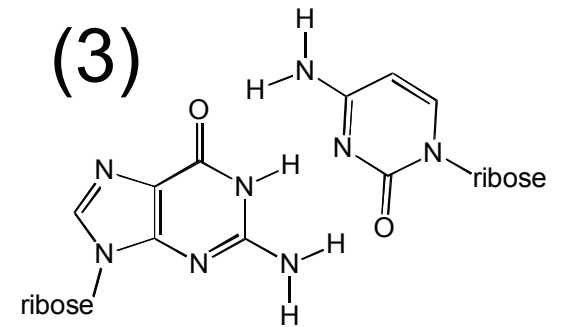
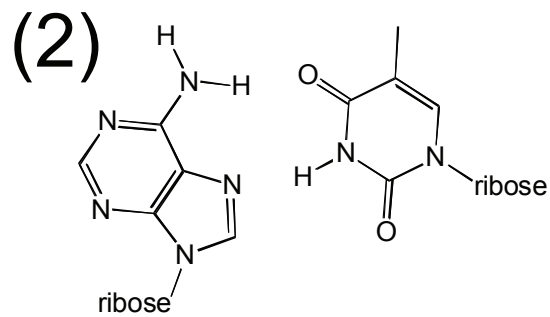
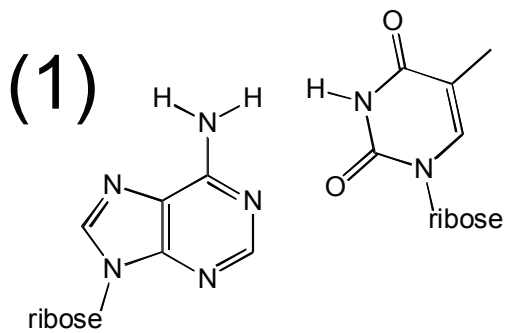
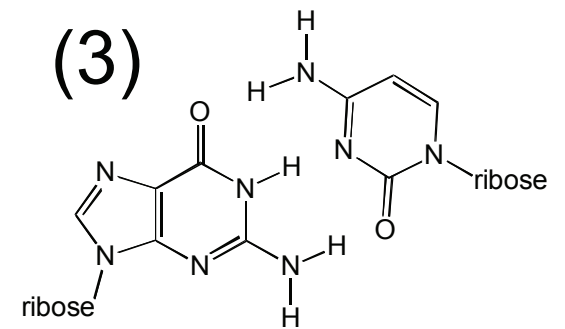
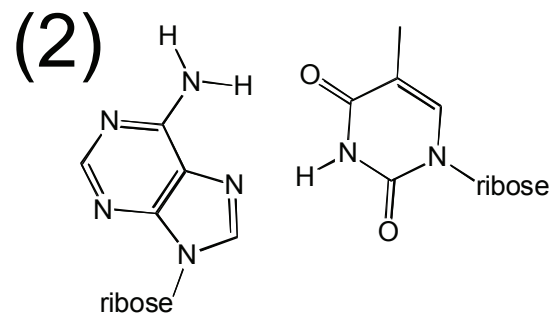
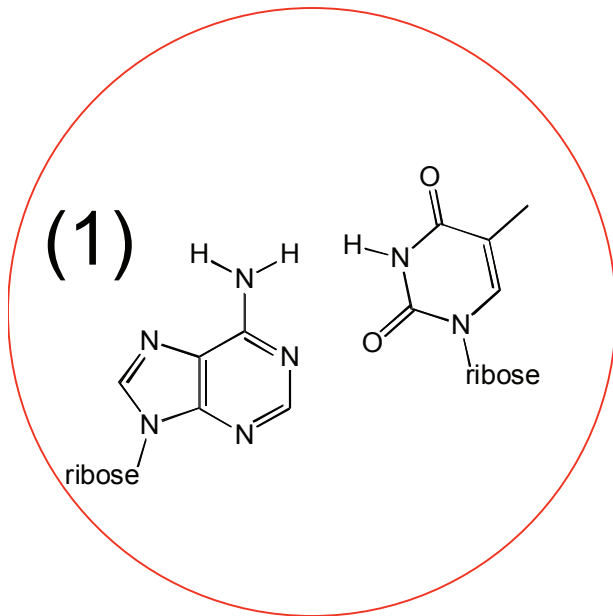


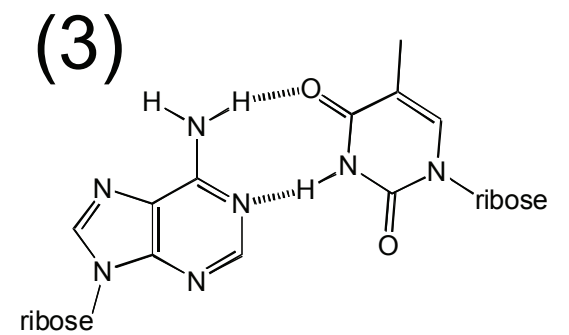
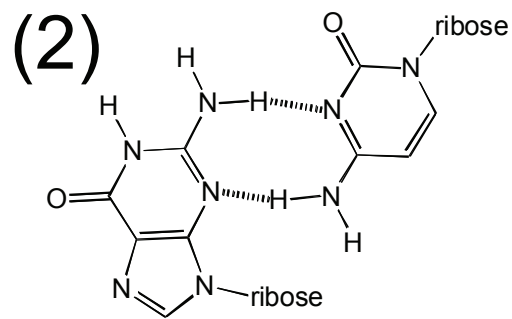
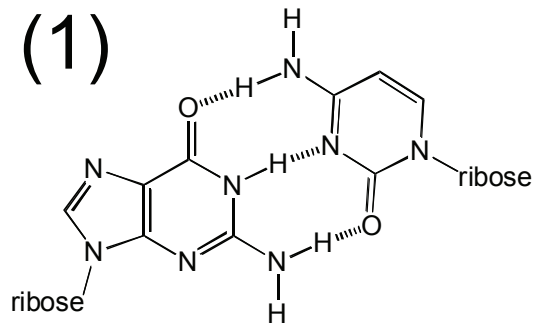
Which base pair has the weakest interaction?



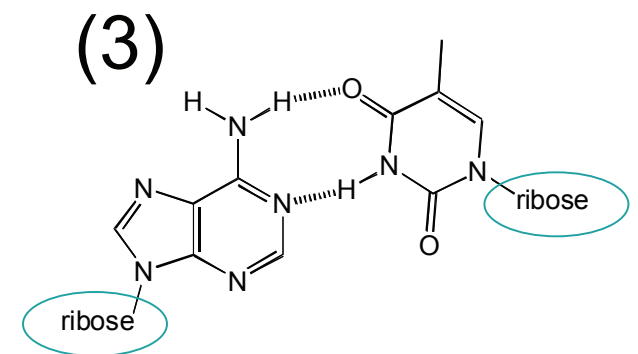
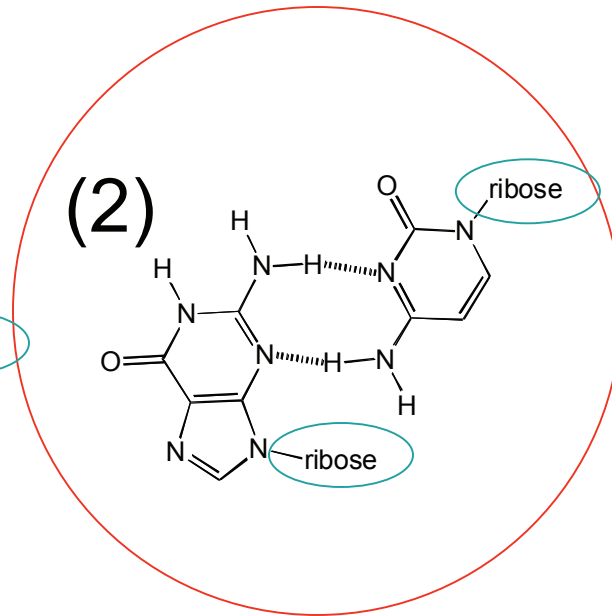
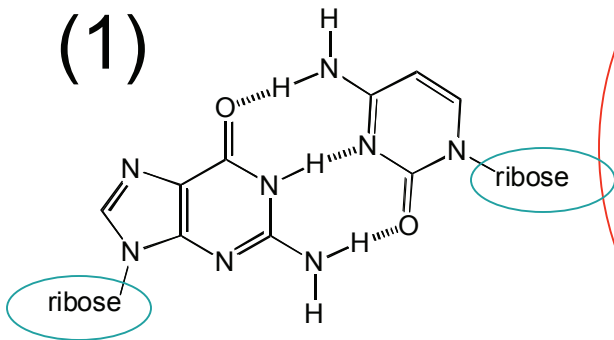
Which base pair has the
weakest interaction?



Which base pair is **not** Watson-Crick?



Which base pair is **not**
Watson-Crick?

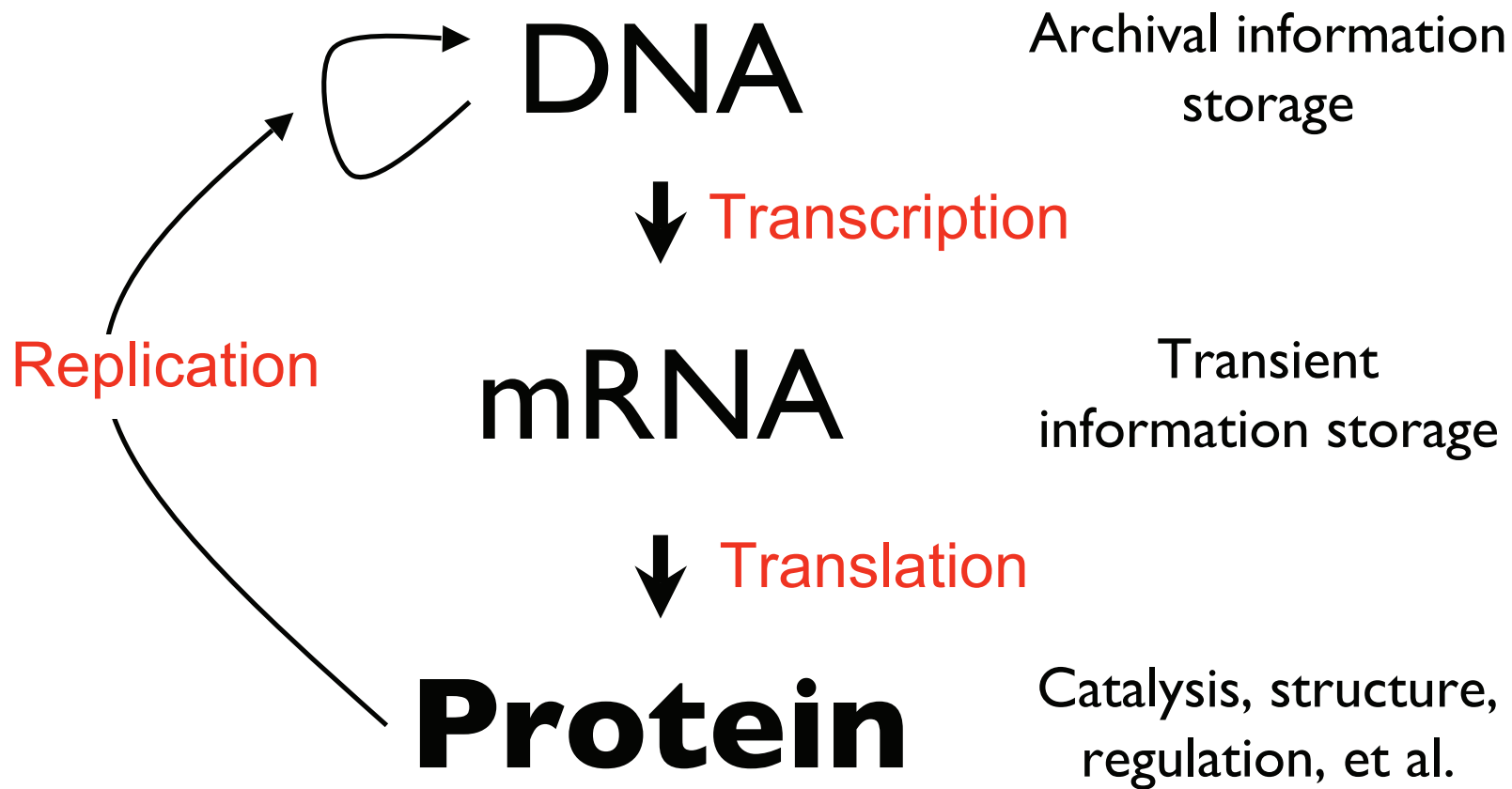


DNA, RNA, Proteins

revolutions at the turn of the century

opportunities for the 21st century

The Central Dogma



The Genetic Code

- Three bases (in a row) of DNA (RNA) encode a single amino acid
- An amino acid is covalently attached to a tRNA molecule
- That tRNA molecule contains a triplet (3 base) sequence that binds (Watson Crick) to the complementary sequence in the mRNA

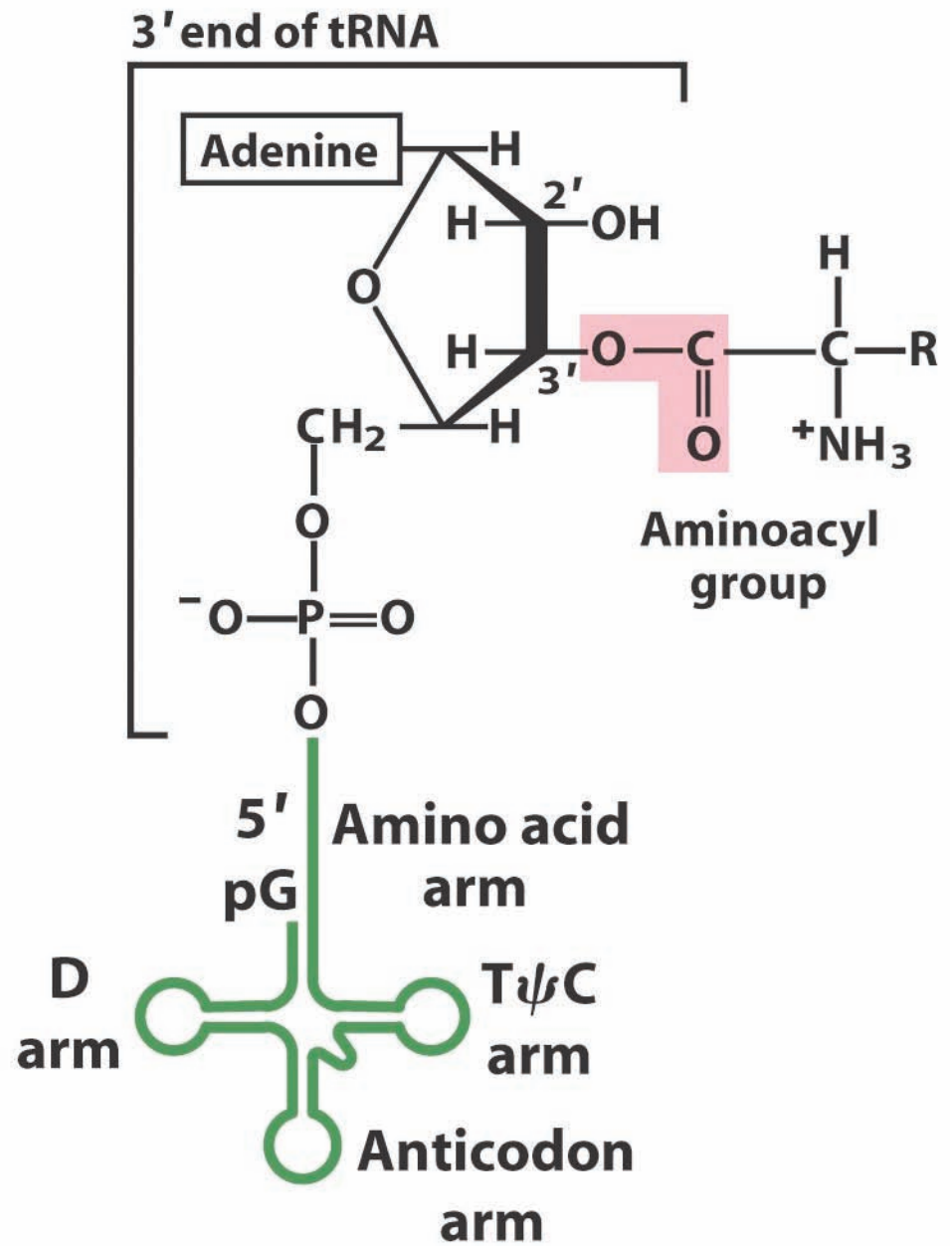
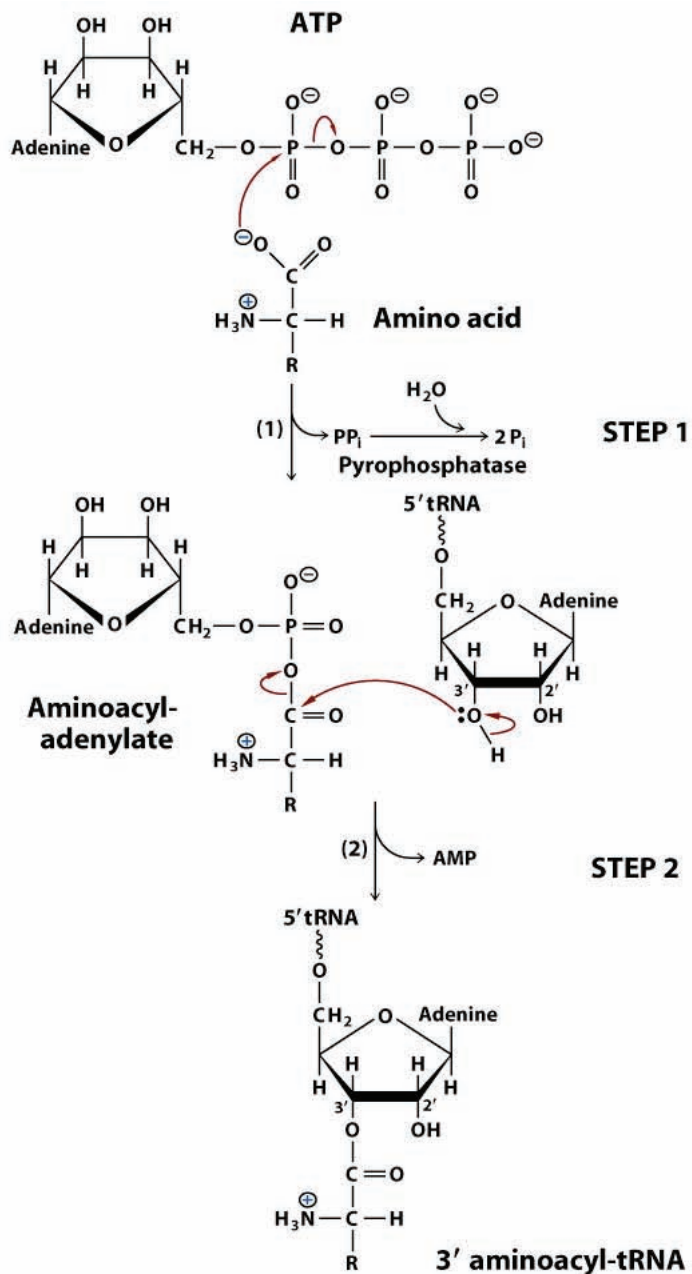
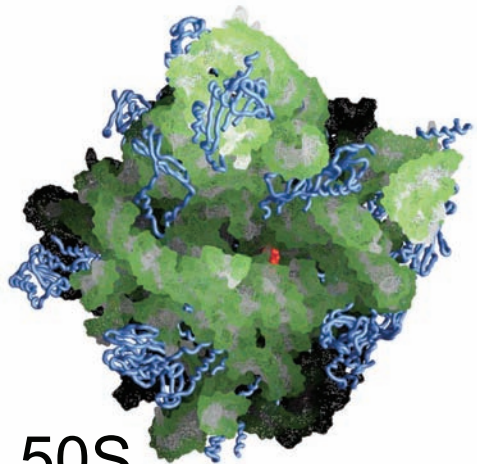
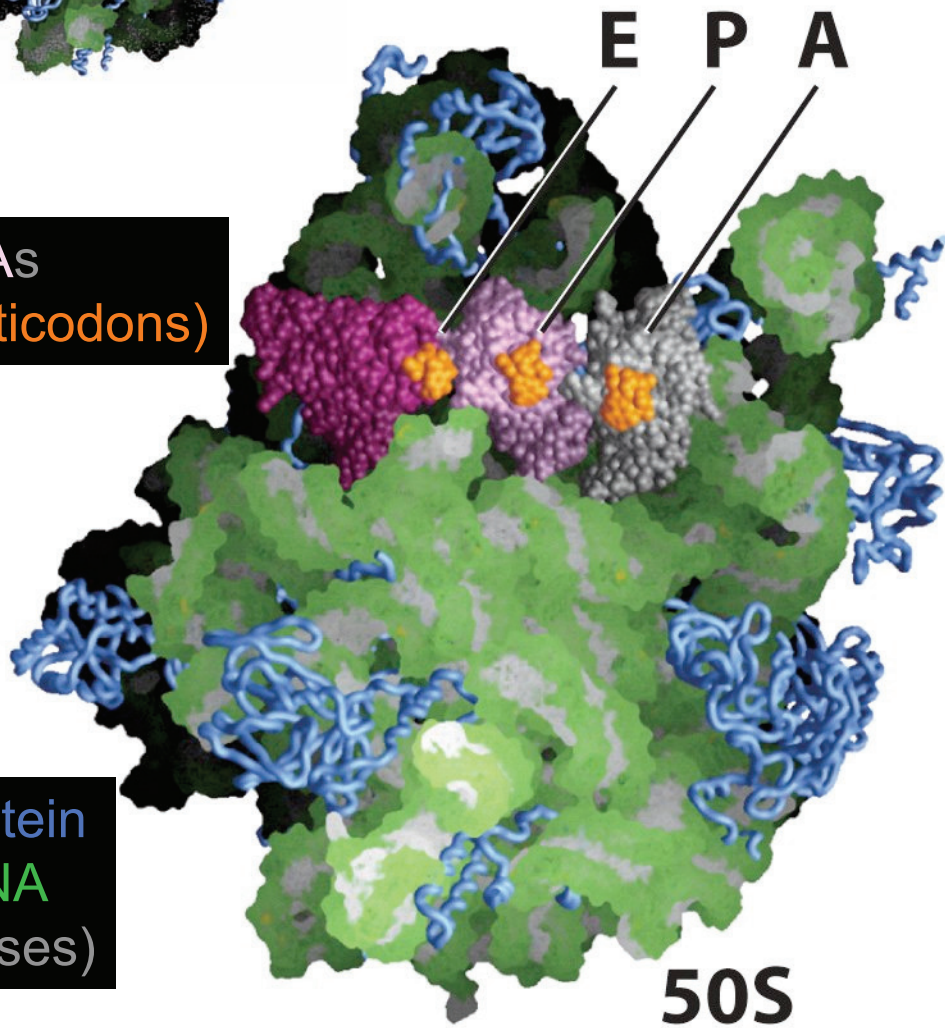


Figure 22-9 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.



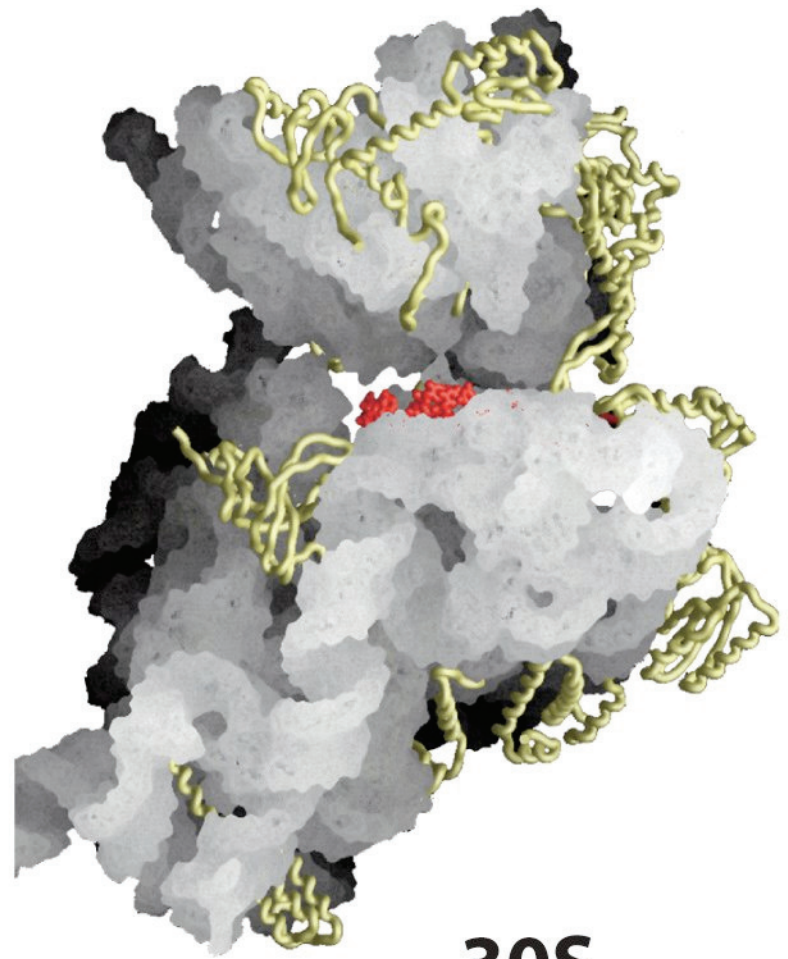
50S

tRNAs
(anticodons)



Protein
rRNA
(bases)

Protein
rRNA
mRNA codon



30S

First letter of codon (5' end)



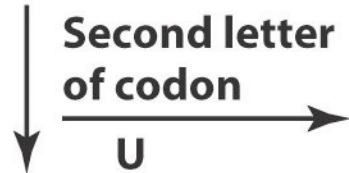
	U		C		A		G	
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys
U	UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop
	UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg
C	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
A	AUU	Ile	ACU	Thr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

Stop Codons

- UAA
- UAG
- UGA

Initiation
AUG

First letter of codon (5' end)



2 have 1 codon

Rest have multiple --

genetic code is degenerate

mostly due to changes in 3rd base

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

First letter of codon (5' end)

Second letter of codon

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

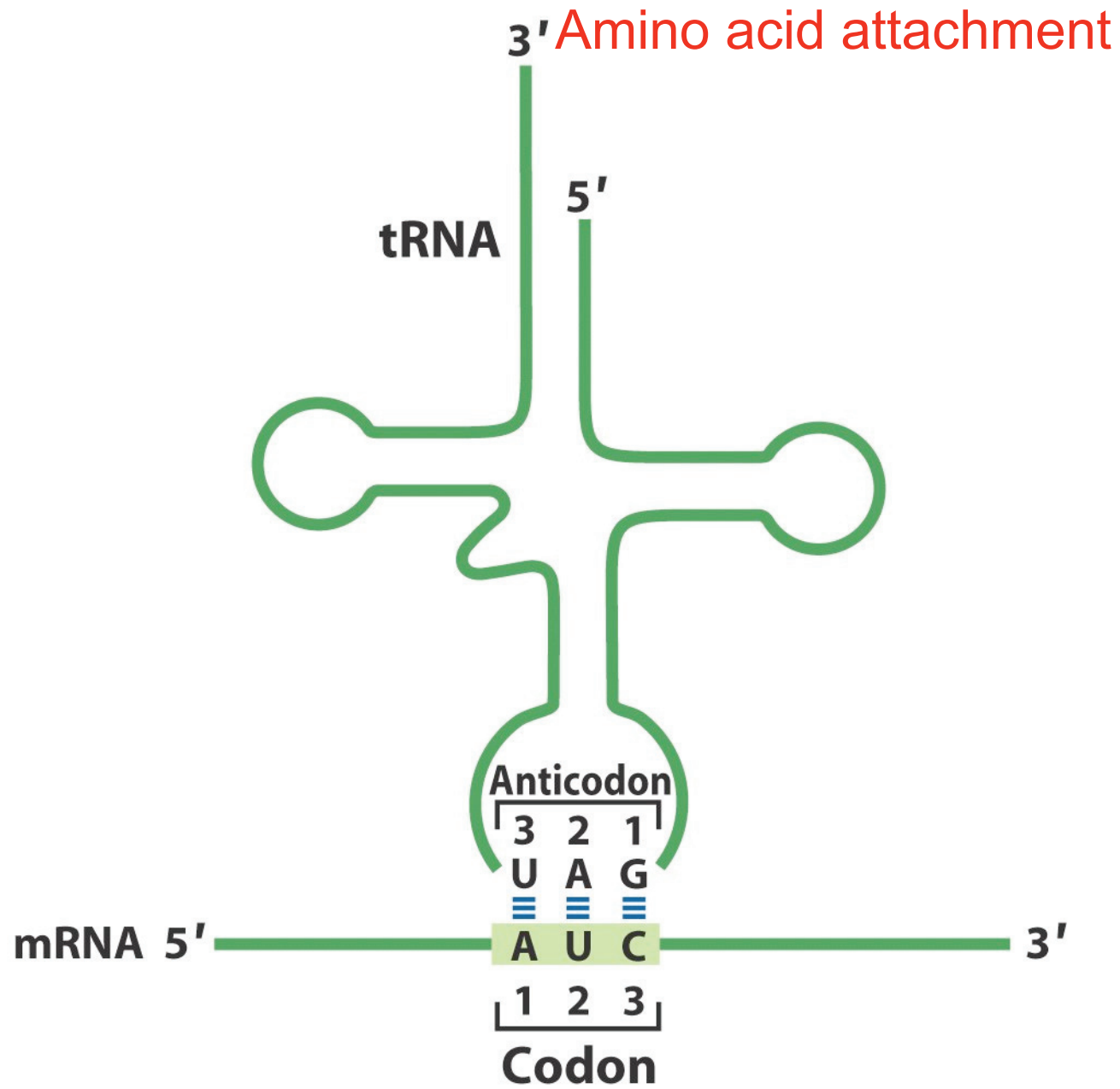
2 have 1 codon

Rest have multiple --

genetic code is

degenerate

mostly due to changes in 3rd base



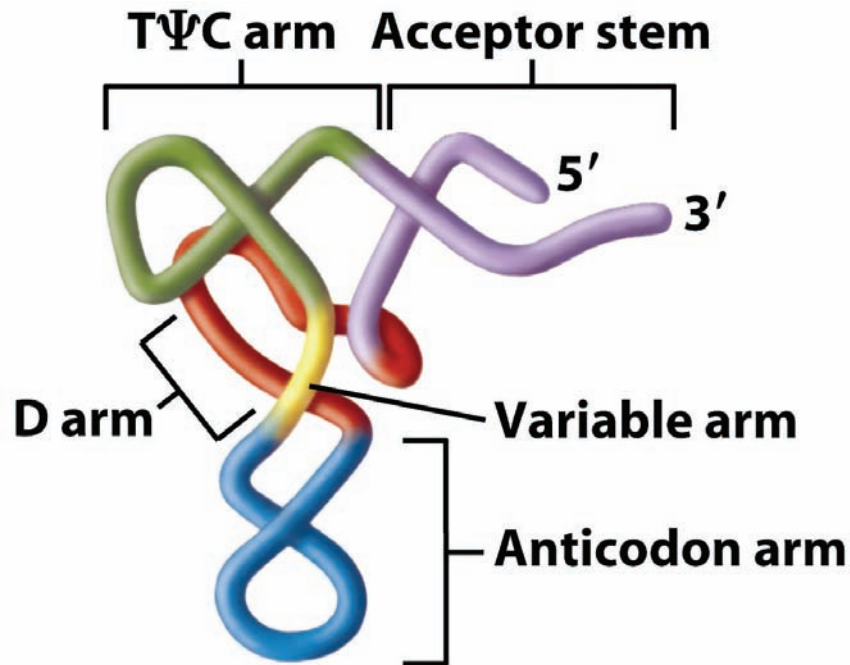


Figure 22-5 Principle:
© 2006 Pearson Prent

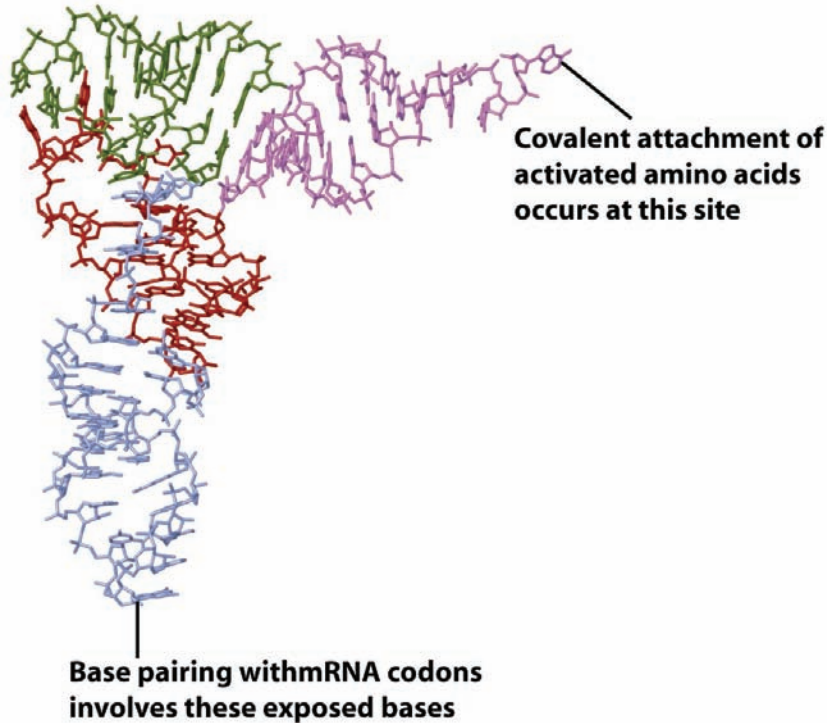
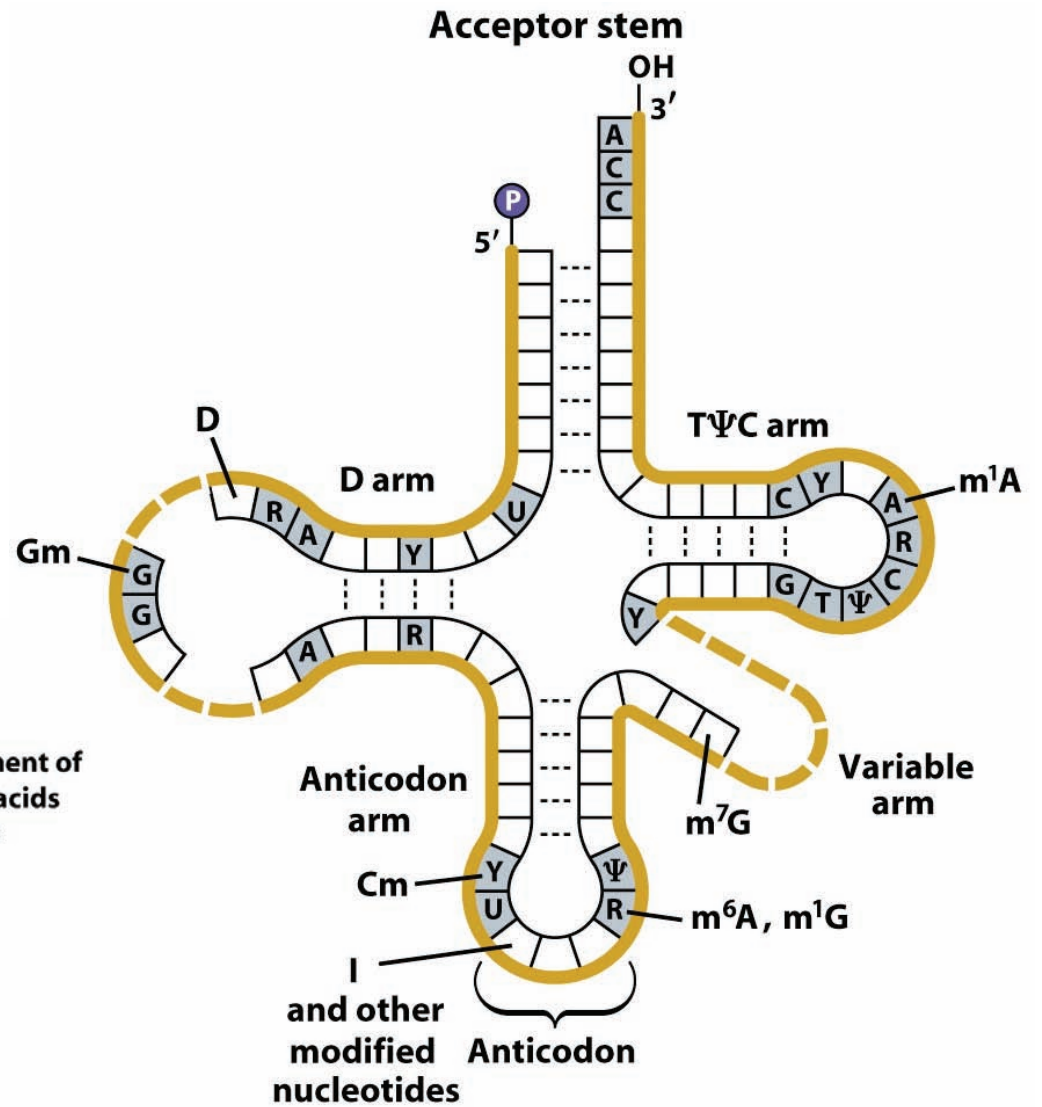


Figure 22-6 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.



22-4 Principles of Biochemistry, 4/e
Pearson Prentice Hall, Inc.

Many mutations are **conservative**

TABLE 22.1 Predicted base pairing between the 5' (wobble) position of the anticodon and the 3' position of the codon

<i>Nucleotide at 5' (wobble) position of anticodon</i>	<i>Nucleotide at 3' position of codon</i>
C	G
A	U
U	A or G
G	U or C
I ^a	U, A, or C

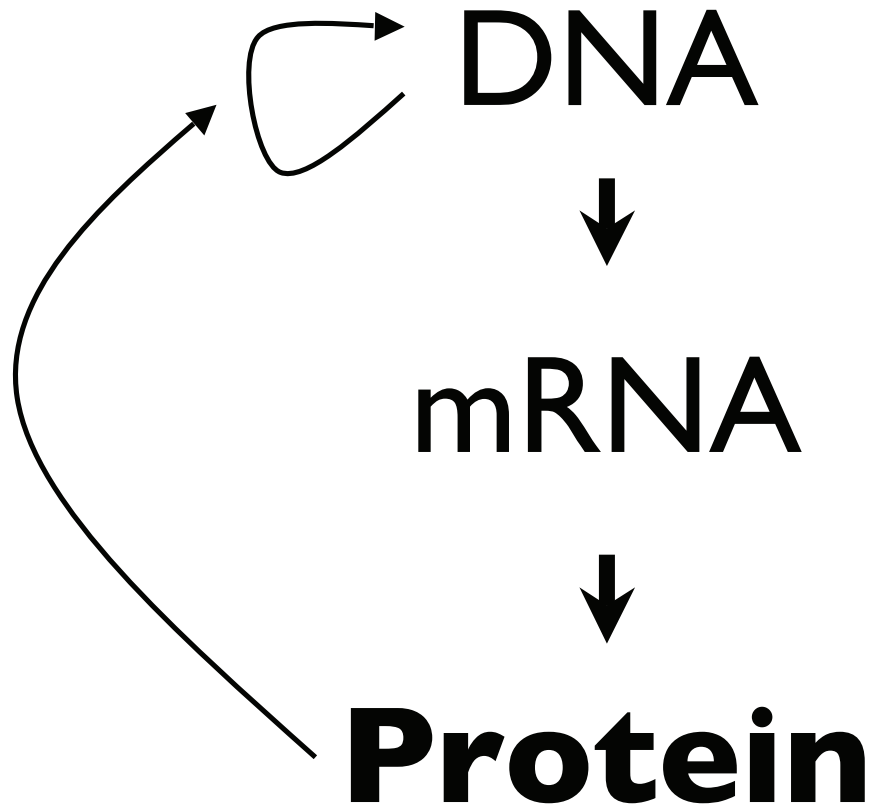
^aI = Inosinate

Table 22-1 Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

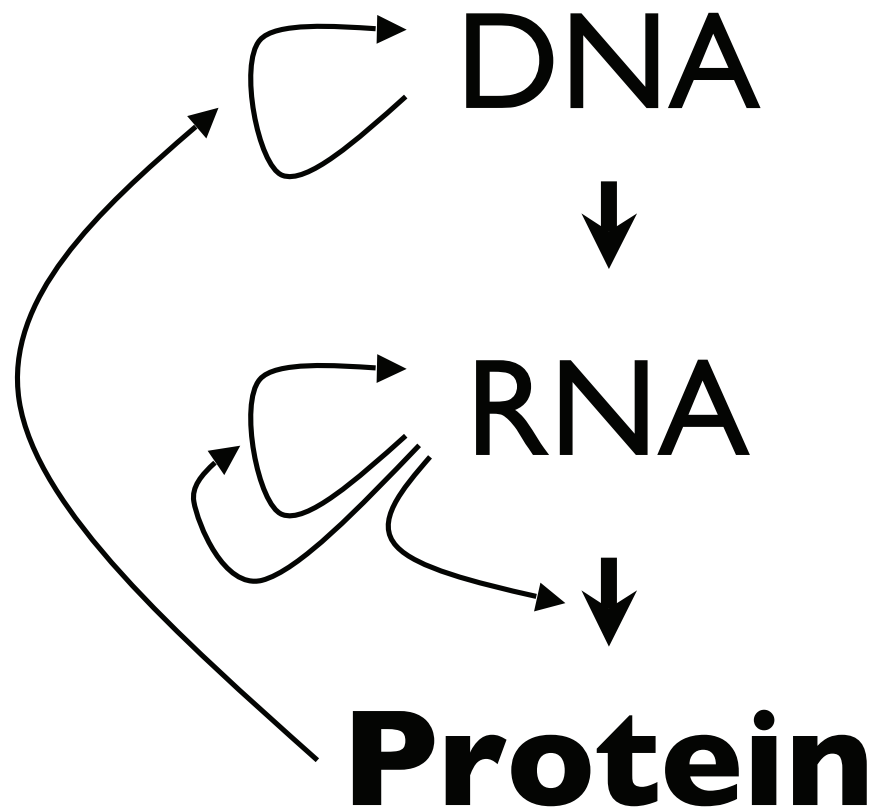
First letter of codon (5' end)

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

Chicken & Egg?



RNA can do everything



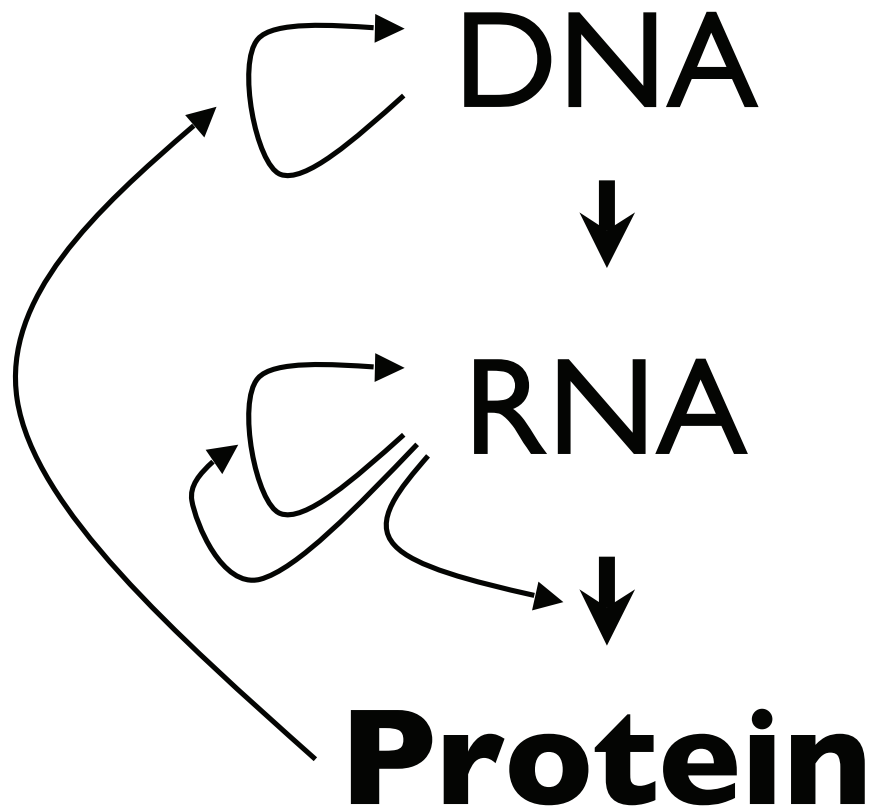
Archival information
storage

Transient
information storage
Catalysis!

1980-2000

Catalysis, structure,
regulation, et al.

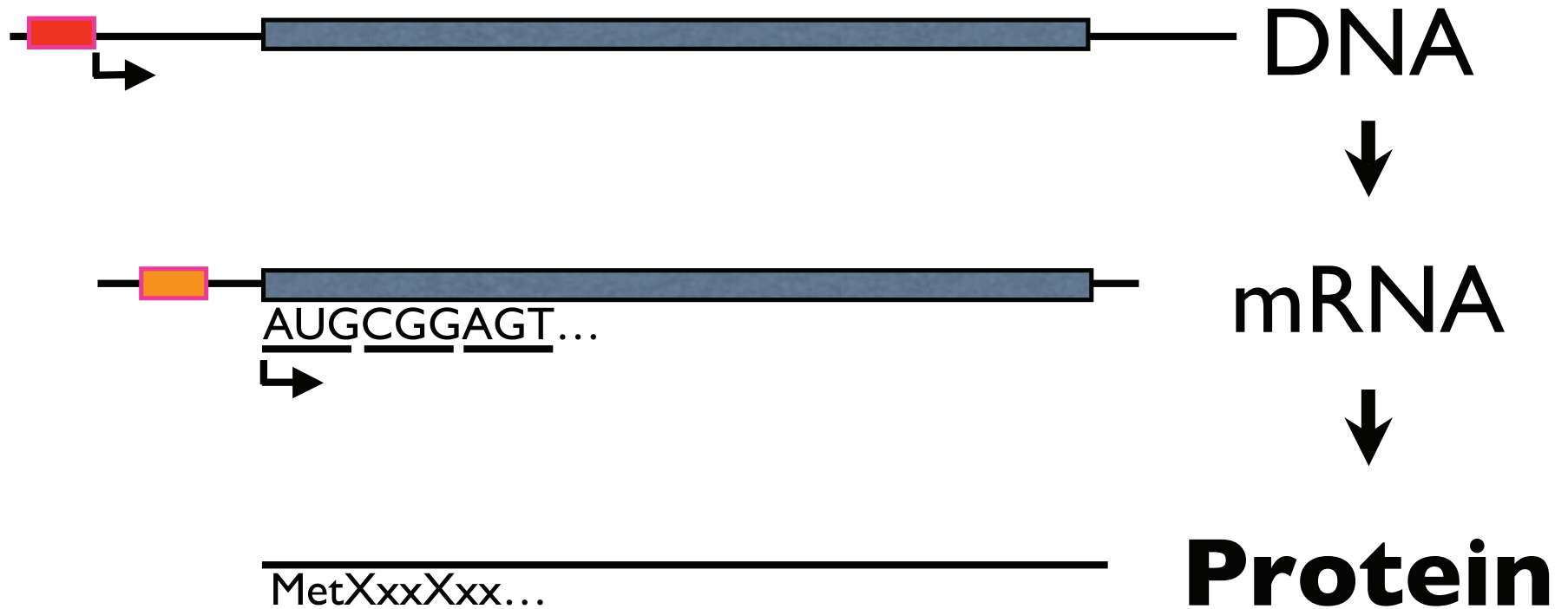
RNA - primordial molecule



So RNA was probably the first, primitive do-everything biomolecule.

But **proteins** came along to supplant everything and make the world, evolutionarily, what it is today. All hail the protein!

Seems simple...



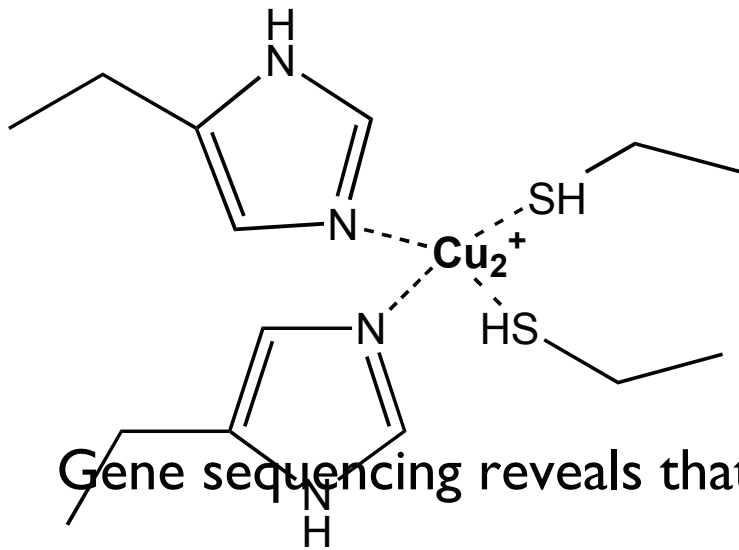
Simple?

A personal aside

PhD Dissertation

1984

ENDOR and EPR evidence for a model in which two absolutely conserved Cys residues coordinate copper



1983

Gene sequencing reveals that wheat cytochrome oxidase has one of the tw

Cu_A - cytochrome c oxidase

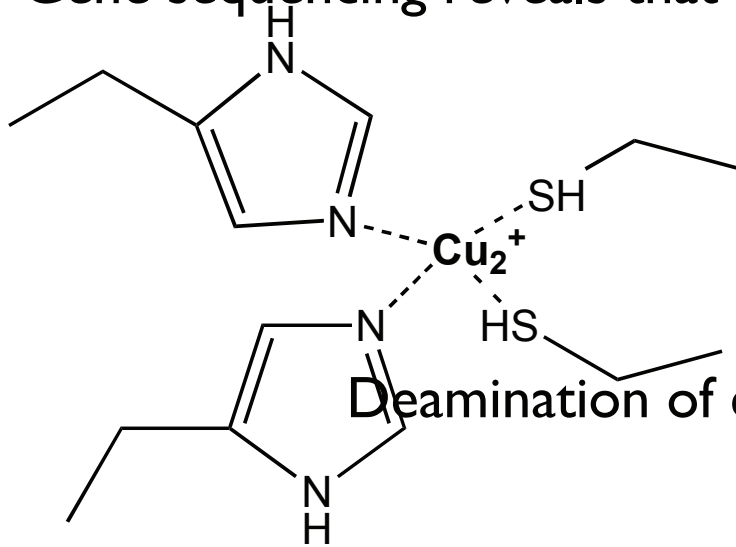
Aaaarrggh!

Simple?

A personal aside

1984

Gene sequencing reveals that wheat cytochrome oxidase has one of the two



1985-1995

Deamination of cytosine converts C to U, changing Arg codon

The RNA is edited!

Cu_A - cytochrome c oxidase

A sequence-specific and precisely controlled process

Simple?

A personal aside

1984

Trypanosomes completely lack subunit III

12 subunits

Subunits I, II, III
- major subunits

Subunits IV-XII
- minor subunits

1985

Subunits are post-transcriptionally added to the pre-mRNA, generating

Two thirds of the final gene is added U's!

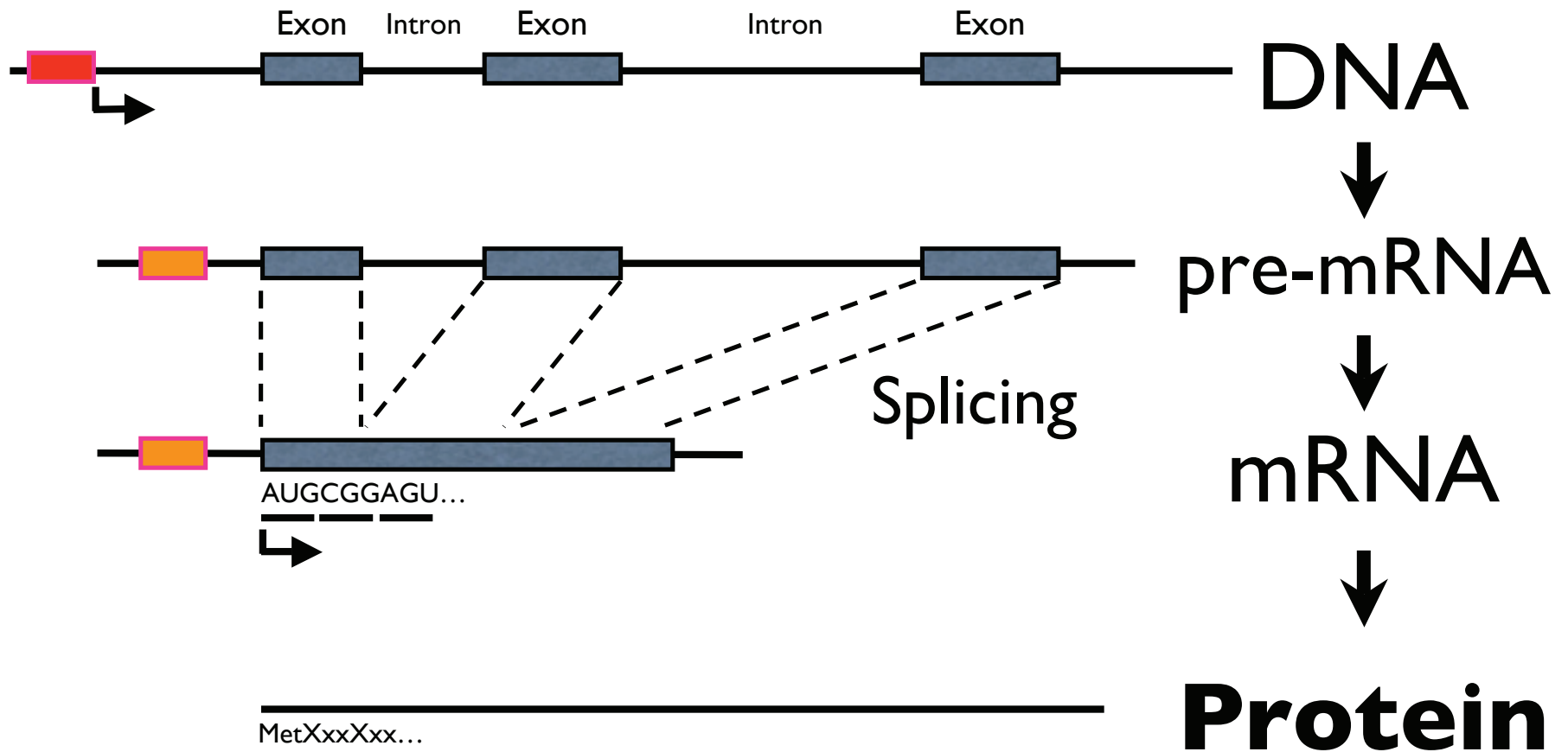
Cytochrome c oxidase

A sequence-specific and precisely controlled process

Seems simple?

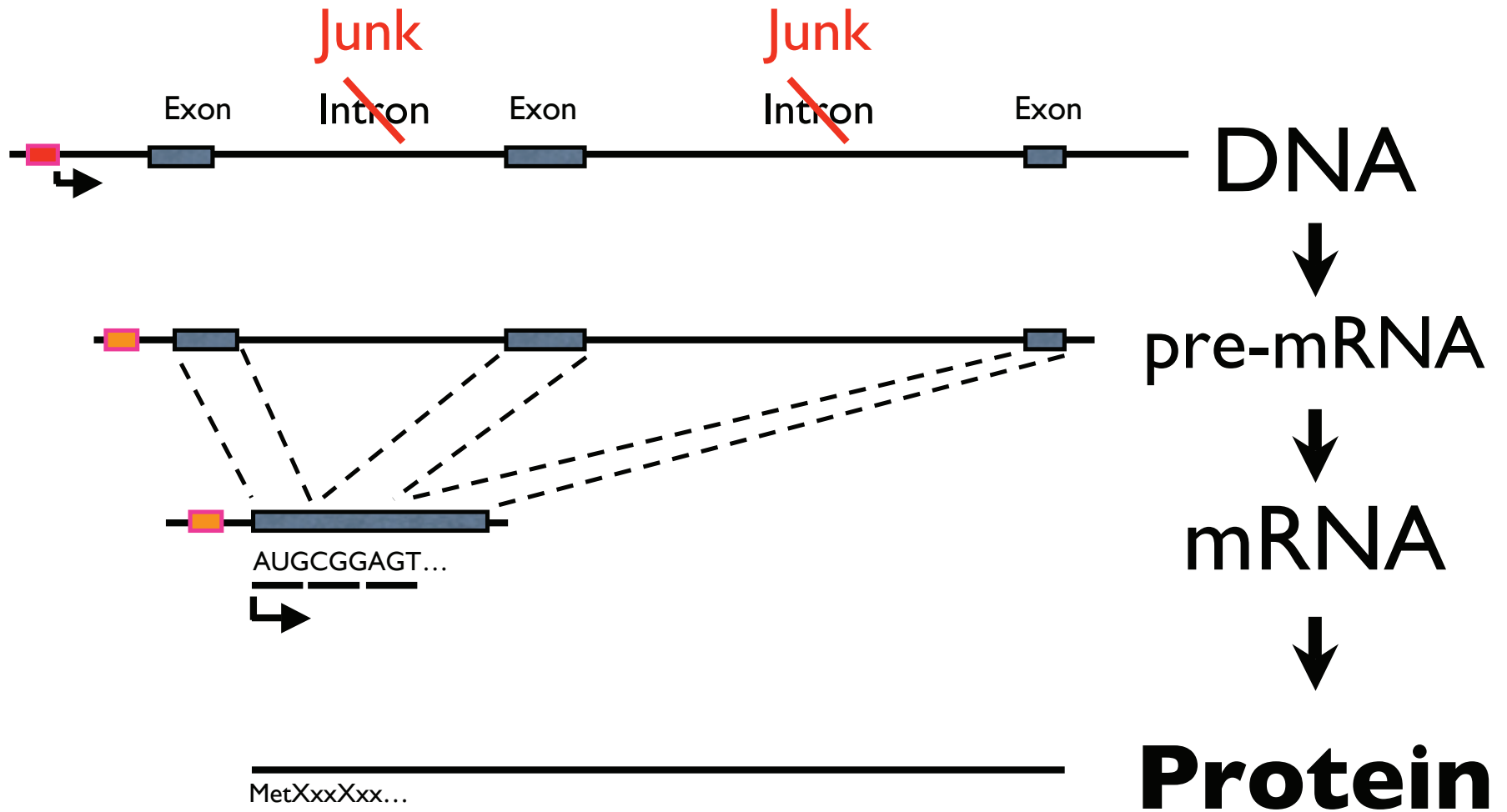


Even more complicated...



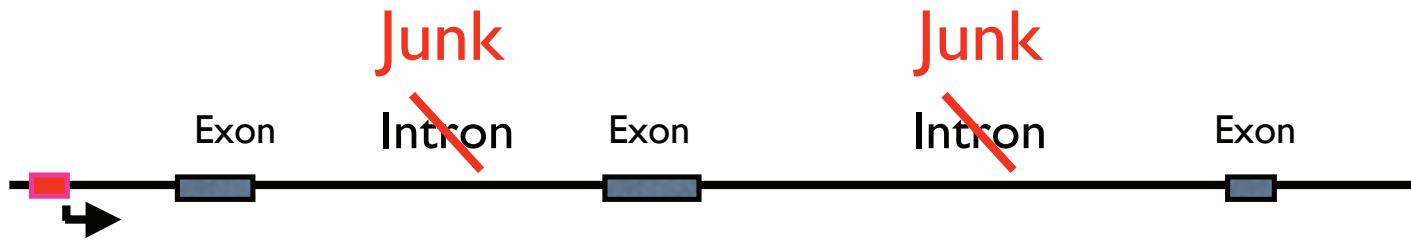
Reality...

and our perception of it



Reality...

and our perception of it



Genome project goals

Identify and characterize the proteins.

What are their structures?

What do they do?

How do they interact?

20th Century View

Kinases

Polymerases

Hydrogenases

Receptors

Oxygenases

Proteases

Nucleic acids → Trash

21st Century News

Genome project

Number of protein-encoding genes in the human genome: 25,000

Number of protein-encoding genes in the 1,000 cell *C elegans* genome: 19,500

Number of protein-encoding genes in the corn genome: 40,000

21st Century News

Genome project

Number of protein-encoding genes in the human genome: 25,000

Number of different proteins: \gg 25,000

How? RNA editing and alternative splicing

Up to 3/4 of all human genes are subject to alternative editing

The prevalence of alternative editing appears to increase with an organism's complexity

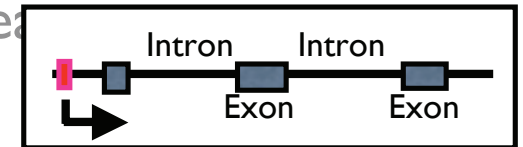
21st Century News

Genome project

How? RNA editing and alternative splicing

Up to 3/4 of all human genes are subject to alternative editing

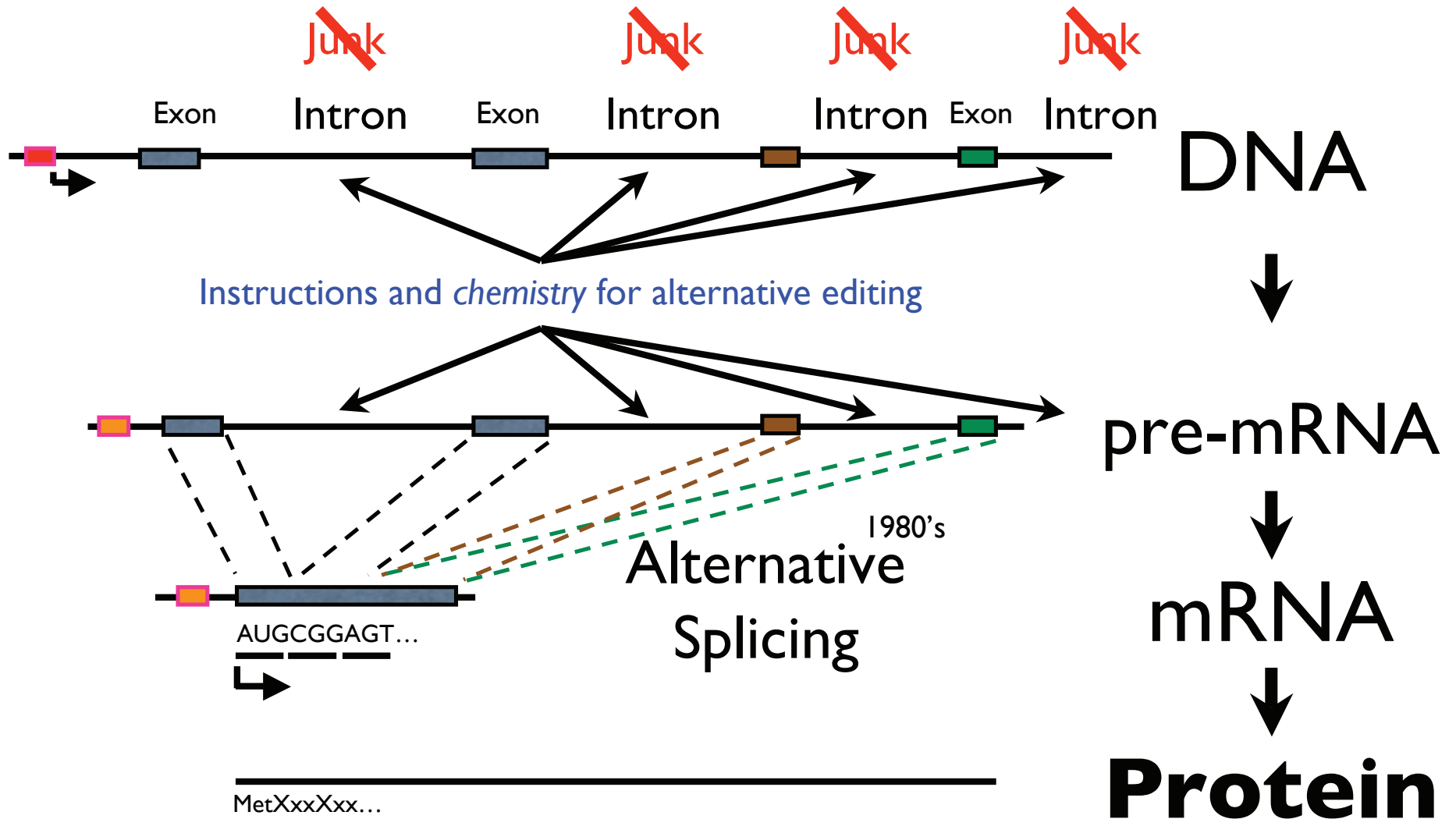
The prevalence of alternative editing appears to increase with an organism's complexity



Humans have the highest number of introns per gene of any organism

At least 15% of the gene mutations that produce genetic diseases and cancers do so by effecting pre-mRNA editing

21st Century News



Back in the lab...

With the advent of PCR, nucleic acids have been recognized as extremely powerful combinatorial tools in the test tube

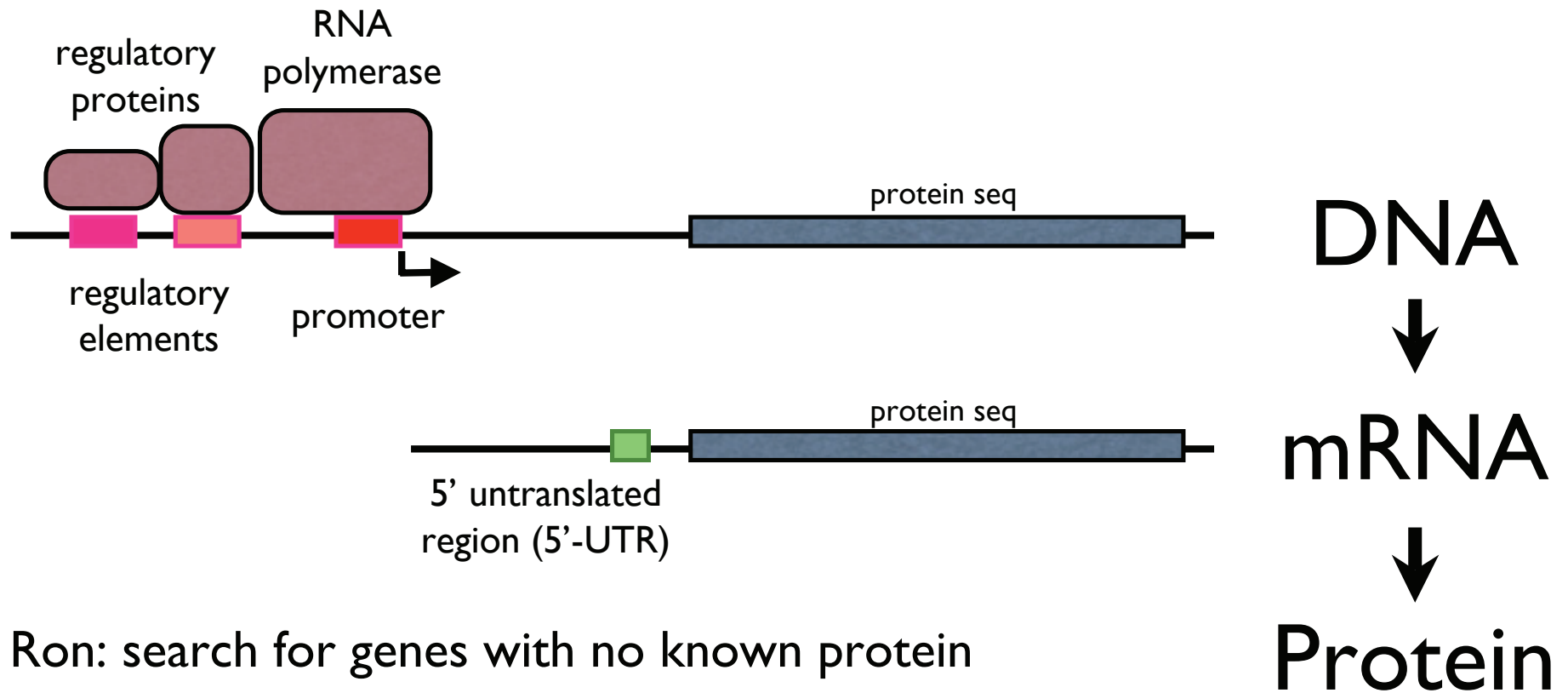
Aptamers can be selected that bind to
“your favorite molecule”

Can create riboswitches

Ron Breaker:

“if it’s so easy for us, I’ll bet nature exploits this”

More gene regulation



Ron: search for genes with no known protein regulator and which have a highly conserved 5' UTR

Does RNA from that conserved 5' UTR bind the product or substrate of the encoded enzyme?

→ One week: a Nature paper!

2004

Project Encode (2007)

(More) rewriting of textbooks

June 2007, published in the journal Nature

- ★ Some regions of DNA far from protein-coding genes (extreme “junk?”) are nevertheless highly conserved
- ★ Most of both strands of the DNA is transcribed (far beyond that required for protein-coding genes)

21st Century Opportunities

