DNA to Proteins

Our body is merely the mechanism our genes use to reproduce.

The above statement is very profound and is worthy of some careful thought. Our genes determine the characteristics of our bodies so that our bodies can reproduce successfully and pass those genes on to future generations. Our bodies are discarded, but copies of our genes can continue. Genes that fail to produce successful bodies are lost. Once we have reproduced and raised our offspring, our genes have little interest in our bodies. Grandparents can be useful, but not nearly as much as parents so natural selection aims to keep younger people healthy.

Commonalities between the genes of siblings and their parents enhance familial ties, but you can think of this as a conspiracy of those genes to enhance their chance of survival. Twins sense a specially strong bond because the share a complete set of genes. Ant colonies work as an incredibly efficient team with individual survival subservient to the survival of the colony.

DNA (deoxyribonucleic acid) is the way life's recipe is stored. It consists of a stack of nucleotide pairs used in groups of 3 to specify a specific amino acid. Each DNA sequence that is used together to make the amino acid sequence in a protein is considered a gene. One gene for each type of protein, thousands of genes per chromosome, and 23 chromosome pairs in humans.

Our oxytocin is a very small protein called a polypeptide and was made from 9 amino acids:

gly-leu-pro-cys-asn-gln-ile-tyr-cys

Here the abbreviations for the amino acids shown on page 7 of your molecular diagram handout are used. When we built this, we saw that the sulfur atoms in the two copies of cysteine (cys) find each other and bond to form a loop.

The DNA code is made up of flat nucleotides: adenine (A), guanine (G), cytosine (C), and thymine (T). These are shown on page 9 of our molecular diagram handout. These are strung together in pairs to form a long double-helix structure with A always matched with T and G always matched with C.

The DNA code for oxytocin uses 3 nucleotides for each amino acid and is paired to form its famous double-helix structure. The section for making oxytocin looks like:

ACG-ATG-TAT-GTT-TTG-ACG-GGA-GAC-CCC
TGC-TAC-ATA-CAA-AAC-TGC-CCT-CTG-GGG

Dashes are put between groups of three nucleotides our convenience, nature just rams them together. Each group of three corresponds to an amino acid, but has start and stop groups that allow an unambiguous readout when the amino acids are interpreted.

When DNA is replicated, these strands are separated and a new match is created for each, A with T and C with G. Special proteins perform the unzipping and matching, making two exact copies of the original strand. Other proteins are used to check and repair the DNA. Replication of the DNA only happens when cells divide.

Using the DNA information for the manufacture of proteins like oxytocin involves copying the DNA section shown above. The DNA is in the cell nucleus, but the oxytocin and other proteins are made in the surrounding cytoplasm. So the DNA code is copied to an intermediate temporary code transfer-RNA (t-RNA) except that thymine (T) is replaced by a different nucleotide, uracil (U). For oxytocin, the t-RNA then looks like

UGC-UAC-AUA-CAA-AAC-UGU-CCU-CUG-GGG

These groups of 3 nucleotides are called codens and the following table shows how they correspond to different amino acids and start/stop instructions. Note that there are redundancies.

First nucleotide	Second nucleotide								Third nucleotid
	U		C		A		G		
U	UUU	phenylalanine	UCU	serine	UAU	tyrosine	UGU	cysteine	U
	UUC		UCC		UAC		UGC		C
	UUA	leucine	UCA		UAA	stop	UGA	stop	A
	UUG		UCG		UAG		UGG	tryptophan	G
C	CUU	leucine	CCU	proline	CAU	histidine	CGU	arginine	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	glutamine	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	isoleucine	ACU	threonine	AAU	asparagine	AGU	serine	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	lysine	AGA	arginine	A
	AUG	methionine (start)	ACG		AAG		AGG		G
G	GUU	valine	GCU	alanine	GAU	aspartic acid	GGU	glycine	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	glutamic acid	GGA		A
	GUG		GCG		GAG		GGG		G

Using this chart, we can convert the t-RNA sequence above to the amino acid sequence:

cys-tyr-ile-gln-asn-cys-pro-leu-gly

Scientists have recently synthesized DNA sequences for new enzymes, inserted them into the DNA of animals, or bacteria and made modified life. This genetic engineering will become easier and must be controlled to prevent the creation of dangerous life forms.