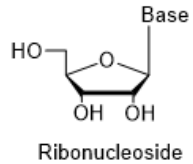
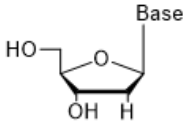
Nucleic Acid Structure

Nucleotides

We started off the discussion of nucleic acids with the components that make up these critical biological polymers. The individual nucleotides of DNA and RNA are shown below. We have purines and pyrimidine forms of ribonucleotides and deoxyribonucleotides. Note that ribonucleotides have a phosphate group linked while ribonucleosides do not. Also note the naming conventions for each

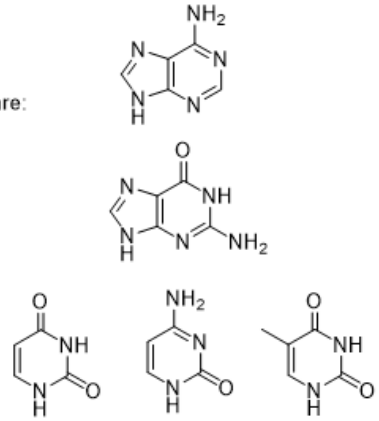


Ribonucleoside



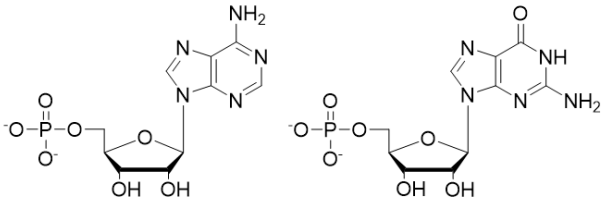
Deoxyribonucleoside

Bases by themselves are:
Adenine
Guanine
Uracil
Cytosine
Thymine



Ribonucleotides

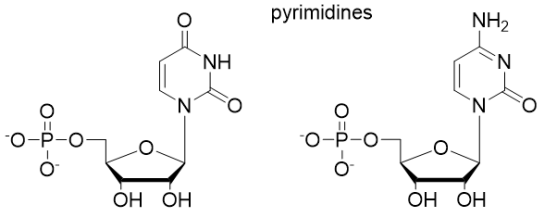
purines



Adeonosine monophosphate

Guanosine monophosphate

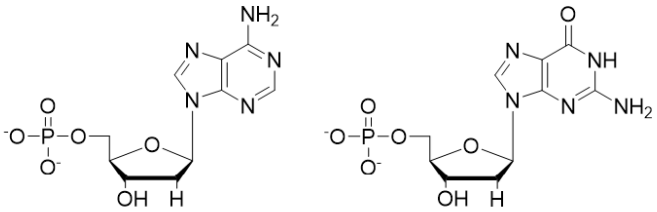
pyrimidines

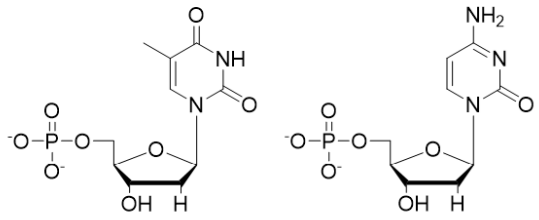


Uridine monophosphate

Cytidine monophosphate

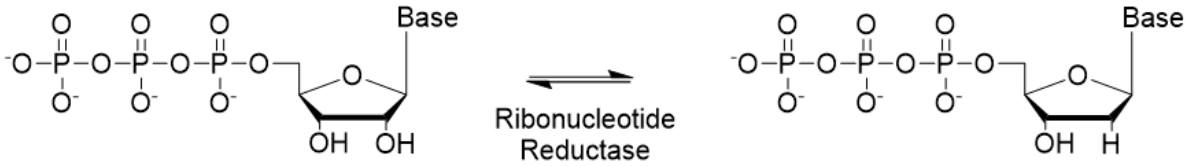
Deoxyribonucleotides





Deoxythymidine monophosphate

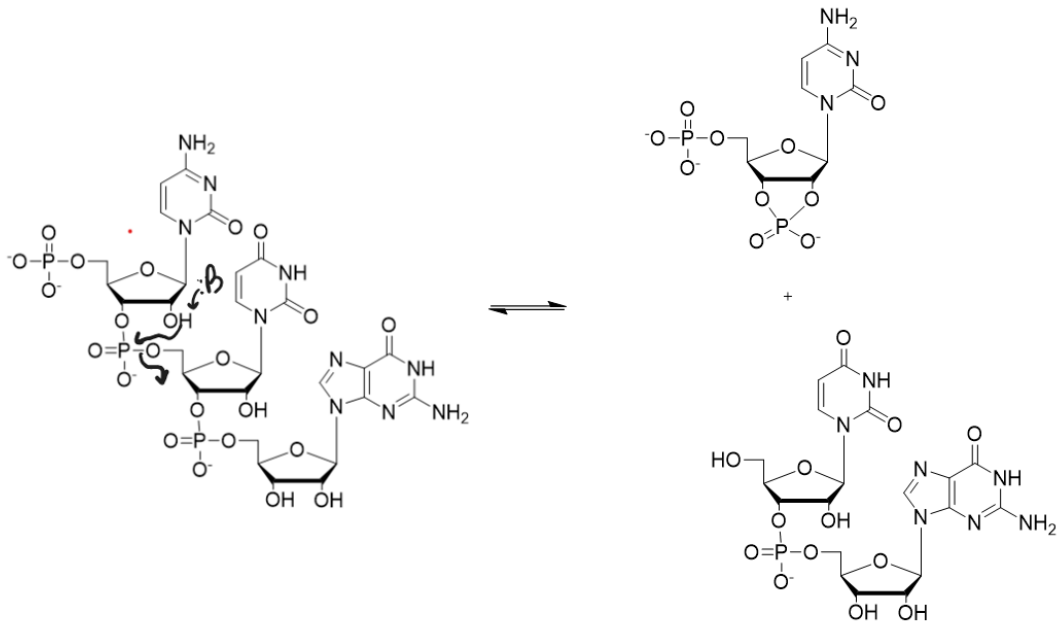
It is important to note that the deoxyribonucleotides are formed from ribonucleotides through the activity of the enzyme ribonucleotide reductase. Ribonucleotide reductase is present in every organism on the planet and was likely critical for the evolution of DNA as our primary information storage molecule. NTPs (nucleoside triphosphates) are required for the formation of RNA polymers, while dNTPs (deoxynucleoside triphosphates) are critical for DNA biosynthesis.



NTP dNTP

Nucleic Acids

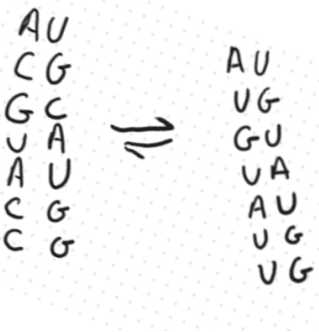
The half-life of the average RNA molecule is about 16 hours, while the half-life of DNA can be years. DNA is a significantly more stable molecule with respect to the phophodiester backbone. The primary reason for this enhanced stability is the lack of the 2'-OH group associated with DNA. This 2' OH group of RNA is prone to deprotonation which then allows for nucleophilic attack on the phosphodiester backbone and strand cleavage as shown below for the RNA sequence CUG. The cyclic phosphate formed can then hydrolyze to give a 2' or 3' phosphate.



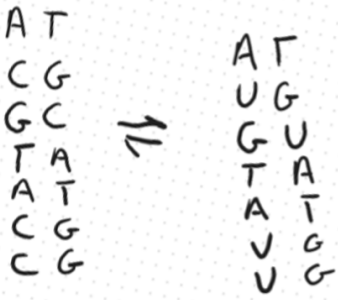
The ease with which RNA readily cleaved in a basic environment may have been beneficial early in the evolution of life as it would allow for sampling many different sequences and if those sequences were not protected or beneficial (polymer forming or self-replicating) then they could readily degrade.

The main difference between RNA and DNA is the presence or absence of the 2'-OH group. However, RNA can also be much more diverse than DNA. Many modified nucleotides have been observed in RNA molecules (alternative base structures). Evolutionarily speaking, one of the more important differences is that RNA has uracil bases while DNA has Thymine bases. The only difference between the two is the presence of an extra methyl group on Thymine. This development was also critical in the evolution of DNA as our primary information storage molecule. Cytosines can readily deaminate to form uracil. In RNA which has a short life-time this is not a significant problem. However, DNA can last for years and it is possible that over that time every cytosine in a DNA sequence could deaminate to uracil. In biology we have DNA repair enzymes that recognize the presence of uracils, and these enzymes will promote the removal of these uracils with cytosines. If DNA contained uracil instead of thymine these repair enzymes would not be able to distinguish which base is wrong (the uracil or the complementary strand guanine).

If DNA contained U instead of T it would be very difficult to determine which bases were supposed to be U and which were deaminated cytosine



Instead, since DNA does contain T instead of U, repair enzymes can quickly and easily detect the presence of deaminated C



Nucleic acid structure

Most of us are well aware of the classic double helical structure of DNA in which a sequence of DNA is bound to a complementary sequence. The complementary sequence is dependent on hydrogen bonds formed between A and T (two hydrogen bonds) and G and C (three hydrogen bonds). This classic Watson-Crick base pairing is shown below for a strand of DNA. There are other types of base pairing that exists in nature and that is taken advantage of in nucleic acids nanotechnology, but this is the primary type important for this course. The Watson-Crick structure of DNA is referred to as B-form DNA and is the most common structure that the molecule will form. Other forms are A-form and Z-form. A-form is more common in dsRNA and Z-form has been found in short stretches of prokaryotic and eukaryotic DNA. In the diagram below note that for A, B and Z form DNA it is showing 36 base pairs.

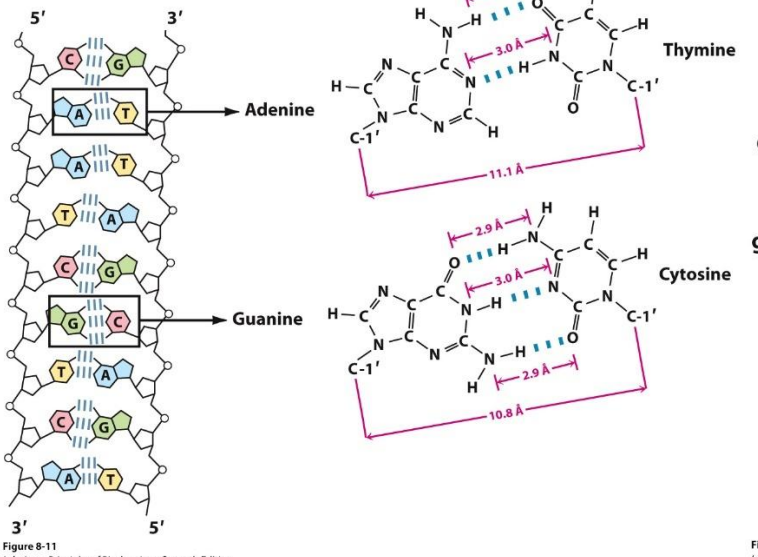


Figure 8-11
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

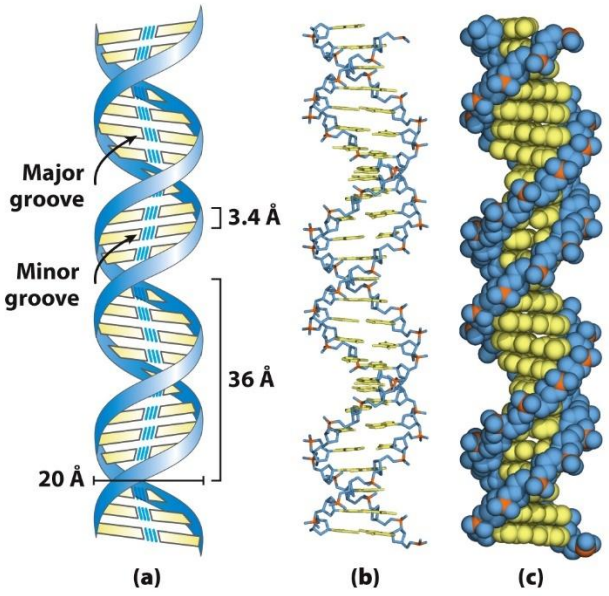


Figure 8-13
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

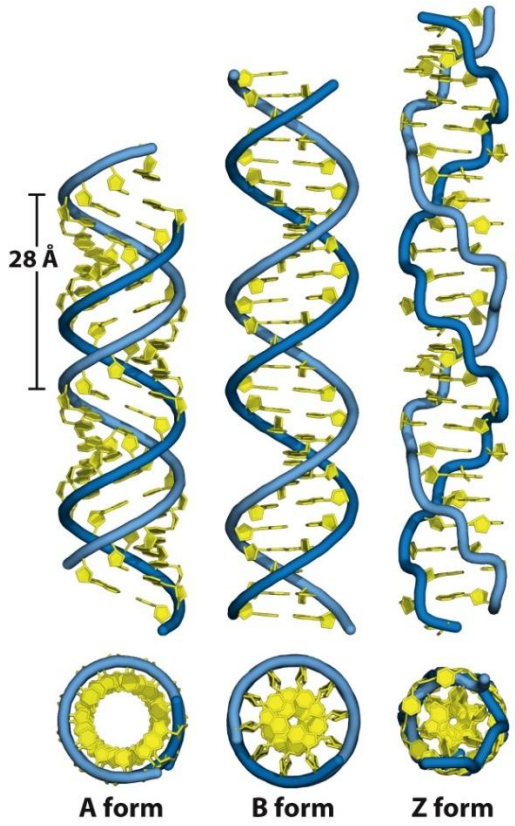


Figure 8-17 part 1
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

	A form	B form	Z form
Helical sense	Right handed	Right handed	Left handed
Diameter	~26 Å	~20 Å	~18 Å
Base pairs per helical turn	11	10.5	12
Helix rise per base pair	2.6 Å	3.4 Å	3.7 Å
Base tilt normal to the helix axis	20°	6°	7°
Sugar pucker conformation	C-3′ endo	C-2′ endo	C-2′ endo for pyrimidines; C-3′ endo for purines
Glycosyl bond conformation	Anti	Anti	Anti for pyrimidines; syn for purines

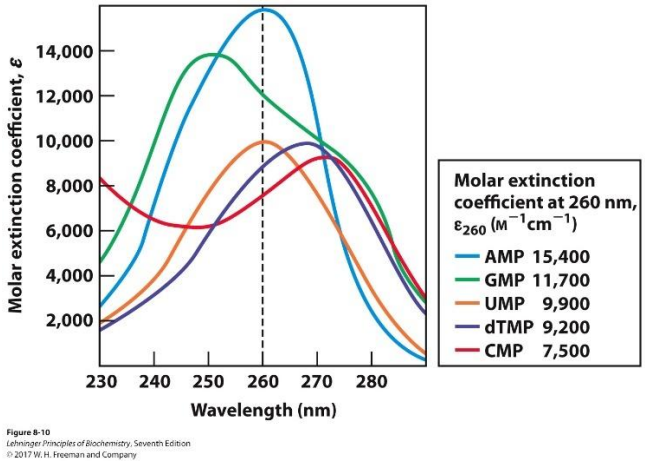
Figure 8-17 part 2
Lehninger Principles of Biochemistry, Seventh Edition
© 2017 W. H. Freeman and Company

The difluorotoluene experiment

Now the key question that we have to consider is why does DNA form the structure that it forms. Why is this important? First, the structure is important because structure dictates function. The double helical nature of the molecule leads to the information content (base sequence) being buried on the inside of the molecule. Weak forces maintain this structure which can be readily pulled apart to access that information (think Velcro analogy) whether we are talking about DNA replication or transcription of particular genes. Most of the time the primary focus when DNA structure is discussed is on hydrogen bonding. However, it is critical to note that hydrogen bonding is not the only weak

force at play and I would argue that hydrogen bonding is more important for the specificity of the interactions between the two strands and that the bases will face each other anyway due to the hydrophobic effect.

Before discussing the weak forces, let's first talk about experiments conducted by Eric Kool's group at Stanford University. These are especially interesting because they use chemical probes to look at the impact of hydrogen bonding on the stability of the DNA double helical structure. In this work, the group synthesized an analogue of thymidine which had a similar shape to thymine but was incapable of hydrogen bonding. They found that this synthetic nucleotide could be incorporated into DNA across from adenine by a DNA polymerase even though it did not hydrogen bond, suggesting that the DNA polymerase was influenced more by the shape of the structure than the ability to hydrogen bond. It should also be noted that this new base (difluorotoluene) was significantly more hydrophobic than thymine. Below is data from Kool's work where the group reports the melting temperature of double stranded DNA as a measurement of double-strand stability. These measurements are relatively straightforward where a sample of DNA is heated to increasing temperatures and a shift in absorbance occurs when the majority of the DNA in the sample melts to single stranded DNA. This shift in absorbance observed in these experiments can be attributed to two possibilities 1) pi stacking (to be discussed below) is not as prevalent in single-stranded DNA and this loss of stacking interactions between bases will impact the absorbance of the DNA. 2) When DNA forms hydrogen bonds with the complementary base this locks the base into its classic structure that you memorized above. However, when DNA is single-stranded the conformation of the bases can change and these structures would have a different absorbance associated with them. Even if only a small percentage of the bases undergo this change it can still have an impact on the absorbance. Also, note that the absorbance at 260 nm can be used to determine the concentration of an unknown solution of nucleic acid using Beers-Lambert law A=ECl where A is the absorbance at 260 nm, E is the extinction coefficient (~10000 $M^{-1}cm^{-1}$ per nucleotide), C is the concentration and l is the path length (typically 1 cm).



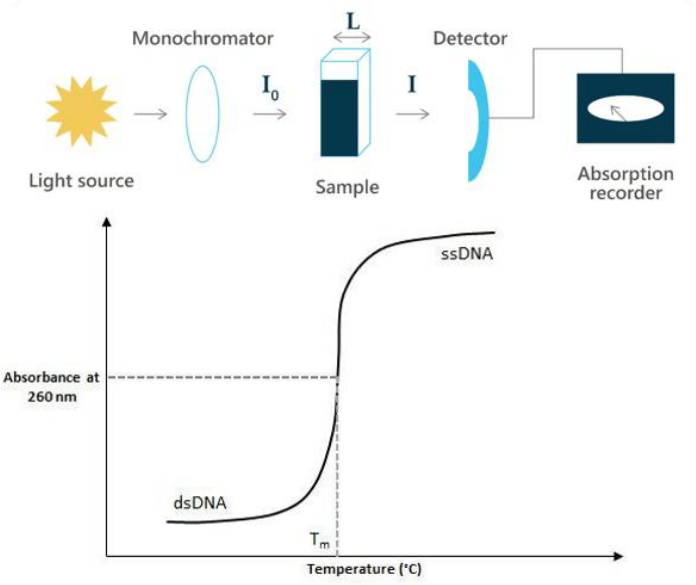
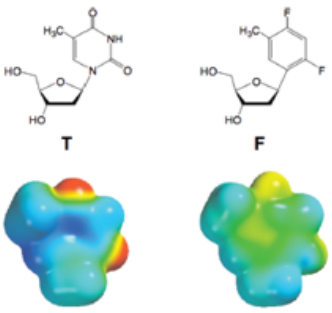


Figure 8-10
Lehninger Principles of Biochemistry, Seventh Edition



JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Measurement and Theory of Hydrogen Bonding Contribution to Isosteric DNA Base Pairs

Omid Khakshoor,[†] Steven E. Wheeler,[*,§] K. N. Houk,[‡] and Eric T. Kool[*,†]

[†]Department of Chemistry, Stanford University, Stanford, California 94305, United States
[‡]Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States
[§]Department of Chemistry, Texas A&M University, College Station, Texas 77843, United States

12mer duplex		T_m @ 3 μM (°C)
X•Y	5'-TGTAT X CGTGCG 3'-ACATA Y GCACGC	
T•A	5'-TGTAT T CGTGCG 3'-ACATA A GCACGC	51.3
T•G	5'-TGTAT T CGTGCG 3'-ACATA G GCACGC	43.4
T•T	5'-TGTAT T CGTGCG 3'-ACATA T GCACGC	38.1
T•C	5'-TGTAT T CGTGCG 3'-ACATA C GCACGC	37.4
F•A	5'-TGTAT F CGTGCG 3'-ACATA A GCACGC	38.7
F•G	5'-TGTAT F CGTGCG 3'-ACATA G GCACGC	36.0
F•T	5'-TGTAT F CGTGCG 3'-ACATA T GCACGC	34.6
F•C	5'-TGTAT F CGTGCG 3'-ACATA C GCACGC	33.9

Table 3. Pairing and Stacking Contributions As Measured in a Dangling End Context[a]

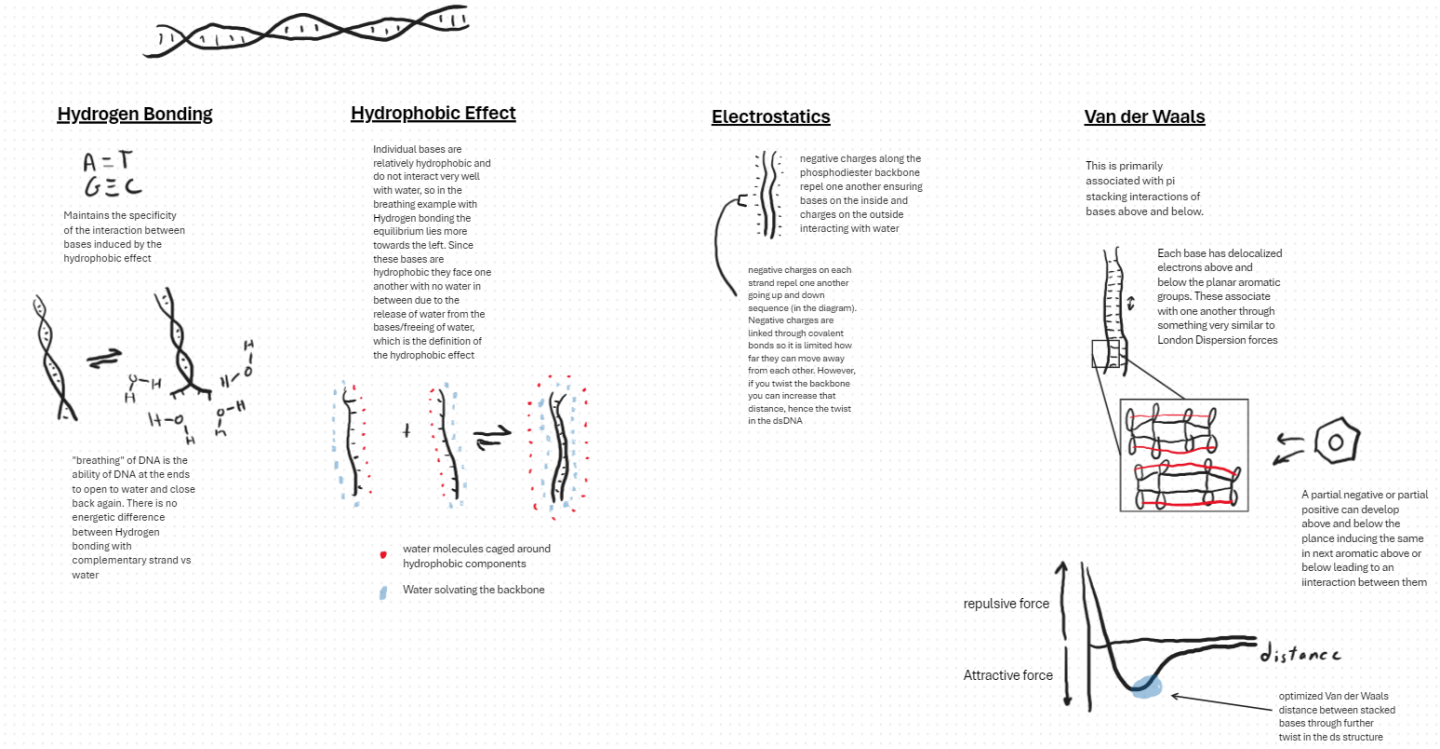
DNA duplex	T_m @ 6 μM (°C)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (cal/molK)	$\Delta G°_{37}$ (kcal/mol)
5'-CGCGCG GCGCGC-5'	43.1	−53.1 (±2.1)	−144 (±7)	−8.44 (±0.06)
5'-TCGCGCG GCGCGCT-5'	48.7	−56.5 (±2.1)	−152 (±7)	−9.44 (±0.07)
5'-CGCGCGA AGCGCGC-5'	46.3	−56.3 (±1.8)	−152 (±6)	−9.06 (±0.06)
5'-TCGCGCGA AGCGCGCT-5'	51.4	−61.5 (±1.8)	−166 (±5)	−10.11 (±0.08)
5'-FCGCGCG GCGCGCF-5'	54.0	−60.0 (±1.9)	−160 (±6)	−10.49 (±0.10)
5'-FCGCGCGA AGCGCGCF-5'	54.5	−61.8 (±2.2)	−165 (±7)	−10.70 (±0.13)

[a]Conditions: 100 mM NaCl, 10 mM Na_2HPO_4 (pH 7.0), 0.1 mM EDTA; [DNA] = 2−12 μM . Thermodynamic values obtained from averaging values from curve fits and van't Hoff plots (see Experimental Section).

Now we know how the measurements were made we can start to look at the data (note that the table on the right above also measured full thermodynamic parameters using a technique called isothermal titration calorimetry, we may talk about this method later in the class). The group prepared dsDNA with the TA base pair replaced with an FA base pair either in the middle (table on the left) of the sequence or at the end of the sequence (table on the right). They found that in the middle of the sequence the change resulted in a lower melting temperature and therefore destabilization of the double helix. However, when the AF was added to the end of the sequence they observed an increase in stability of the double-stranded DNA as apparent from the increase in Tm. In class, I discussed why this might be. First, we have to consider the role of hydrogen bonding in the interaction of the complementary sequences, and a key issue that is often overlooked is that DNA exists in water. Hydrogen bonding to water is just as likely as hydrogen bonding to a complementary base. This leads to the ends of the DNA sequence "breathing". Meaning it can open and close because there is no real preference for hydrogen bonding with either one and hydrogen bonding with water may in fact be more favorable than with the complementary base. So, something else must be responsible for the bases to face inwards. The primary reason that the bases face each other is that they are relatively hydrophobic. Adenine solubility in water is 1 mg/mL, adenosine nucleoside is 10 mg/mL and AMP is 70 mg/mL. The base itself is quite hydrophobic. Difluorotoluene is even more hydrophobic than thymine and cannot hydrogen bond. This leads to the end of the molecule being more closed off from water than the natural base. There is less competition with water binding to difluorotoluene and therefore the double helix is stabilized when the analogue is at the end. When the difluorotoluene is in the middle of the sequence there is a different impact. Above and below the difluorotoluene are base-paired complementary sequences. The hydrogen bonding associated with the base pairs leads to a close association between the strands. If there is no base

pairing the sequence will bulge where the difluorotoluene is. This disrupts pi stacking above and below and weakens the interaction between the two strands destabilizing the dsDNA.

Weak forces and the structure of nucleic acids

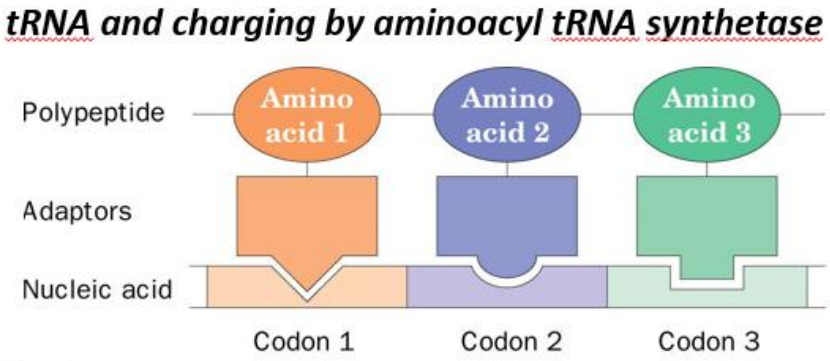


So we addressed hydrogen bonding and I argue that hydrogen bonding is more relevant to the specificity of interactions between strands rather than the fact that bases face inwards in the structure of double-stranded DNA. Now we need to look at the other weak forces and how they contribute. Above I also noted the importance of the hydrophobic effect on DNA double strand structure. The bases are relatively hydrophobic and therefore face away from water pointing towards the interior of the double-stranded molecule. The hydrophobic effect is by far the most important for the overall structure and leads to the bases facing one another, which then leads to hydrogen bonding maintaining those interactions through classic Watson-Crick base pairing. The next weak force that we considered was the role of electrostatic interactions. The phosphodiester backbone of DNA has a negative charge at each nucleotide due to the phosphate backbone. Due to this negative charge, the backbone of each strand could not interact directly with one another due to charge-charge repulsion. In addition, water surrounding DNA would promote favorable solvation of these ions in the phosphate backbone so they will face water. This addresses the negative charge on each of the strands, so now we have to consider the negative charge going up or down either strand. As shown in the diagram below, the negative charges would stack on top of one another in the DNA sequence. However, this stacking would lead to additional charge-charge repulsion. The fact that these charges are linked through covalent bonds to one another, restricts the ability of these charges to move away from one another. Since they cannot move away from one another vertically, the double-stranded molecule twists to optimize the distance between these negative charges. Therefore, the negative charge in the backbone helps to induce a twist in the overall structure. There is one more contributor to the twist in the overall DNA structure. Each of the bases is aromatic, which makes them relatively planar with delocalized electrons above and below the rings. These rings are stacked on top of one another with delocalized electrons above and below the plane. The distance between these ring structures, like the distance between negative charges in the backbone, is limited in how far away they can get from one another due to the covalent linkages between them (through the phosphodiester-ribose backbone). However, similar to London Dispersion forces (the third category of Van der Waals forces discussed) the pi orbitals will fluctuate between more negative and more positive. As in London Dispersion forces this will lead to induction of an optimized interaction between bases that are stacked on top of one another. Since there

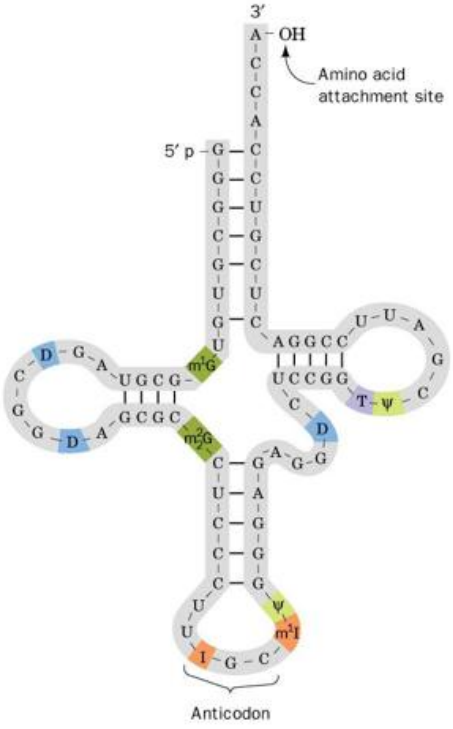
is no way to push them away from each other the bases must twist relative to each other to optimize the distance for attractive forces.

RNA structure

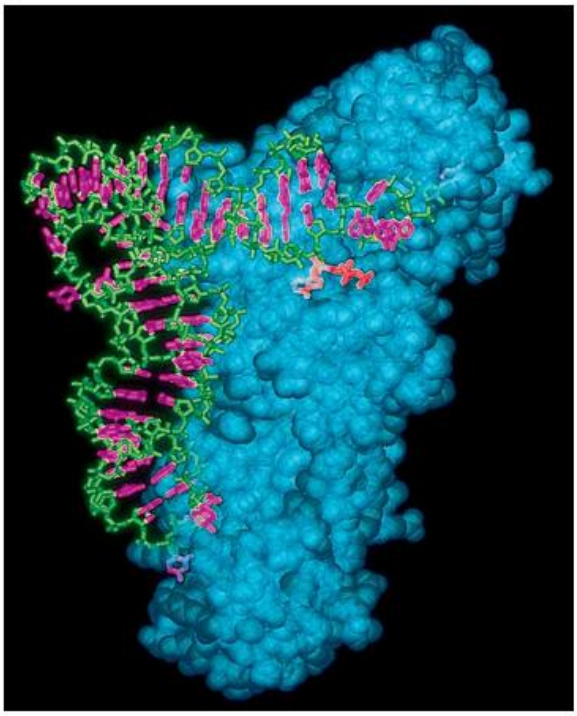
DNA is almost always found in double-stranded form where it is bound to a complementary sequence as described above. However, RNA is rarely found in double-stranded form. However, this is not because RNA will not form double-strands. This is primarily because there is no complementary sequence available. Of course viral RNA is often double-stranded RNA, so it is definitely possible, just not typically how we find it. Instead, RNA is usually single-stranded and folds onto itself for the same reasons that DNA forms the classic double helix structure. The weak forces apply in the exact same way, the structures are different simply because of the lack of a complementary sequence with RNA. The primary example used in class for RNA structure was tRNA (transfer RNA). This RNA is responsible for carrying amino acids to the ribosome and localizing them through their anti-codon loop to codons on an mRNA (the sequence transcribed from DNA that serves as the template for protein. tRNA structure was shown in two forms: 1) a secondary structure depiction and 2) a tertiary structure depiction. The secondary structure simply shows hydrogen bonding networks and can be useful for simplifying the overall 3D structure (tertiary structure



Secondary structure depiction of tRNA



Tertiary structure depiction of tRNA



In the diagram above, I show the tertiary structure of tRNA bound to an aminoacyl tRNA synthetase enzyme. This enzyme is responsible for catalyzing the addition of an amino acid to the appropriate tRNA. In this depiction at the

top to the right part of the tRNA is buried within a pocket in the enzyme where the addition of the amino acid takes place. On the bottom left part of the diagram the anti-codon loop which specifies the codon that the tRNA interacts with is buried in the bottom part of the enzyme. The anticodon loop interacts through classic Watson-Crick base pairing with an mRNA. The sequence of the tRNA in the anticodon loop is typically a classic Watson-Crick complement to the specified mRNA codon with the first two residues then the third residue is often a non-canonical nucleotide that can interact with multiple nucleotides. The is "wobble" base pairing of that third nucleotide is what makes the genetic code degenerate where more than one sequence can encode a particular amino acid.

First position (5′ end)	Second position				Third position (3′ end)
	U	**C**	**A**	**G**	
U	UUU Phe	UCU	UAU Tyr	UGU Cys	**U**
	UUC	UCC Ser	UAC	UGC	**C**
	UUA Leu	UCA	UAA Stop	UGA Stop	**A**
	UUG	UCG	UAG Stop	UGG Trp	**G**
C	CUU	CCU	CAU His	CGU	**U**
	CUC	CCC	CAC	CGC	**C**
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	**A**
	CUG	CCG	CAG	CGG	**G**
A	AUU	ACU	AAU Asn	AGU Ser	**U**
	AUC Ile	ACC	AAC	AGC	**C**
	AUA	ACA Thr	AAA Lys	AGA	**A**
	AUG Met[b]	ACG	AAG	AGG Arg	**G**
G	GUU	GCU	GAU Asp	GGU	**U**
	GUC	GCC	GAC	GGC	**C**
	GUA Val	GCA Ala	GAA Glu	GGA Gly	**A**
	GUG	GCG	GAG	GGG	**G**

[a]Nonpolar amino acid residues are tan, basic residues are blue, acidic residues are red, and nonpolar uncharged residues are purple.
[b]AUG forms part of the initiation signal as well as coding for internal Met residues.

The ribosome is primarily made of RNA and includes accessory proteins to enhance its function. The first structures of the ribosome led to Nobel prizes back in 2009. Below is the structure of a eukaryotic ribosome in blue is the small subunit, in gray is the large subunit, purple are accessory proteins that interact with the ribosome. In the figure to the right, the small and large subunit of the ribosome are separated computationally so that you can see the location of tRNAs bound to the structure. These tRNAs are in yellow, orange, and red and can be seen located in what are termed the ribosome E, P, and A sites for exit, peptidyl transfer, and acceptor sites. Note that on the small subunit where and mRNA would bind you can see the anti-codon loops of the tRNAs. In the large subunit you can see how the region that would be linked to an amino acid is buried within a pocket of RNA. This is a fantastic demonstration of the importance of RNA in protein synthesis as you can see that the site where peptides are formed is completely dependent on RNA. The ribosome is a ribozyme where RNA is responsible for catalyzing the formation of a peptide. This holds well for our discussion of the evolution of life, and is the focus of a paper I describe below. The chemistry that is actually catalyzed by the ribosome is shown below. Where a growing peptide chain is linked to a tRNA in the P-site. The next tRNA binds to the

acceptor site and the amine of that amino acid attacks the peptide on the P-site tRNA where the tRNA in the P-site is the leaving group. This continues along the mRNA as the protein is synthesized.

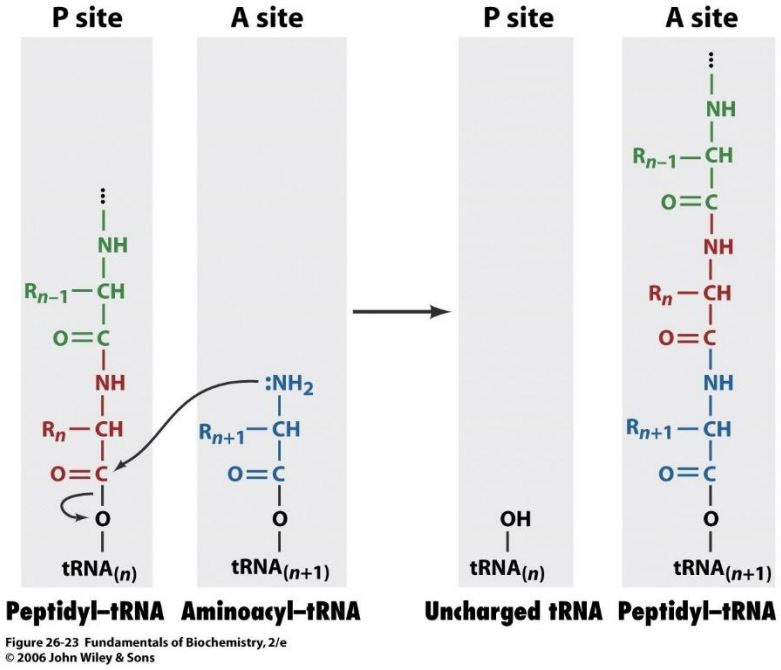
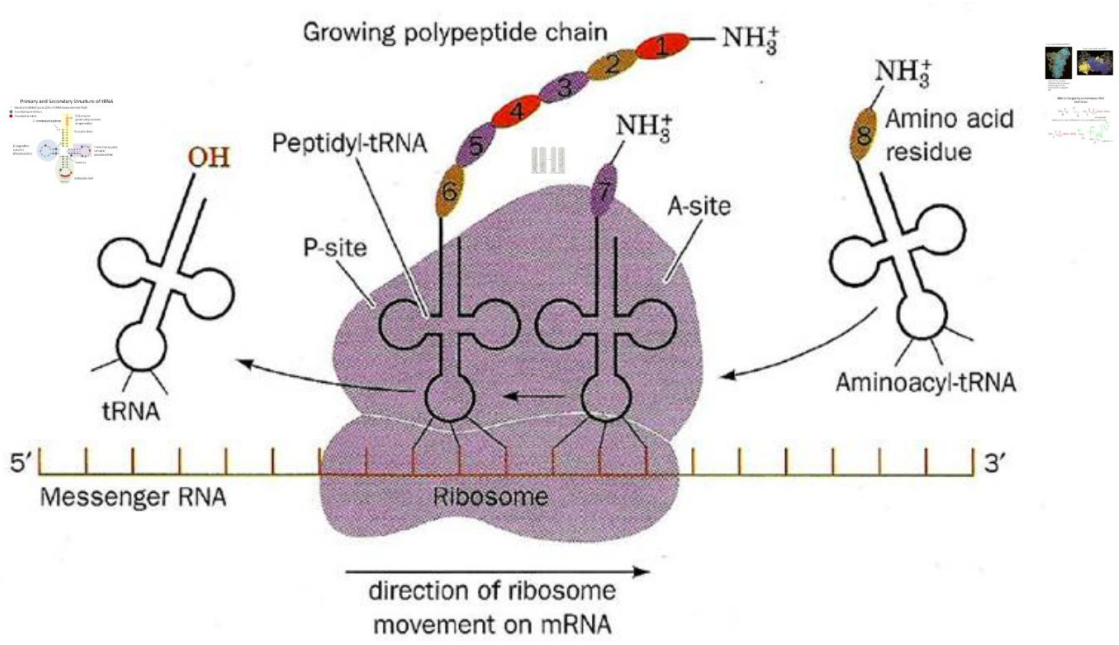
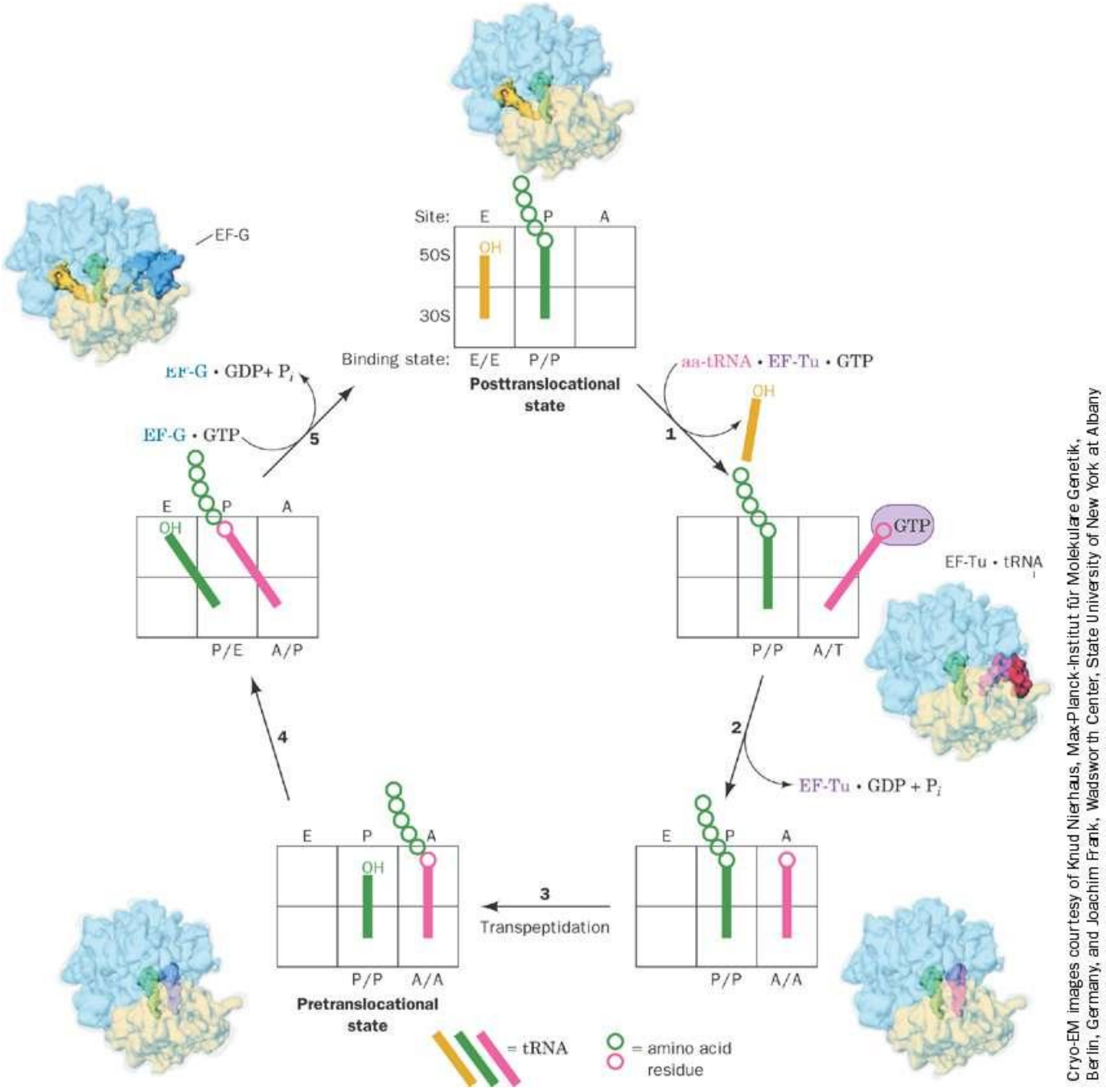


Figure 26-23 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons



The ribosome can be described as a molecular ratchet that ratchets the mRNA through it using its interactions with the tRNA and the tRNA interaction with mRNA. The diagram below is intended to show this, where at the top you see a free tRNA in the E site and an elongating peptide linked to tRNA in the p-site, with an empty A site. The mRNA is at the bottom of the box where the anticodon loop interacts and the circles represent amino acids. When the next tRNA binds it binds first through the anti-codon loop to an mRNA codon at the A-site. GTP hydrolysis promotes a shift in conformation adjusting the position of the new amino acid in the A-site in close proximity to the peptide chain in the P-site. In step 3 we have catalysis where the amino acid in the A-site attacks the peptide at the amino acid linked to tRNA. The tRNAs then shift the region that carries or carried the amino acid into the E and P-site then the tRNA drags the mRNA into the next position placing the next codon into the A-site and the peptide in the P-site. After this the cycle just

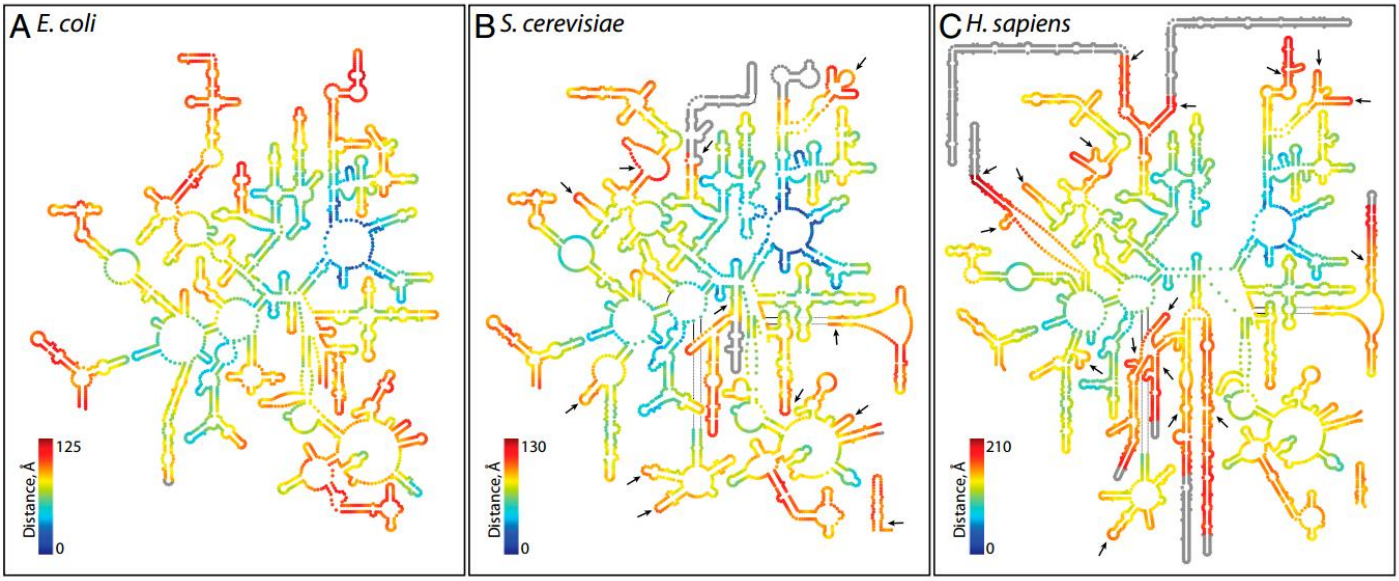
continues on over and over until the protein is completed and the final protein is cleaved from the last tRNA and released into the cell.



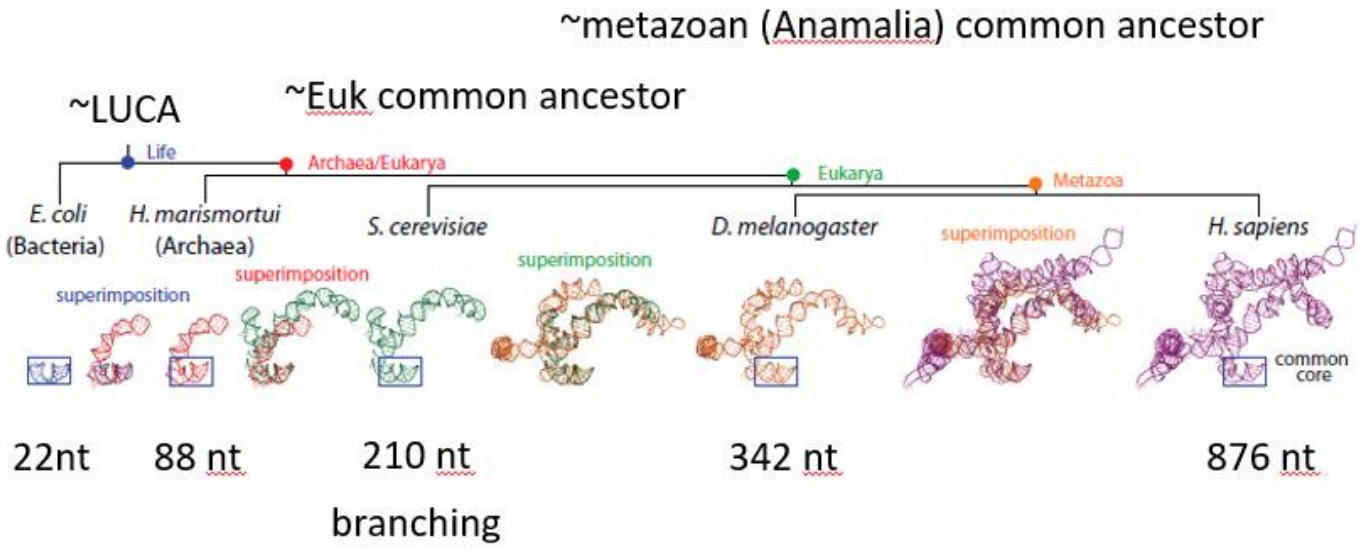
Evolution of the Ribosome at Atomic Resolution

Next up I described one of my favorite research articles of all time, which is a paper from Loren William's group at Georgia Tech. This work took advantage of the fact that we have the structure of the ribosome from many different organisms at various levels across the evolutionary scale. The idea behind this work was that we should be able to map the evolution of the ribosome all the way back from the last universal common ancestor (LUCA) by analyzing the structures of these ribosomes and finding regions of expansion and applying what we find at those regions to the simplest ribosomes to trace back the evolution of the ribosome. In the first set of data shown I provided the secondary structure depictions of a bacterial, yeast, and human ribosome large subunit. The distance from the peptidyl transfer center (PTC) of the ribosome in 3D space is mapped onto these depictions where very close regions are in blue and

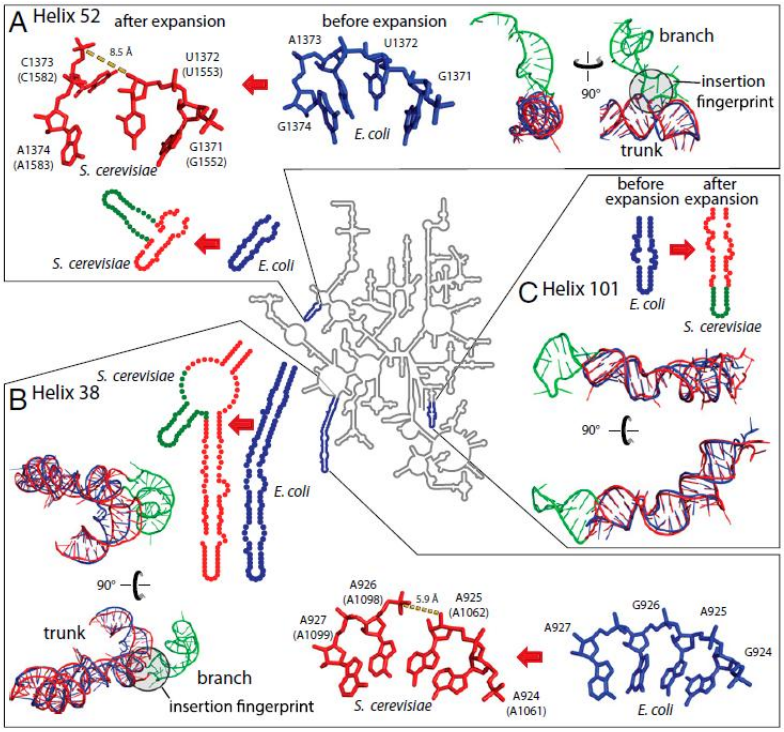
regions far away are in red. The basis of this was the idea that the peptidyl transfer center was likely the first component that evolved, and this goes along with my earlier arguments about the RNA world and the need for a self-replicating sequence to emerge as well as a polymer forming sequence (the PTC of the ribosome).

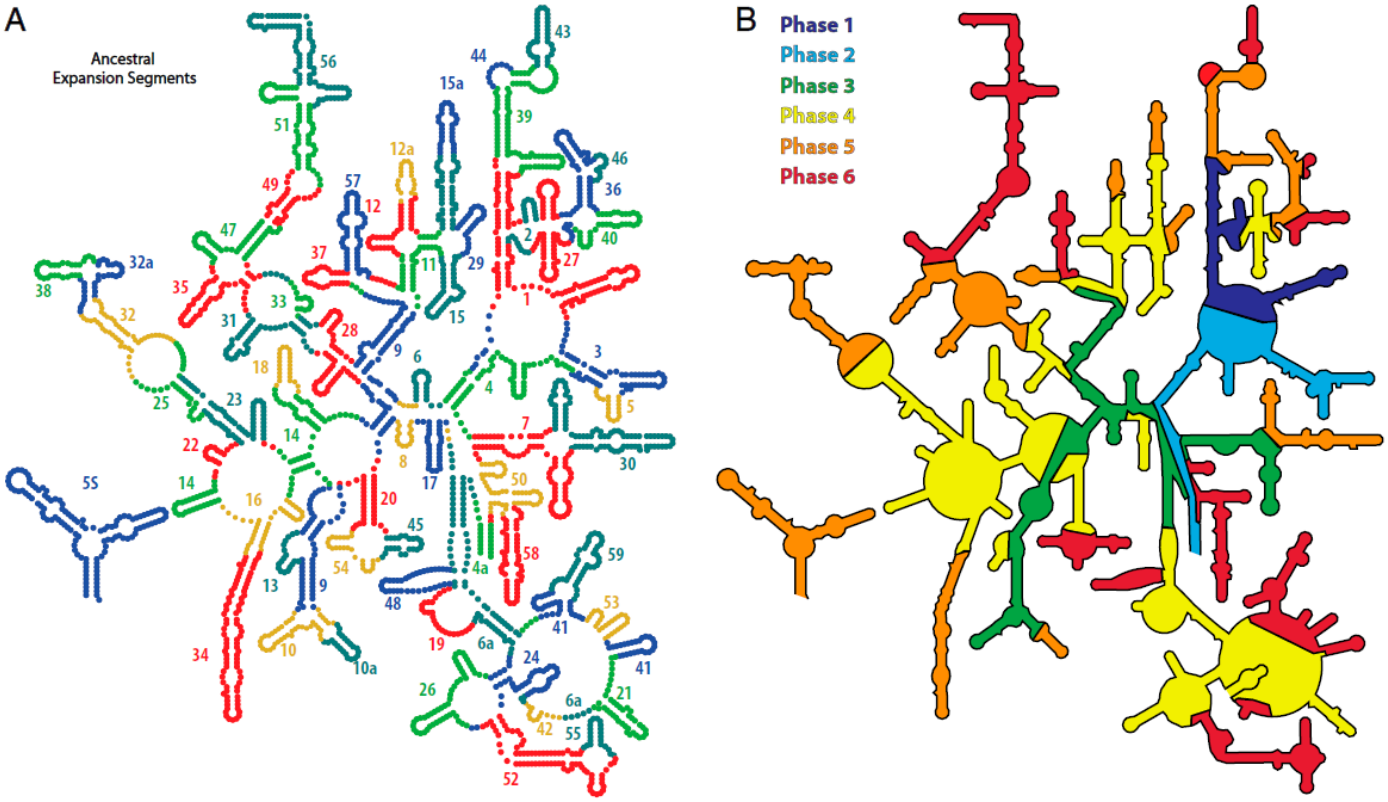


From this, the group found numerous regions that were similar between the structures and so they began to look at how these regions change as we get to more complex organisms as shown below for bacteria, archaea, yeast, fruit flies, and humans. What you see in the diagram below is that the 22 nucleotide region in bacteria is conserved all the way up to humans. As we increase organism complexity we can see increasing complexity adjacent to this conserved region and you can superimpose these structures to see these changes.



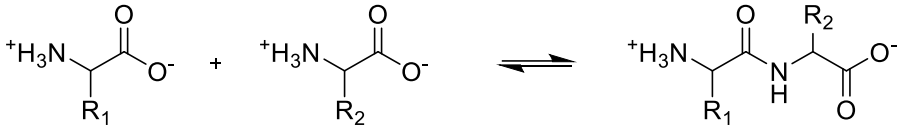
Next, Dr. William's group looked into these regions of expansion and looked for insertion fingerprints that they could then use to look at the simplest ribosomes to try to map the evolution of the ribosome. They found specific structural features where there were branches or expansions in these sequences. The group took these molecular features and identified what they called ancestral expansion segments that were used to try to figure out what parts of the ribosome changed and when relative to one another. This led to the final diagram that shows the PTC as the starting point and they described the likely step-by-step evolution as shown in the second figure below.


A Helix 52
after expansion
before expansion
C1373 (C1582)
8.5 Å
U1372 (U1553)
A1374 (A1583)
G1371 (G1552)
S. cerevisiae
A1373
U1372
G1371
G1374
E. coli
branch
insertion fingerprint
90°
trunk
S. cerevisiae
E. coli
B Helix 38
S. cerevisiae
E. coli
90°
trunk
branch
insertion fingerprint
C Helix 101
before expansion
after expansion
E. coli
S. cerevisiae
90°
A926 (A1098)
5.9 Å
A925 (A1062)
A927 (A1099)
A924 (A1061)
S. cerevisiae
G926
A925
A927
G924
E. coli

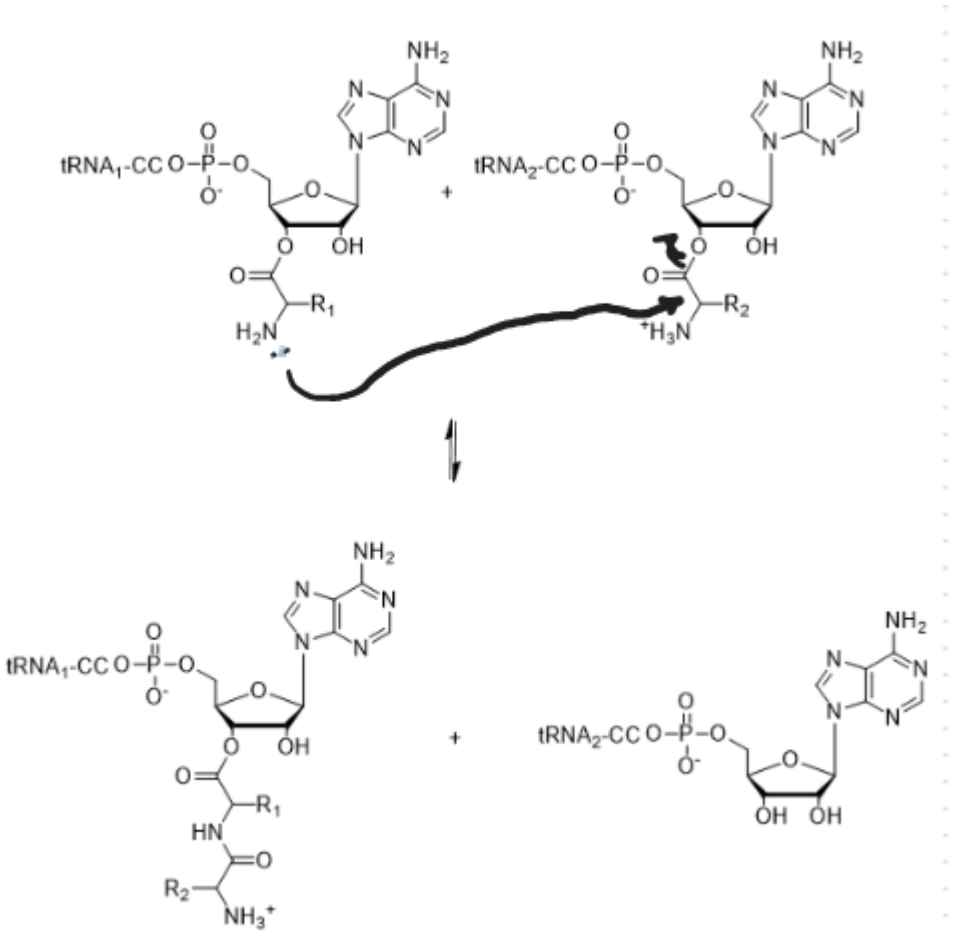

A
Ancestral Expansion Segments
B
Phase 1
Phase 2
Phase 3
Phase 4
Phase 5
Phase 6

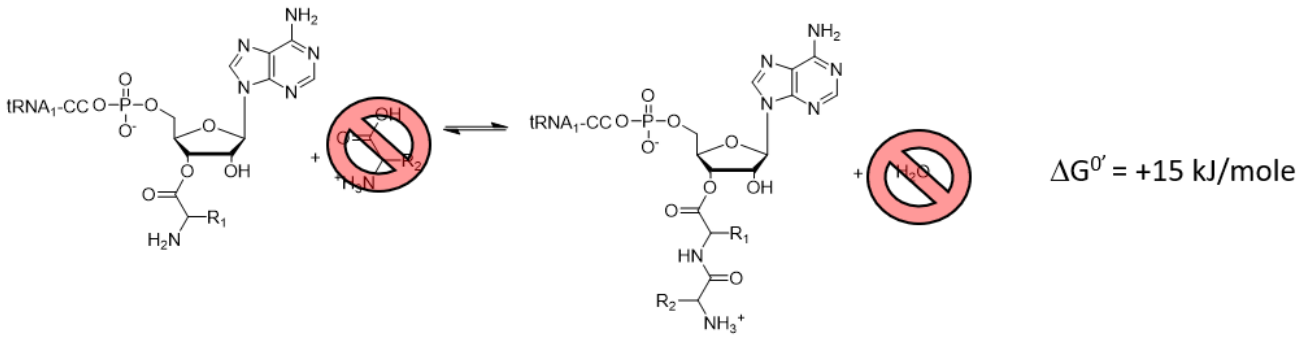
Free energy of peptide formation

Peptides are formed through the condensation of amino acids to give an amide bond linkage between the two amino acid residues. I started this discussion with a key question. When two amino acids come together to form an amide linkage that process is not favorable and has a $\Delta G^{0'}$ of +15 kJ/mole (the value for amide bond formation).



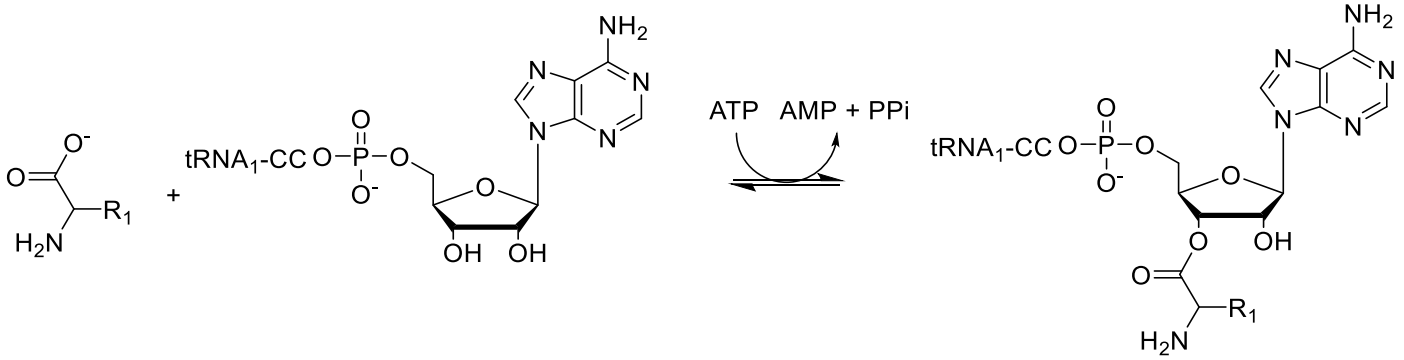
Proteins are "the workhorses of a cell" and indeed to much of the critical work required for complex life to exist. It is not possible that the formation of proteins is an unfavorable process. This means that it is essential to think about how Biology actually achieves the formation of peptide linkages. The key point here is that proteins are not made simply by condensing two amino acids together. Instead, each of those amino acids are linked to a tRNA and that ester linkage to that tRNA is critical for the process to be favorable. As shown below amide bond formation costs +15 kJ/mole. However, ester bond hydrolysis provides -20 kJ/mole. Therefore the overall process in biology is actually closer to -5 kJ/mole which is favorable.



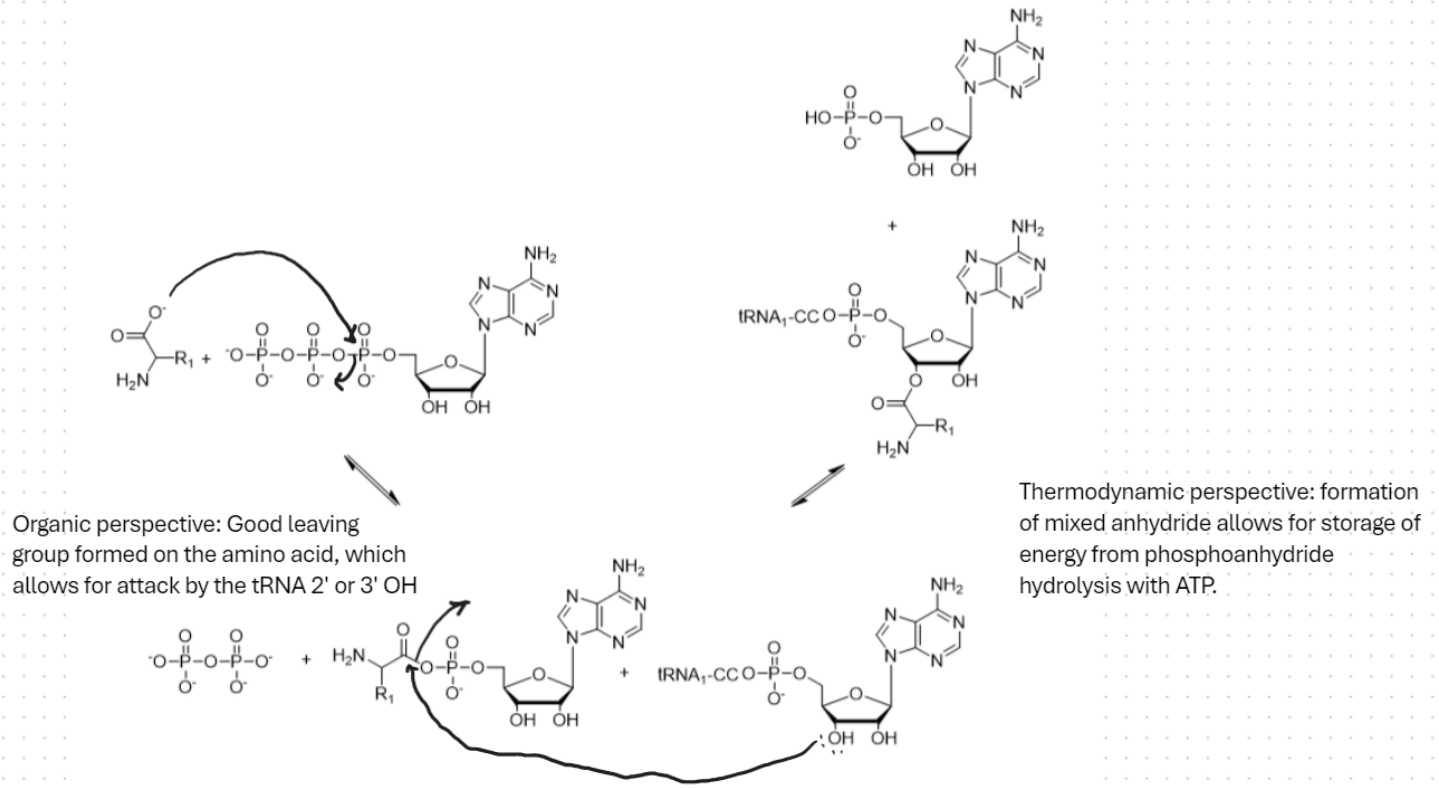


ΔG0′ = -5 kJ/mole

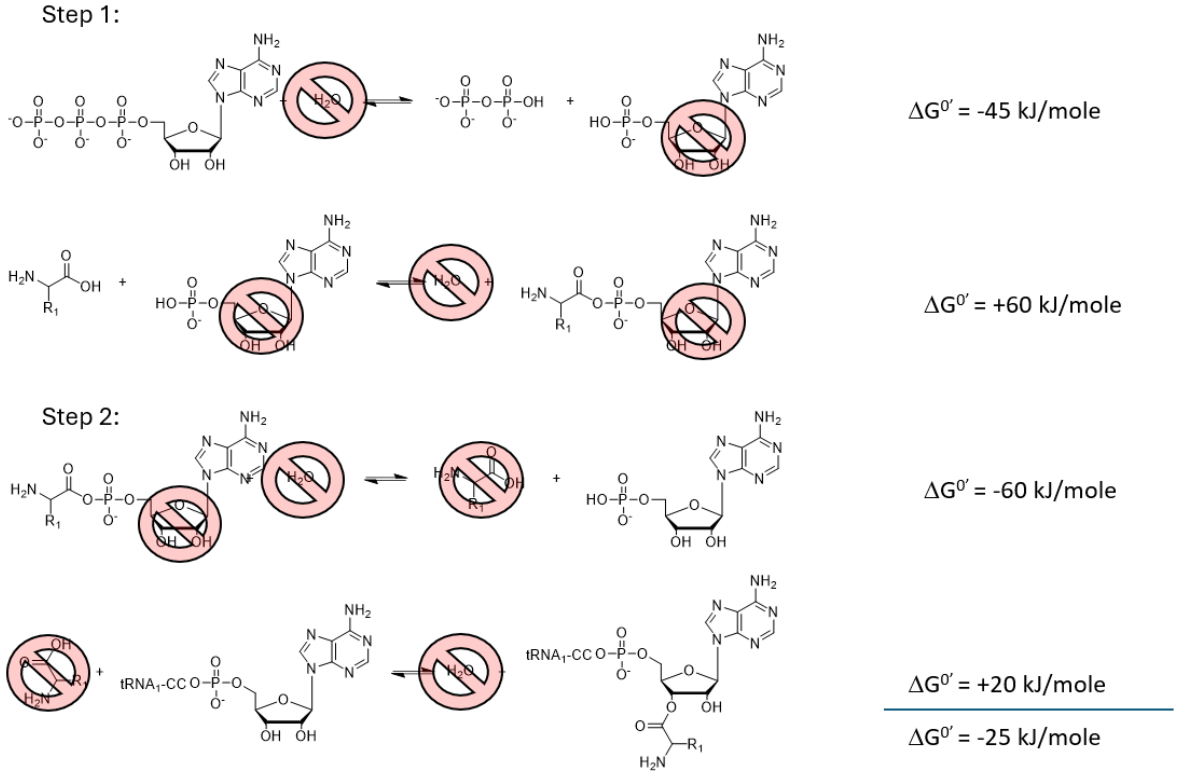
So by the free energy analysis that we did here, we can see that the peptide bond formation reaction is favorable in biology due to the fact that the amino acids are linked through an ester bond to tRNA. It is a real possibility that in the RNA world, the anticodon loop came much later and that initially more random polymers were formed that were initially linked to something like tRNA that led to enhanced favorability for the reaction. However, we have just "kicked the can down the road" here a little. While now we know that this reaction is favorable, we had to form the tRNA-linked amino acid. We know that this is an ester linkage, therefore we now need +20 kJ/mole to be available. This is where the concept of coupling reactions comes into play and it is essential to understand this to have a fundamental understanding of metabolism. All reactions must be favorable therefore we need an energy currency that can be used to push unfavorable reactions forward. ATP is the prime energy currency of biology (in limited cases GTP is also used, but more on that later in the semester….maybe). Unfortunately how it serves this purpose is often very poorly understood in part due to the way these reactions are typically drawn as shown here:



This format often gives students the impression that the energy from ATP hydrolysis to give AMP and diphosphate is somehow transferred to the unfavorable reaction driving it forward. This is not what is happening here. When looking at these reactions, to really understand what is happening, we have to consider it from both an organic and thermodynamic perspective. From the organic perspective this is the arrow pushing analysis that describes what takes place in the reaction:



Here we see the reaction catalyzed by the enzyme aminoacyl tRNA synthetase which is responsible for loading amino acids onto tRNA. First, we have an attack of the amino acid O- on the first phosphate unit of ATP. This results in the loss of diphosphate and the formation of a mixed anhydride intermediate. This acylated phosphate (amino acid linked to AMP) is a key intermediate that is often found in metabolism. The AMP provides an excellent leaving group for the next reaction catalyzed by the aminoacyl tRNA synthetase where the 3’ or 2’ OH of the appropriate tRNA will attack the carbonyl displacing AMP and forming the final amino acid charged tRNA. Here is the analysis of the thermodynamics of this reaction:



The In the thermodynamic analysis what we see is that ATP is hydrolyzed to give diphosphate and AMP and we know that this is worth -45 kJ/mole (plenty for the +20 kJ/mole needed). In this reaction with ATP we have the formation of a mixed anhydride (note hydrolysis of carbon anhydride -90 kJ/mole, phosphoanhydride -45 kJ/mole, phosphoester -15 kj/mole, and the mixed anhydride is between the carbon and phosphoanhydrides at -60 kJ/mole). This first step is not favorable as it has an over free energy of +15 kJ/mole. However, this intermediate is not released from the enzyme. Instead, the tRNA attacks the mixed anhydride so the energy put in for the formation of the mixed anhdyride we get back at -60 kJ/mole. The formation of the ester as we already noted is +20 kJ/mole, and the overall free energy is -25 kJ/mole. What does this suggest about the equilibrium of this reaction and would you expect there to be mostly loaded tRNAs in a cell or would you expect mostly free tRNA assuming an abundance of amino acids?