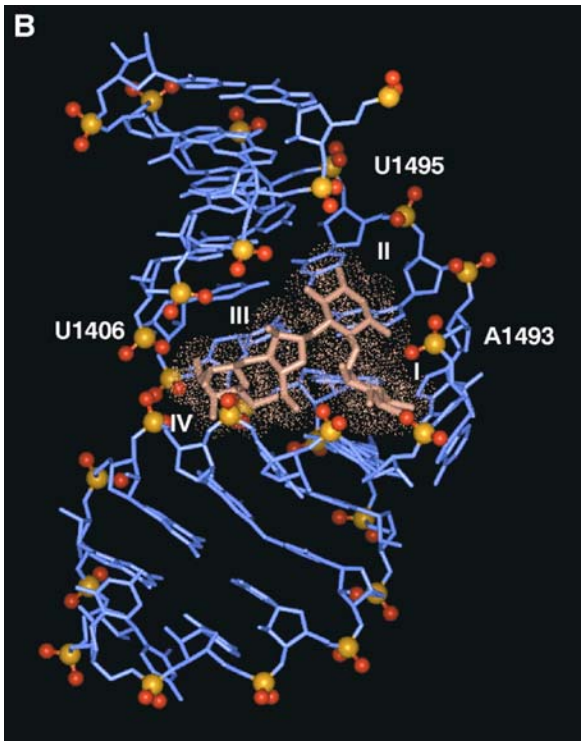


Fourmy *et al.* 1996 *Science*

NMR structure of A site-puromycin complex



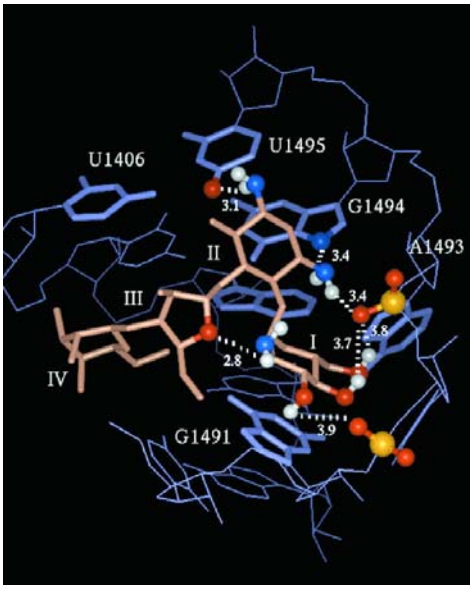
Ensemble: 20 Structures, RMSD 0.7 Å in core, 1.5 Å overall



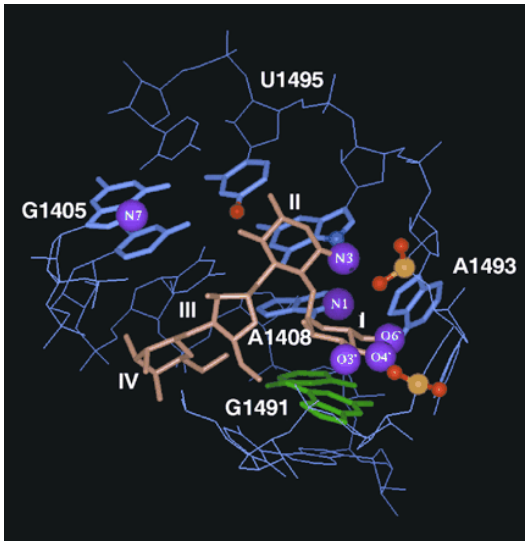
Representative structure

Fourmy *et al.* 1996 *Science*

Puromycin-RNA interactions

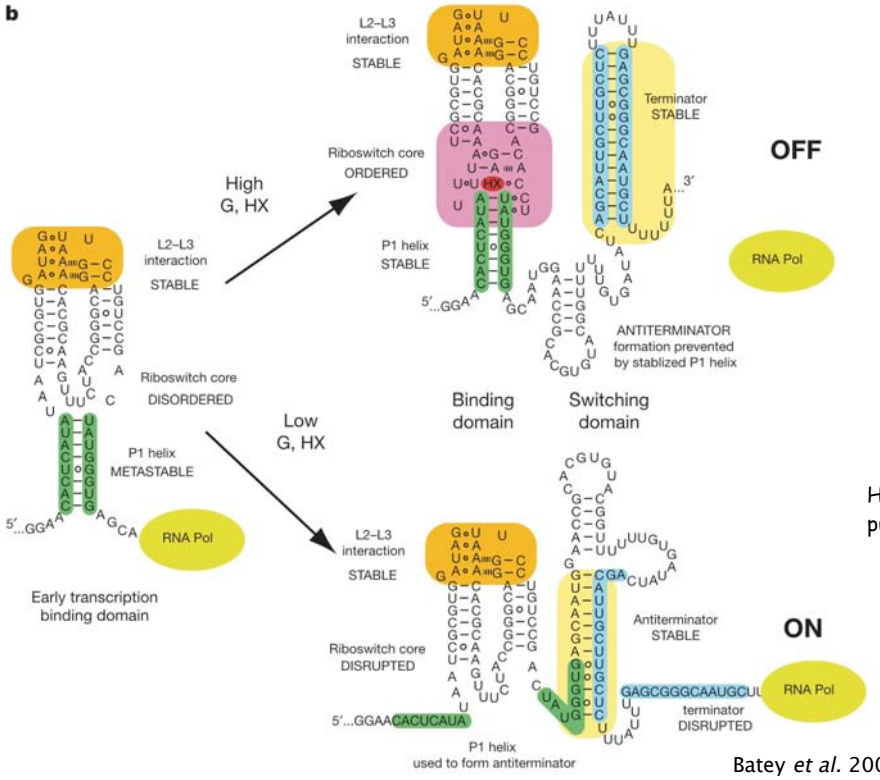


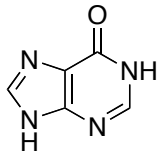
Specific contacts made between rings I and II of paromomycin and A-site RNA. The RNA is in blue, paromomycin is tan, and the view is into the major groove of the RNA core. The U1406–U1495, A1408–A1493 base pairs, as well as G1494 and G1491, are highlighted in the structure. Possible hydrogen bonding contacts are indicated by dashed lines.



Sites of covalent modifications to A-site rRNA and antibiotic that lead to aminoglycoside resistance. The RNA is blue and paromomycin is tan. G1405, U1406–U1495, A1408–A1493, and G1494 are highlighted in blue, and chemical groups in the rRNA that are involved in specific contacts are shown explicitly.

Gene repression by a ligand-responsive riboswitch



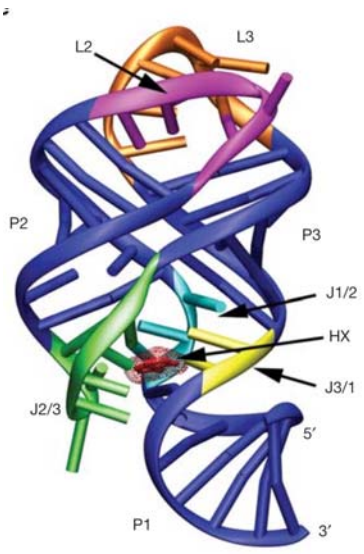


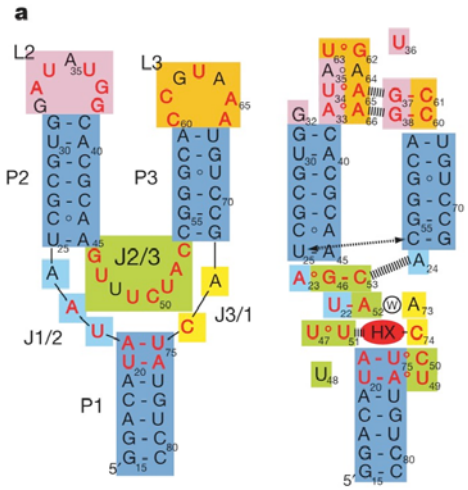
Hypoxanthine: Metabolite from purine salvage pathway

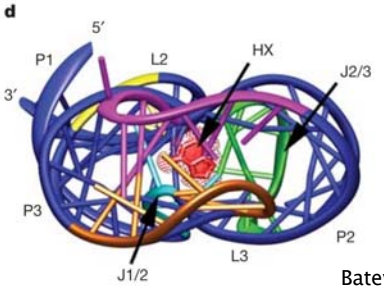
Batey *et al.* 2004 *Nature* **32**(7015):411–5

Crystal structure of the guanine riboswitch–hypoxanthine complex

Induced fit recognition stabilizes RNA 3° structure

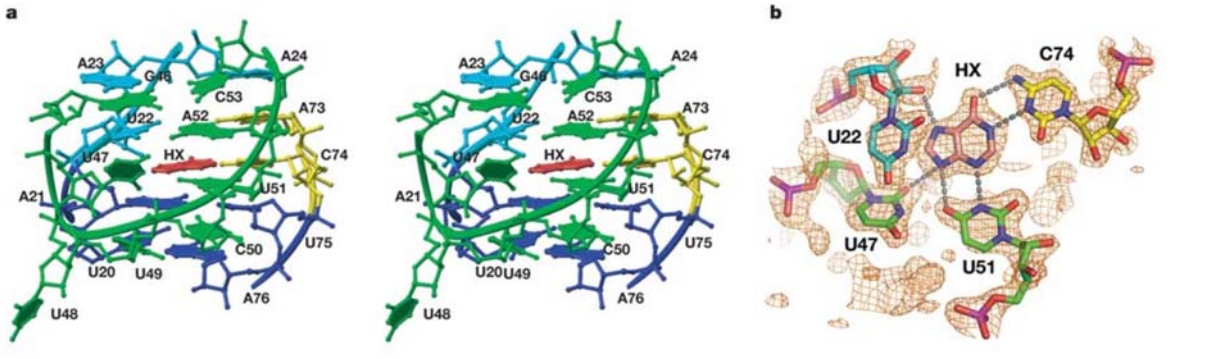




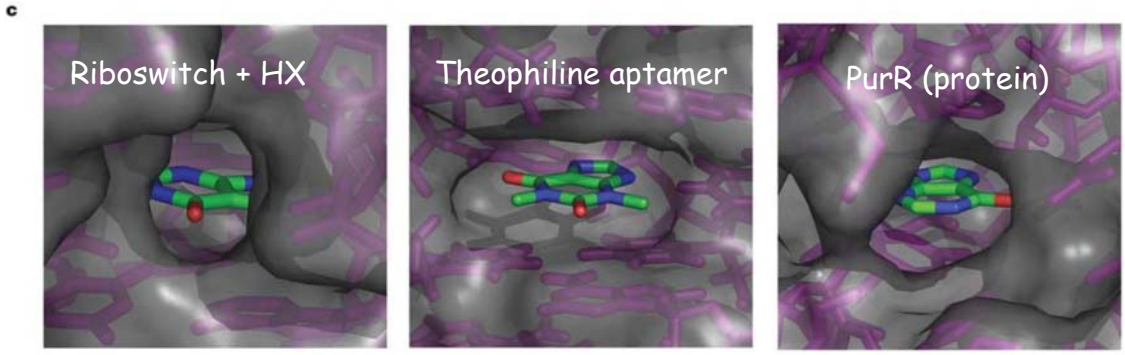


Batey *et al.* 2004 *Nature* **32**(7015):411–5

Recognition of hypoxanthine (HX) by the guanine-binding domain



Batey *et al.* 2004 *Nature* **32**(7015):411-5



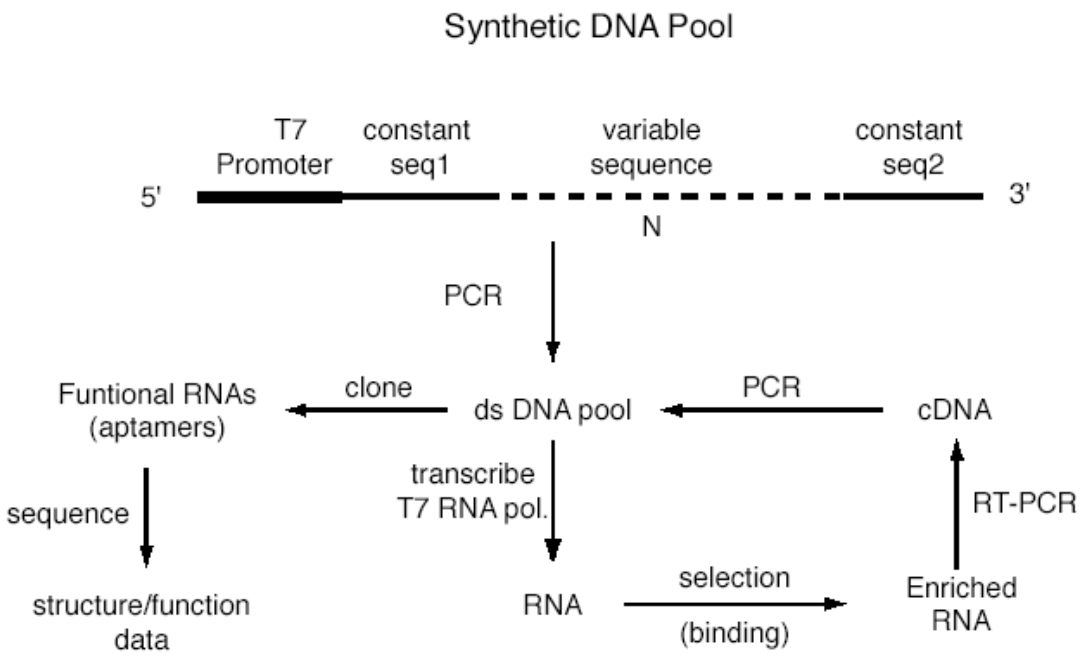
SELEX

- Aptamers
  - Tools
  - Diagnostics
  - Therapeutics
- Rybozymes
  - Gene therapy
- Evolution
  - RNA world

- Versatile
- Any (?) ligand or function can be selected

- What can be constructed with only 4 building blocks?

SELEX: Systematic Evolution of Ligands by EXponential amplification



1. Starting pool of randomized DNA oligos

With ~ 1-10 ug DNA, can sample ~ 10^{14} - 10^{15} sequences
Starting pool contains the intire population of variants
objective is to amplify the winners

Combinatorial problem (4^N molecules needed to sample all possible sequences)

2. Separate the "winners" (functional/competent members of the pool)

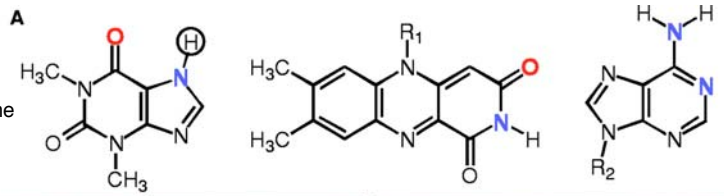
Requires suitable selection scheme
Stringency of selection can be modified

3. Amplify the winners and repeat selection or identify (sequence)

Sequencing is facilitated by constant regions
Usually performed after several iterations (rounds)

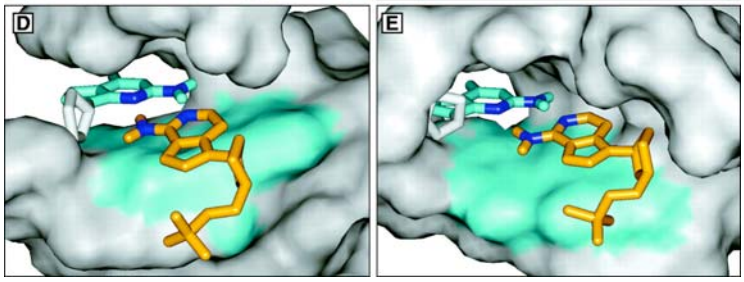
Aptamers to Planar Aromatics (nucleotides)

Theophylline/Caffeine



Theophylline/RNA

FMN/RNA

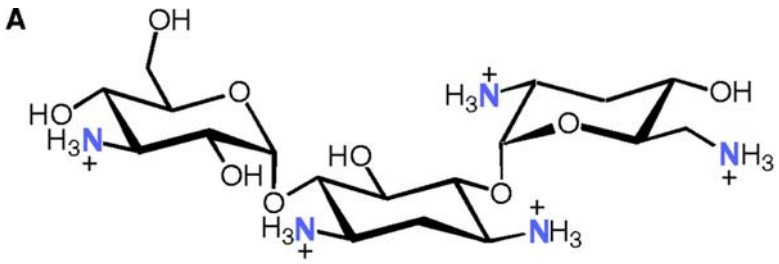


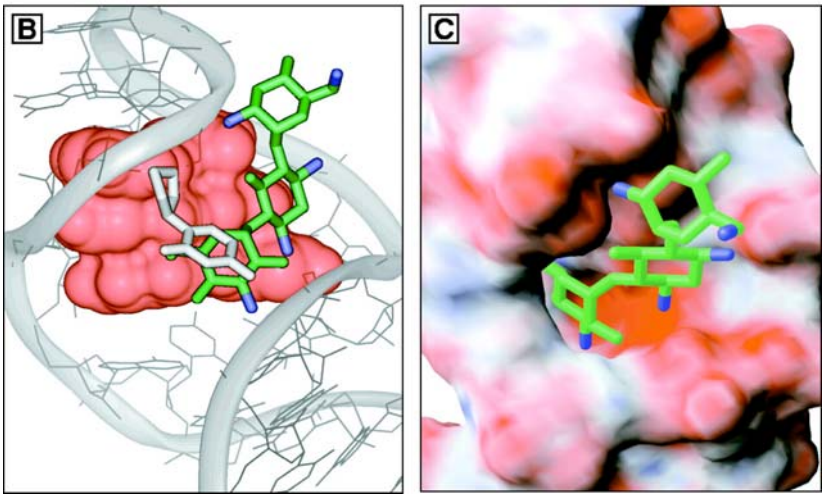
ATP/DNA

ATP/RNA

Hermann and Patel,
Science 2000

Aminoglycoside aptamer

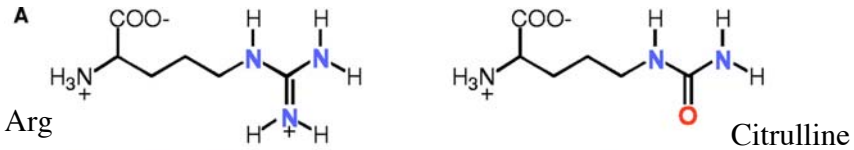




Aminoglycoside antibiotics - tobramycin
In C, red: negative electrostatic potential

Hermann and Patel,
Science 2000

Amino acid recognition aptamers

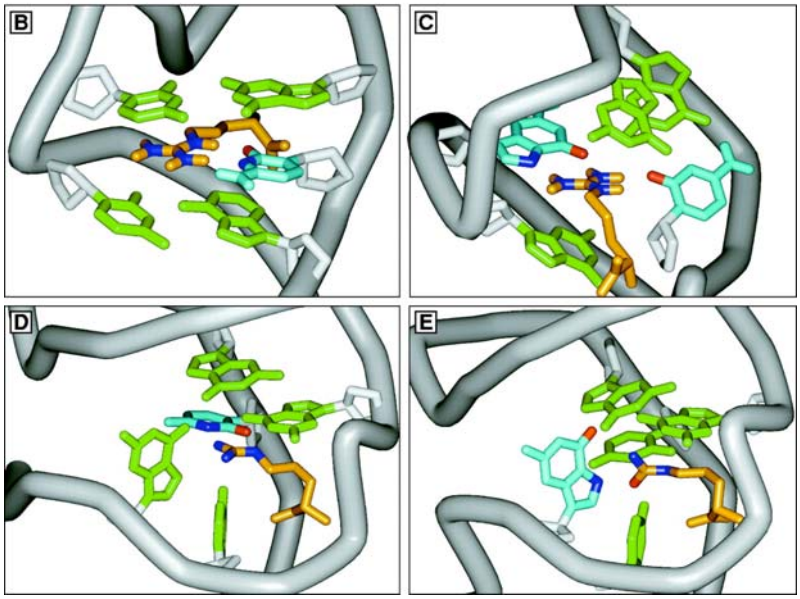


Arg

Citrulline

DNA

RNA



Hermann and Patel, *Science* 2000

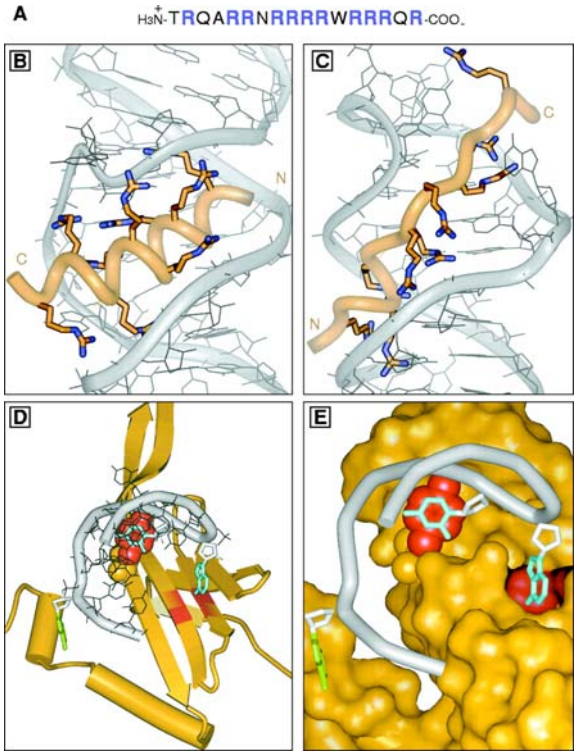
Peptide and protein recognition by aptamers

Arg-rich peptide from HIV REV protein; two aptamers

Groove interaction by helix (B) or an extended structure (C)

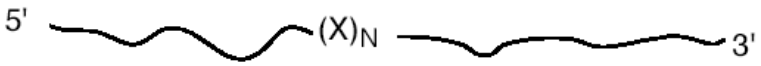
MS-2 Coat protein

Specificity achieved by binding of looped-out nucleotides to cavities and stacking with protein sidechains



SELEX -- A 4^N Combinatorial Problem

Consider an oligo with N randomized nucleotides



If the oligo has 1 randomized nt, it is only necessary to have 4 molecules to have a representative member of each possible sequence (X = A, C, G, T).

Assuming an average MW per nt of ~ 300 Da, one would need:
(Ignoring 5' and 3' constant regions)

$$1 * 4 * 300 \text{ g mol}^{-1} / 6.02 \times 10^{23} \text{ molecules mol}^{-1} = 2 \times 10^{-21} \text{ g}$$

For N = 2, we would need 16 representative molecules:
AA, AC, AG, AT, CA, CC, CG, CT, GA, GC, GG, GT, TA, TC, TG, TT

N	# molecules	required sample
2	16	1e-20 g
3	64	1e-19 g
4	256	5e-19 g
10	~1e6	5e-15 g
20	~1e12	1e-8 g
40 (~ ATP aptamer)	~1e24	66 g
50	~1e30	24e9 g
70 (~ tRNA)	~1e42	3e22 g

Mass of the earth: 6e27 g. Moon: 7.4e25 g)

SELEX requires patience, persistence and faith?

If the pool is sufficiently varied and the selection works,
the number of molecules remaining after the first selection step will be
VERY small.

It is common for no detectable winners to be visible until after multiple rounds of selection/amplification

Aptamer diversity 1999

TABLE 1 Aptamers for small molecules

Target	Estimated K_D (μm)	Reference
Nucleotides and nucleobases		
ATP/adenosine	1	39
ATP/adenosine (DNA)	6	43
Guanosine	32	41
Guanine/xanthine	1.8	184
7-Methyl-GTP	~0.5	185
Theophylline	0.11	42
Amino acids		
Arginine	0.33	61
Citrulline	62	58
Valine	12,000	186
Tryptophan	18	187
Cofactors		
Cyanocobalamin	0.09	188
N-methylmesoporphyrin IX	~14	44
N-methylmesoporphyrin IV (DNA)	~0.5	189
Flavin	0.5	190
NAD	2.5	47
RMP-botin	2	141
Antibiotics		
Tobramycin (aminoglycoside)	0.0008	55
Neomycin (aminoglycoside)	0.1	191
Lividomycin (aminoglycoside)	<0.2	192
Kanamycin (aminoglycoside)	<0.2	192
Streptomycin (aminocyclitol)	~1	193
Viomycin (basic peptide)	12	194
Chloramphenicol (small, neutral)	2.1	195
Transition state analogs		
Diels-Alder reaction	3,500	196
Bridged biphenyl isomerization	542	143
Other		
Dopamine	2.8	197
Peptide (substance P)	0.19	29
Divalent metals	~1	198

[a]Where multiple selections have been performed, this table lists only the highest-affinity case. All aptamers are RNA except where noted ("DNA"). RMP-biotin, ribose-monophospate-biotin, a carboxylate-phosphate anhydride of biotin and ribose-5-phosphate. K_D values were estimated by several different methods of variable accuracy and precision. Consult references for details.

Wilson & Szostak 1999 *Ann Rev Biochem*