>> Dude: Hey Sally.

>> Sally: Hmmm?

>> Dude: How come Buddy is green and I'm not?

>> Sally: Well Dude.

Your parents aren't green and I'm guessing that Buddy's parent was.

>> Dude: That's not what I mean...like, what makes him green?

>> Sally: Envy?

>> Dude: Fine, I'll figure it out myself.

Alright! Here it is.

Hey Sally, check this out.

All of Buddy's DNA is online.

Now all I have to do is find the part that makes him green.

What a nightmare.

It all looks the same.

Just Gs, As, Ts, and Cs.

How am I suppose to know what's what?

Someone needs to write a manual.

>> Sally: We are writing one, but it's a work in progress.

Why is Buddy green?

Well, it's probably because he makes a protein or two that turn him green.

We can figure that out, because sequences that code for proteins have a few things in common.

>> Dude: Oh I remember now.

They all start with ATG then call out for amino acids using three bases at a time.

>> Sally: Those 3 bases are called triplet codons.

Here's a look up table with the genetic code for Buddy.

It tells you which amino acids are encoded by each triplet codon.

>> Dude: OK I got it now.

Here's an ATG for the methionine start, then I see a TAT, that translates to tyrosine, then TAG that will make a ...stop?

That's only two amino acids.

That seems too short to do anything in the cell.

>> Sally: You've just figured out something really important.

Not every ORF will make a protein.

As a rule of thumb, people consider 100 amino acids the minimum to produce a protein, even though there are lots of proteins we know that contain fewer.

>> Dude: Did you say ORF or did you just burp?

>> Sally: didn't burp!

ORF stands for open reading frame.

It's a register for reading the section of DNA that starts with ATG, uses 3 letters at a time to call for amino acids and then uses one or more of the stop codons to close the reading.

>> Dude: So in this DNA sequence there are 3 reading frames that I have to look through?

That will take me forever.

>> Sally: Actually it will take you forever times two since there are 6 reading frames.

Remember DNA is double stranded and the bottom strand, the one you're not seeing here, might also have ORFs that make proteins.

>> Dude: Please tell me there are computer programs to help me with this.

>> Sally: There are lots of good ones, but once you find all the ORFs what will you do?

>> Dude: Yeah...then what?

How will I know which of the ORFs really make proteins?

And what if I've missed a bunch?

Maybe some proteins that turn Buddy green are only a few amino acids long.

>> Sally: Nice!

You've got the start of a really interesting project here.

But remember that discovering important proteins is only part of your job.

If you really want to engineer biology you'll have to refine the genetic material so you and others can build with it.

Before Buddy wakes up, can I show you one great example?