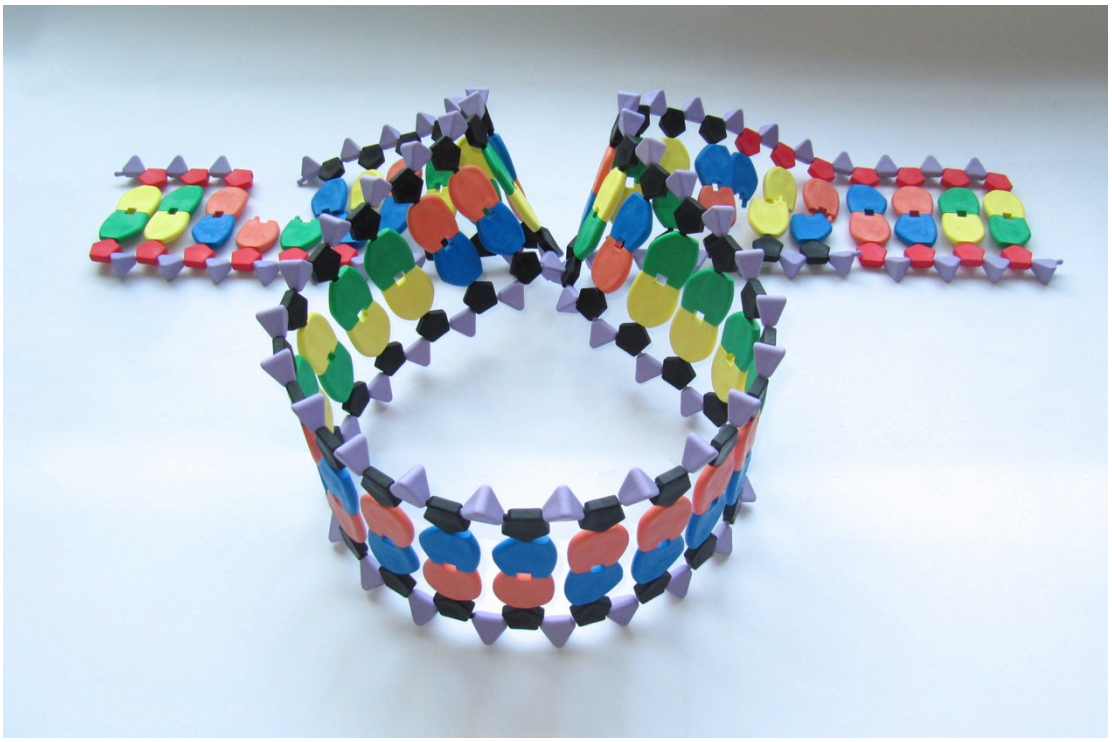
Discovering DNA

Molymod® miniDNA®

HIV – a Genetic Hijacker

Instructions and teacher's guide

Cat no MDNA-HIV-381



We'd love to hear any feedback, comments or questions you have!

Post: Discovering DNA Ltd, PO Box 280 Hertford, SG13 9DG
email: info@discoveringdna.com
tel: 01992 410 140
fax: 01992 410 106

Key: Green = Guanine Yellow = Cytosine Dark Blue = Adenine Orange = Thymine Light Blue = Uracil

Molymod® miniDNA®

HIV – a Genetic Hijacker

Cat no MDNA-HIV-381

Contents:

Black sugar	68	Guanine (Green)	33
Red sugar	20	Adenine (Dark Blue)	31
Claret sugar	39	Thymine (Orange)	23
Phosphate (Purple)	127	Uracil (Light Blue)	11
Cytosine (Yellow)	29		

Introduction

Human immunodeficiency virus (HIV) is responsible for acquired immunodeficiency syndrome (AIDS). HIV is so dangerous because it infects immune system cells so it causes a breakdown of the very system intended to fight viral infections.

There are two types of the virus, HIV-1 and HIV-2. HIV-2 is mostly localized to West Africa and less infectious than HIV-1. HIV-1 is the most pathogenic and widespread form that affects the global population.

HIV has an interesting life cycle and understanding the molecular biology of HIV has enabled scientists to develop treatments based on blocking or inhibiting virus specific processes.

HIV life cycle

HIV is a retrovirus. That means HIV has as an RNA genome that must be copied into DNA (cDNA) before it can be integrated into the host cell genome. To do this, HIV hijacks a transfer RNA (tRNA) molecule from the host cells protein translation machinery to use as a primer to begin the reverse transcription process. The enzyme reverse transcriptase uses this primer to copy the RNA genome to DNA.

Reverse transcriptase

Reverse transcriptase reverses the usual transcription process from DNA to RNA. Thus, the enzyme is widely used in molecular biology research to create cDNA libraries. A cDNA library is a DNA copy of the mRNA expressed in a particular cell type or tissue so it can be studied more easily.

Once integrated, the viral DNA (now called a provirus) is replicated along with the host cell's DNA during cell division. When the infected immune system cell is activated by another infection, the proviral DNA acts as a template for viral RNA transcription producing messenger RNA (mRNA). Some of the viral mRNAs are translated into viral proteins and the full length mRNA becomes the genome for new viral particles.

HIV medicines

Many medicines have now been developed to treat HIV although none offer a cure. HIV is able to mutate very quickly which makes it a difficult target for medicine development. However, combinations of medicines acting on different components of the HIV life cycle (eg reverse transcriptase, integrase) are used to overcome this.

HIV genome

In this set we use the real HIV sequence that has been edited to make it more manageable. The actual HIV genome is 9749 nucleotides long and has 9 genes encoding 19 proteins.

Learning outcomes

Make a model of the HIV virus RNA genome using the real sequence. See how the virus makes a double stranded DNA copy of itself and how this is integrated into the host genome. A simple or advanced option provided - the detailed reverse transcriptase steps can be omitted if desired.

By the end of the session, your students will have learnt:

- HIV
- Viruses
- Retroviruses
- Genetic code
- Translation
- tRNA
- Transcription
- DNA replication
- Reverse transcriptase
- Restriction enzymes
- Complementary base pairing
- Drug discovery

Time requirements:

Preparation - First time will take about 20 minutes but after this preparation is no more than 5 minutes as you can keep components assembled in storage box for future use. **Lesson -** 30 minutes is sufficient to carry out the activity with more time to explain the stages and analyse the outcomes as required.

Key: Green = Guanine Yellow = Cytosine Dark Blue = Adenine Orange = Thymine
Light Blue = Uracil

Preparation

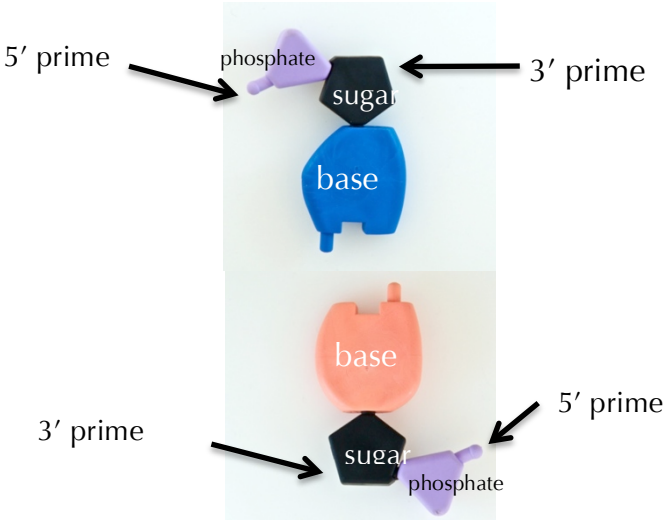
You only need to do this once as the bases can be kept for future lessons!

1 Assemble cDNA bases (black sugars)

DNA bases comprise of a sugar (black), a phosphate (purple) and a base (yellow, green, dark blue or orange).

Attach the purple phosphate to the black sugar by pushing the bent knob from the sugar into the hole in the purple phosphate. Make sure you add the sugar to the 5′ prime end as shown so the knob still sticks out of the purple phosphate not out of the sugar.

Push the coloured base (green, orange, dark blue or yellow) onto the straight knob on the sugar.



Assemble these bases with black sugars:

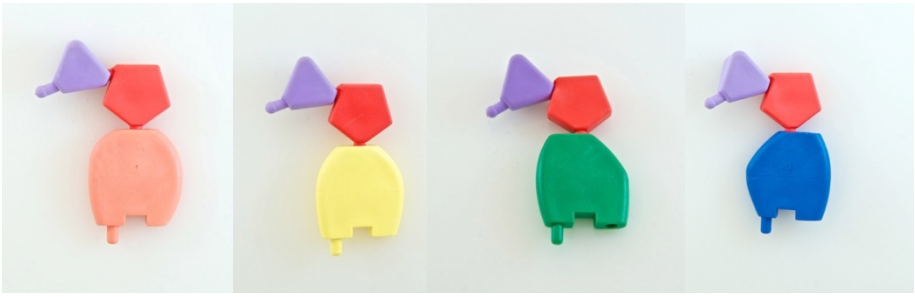
Cytosine (Yellow)	x 16	Adenine (Dark Blue)	x 18
Guanine (Green)	x 16	Thymine (Orange)	x 18

Helpfully, C, G, A, and T are embossed on the edge of the base.

2 Assemble host DNA bases (red sugars)

Assemble the following using red sugars:

Cytosine (Yellow)	x 5	Adenine (Dark Blue)	x 5
Guanine (Green)	x 5	Thymine (Orange)	x 5

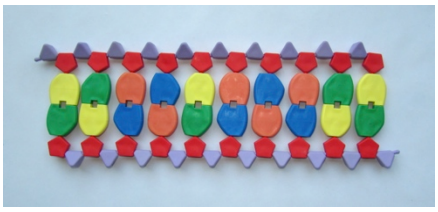


3 Assemble host DNA sequence (red sugars)

Put the red bases in this sequence:

5′- CGTACTATGC -3′
3′- GCATGATACG-5′

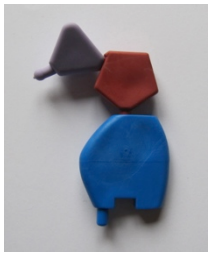
Key: Green = Guanine Yellow = Cytosine Dark Blue = Adenine Orange = Thymine
Light Blue = Uracil



4 Assemble transfer RNA (tRNA) bases (claret sugars)

Assemble the following using claret sugars:

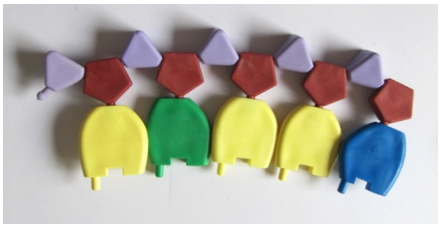
Cytosine (Yellow)	x 3	Adenine (Dark Blue)	x 1
Guanine (Green)	x 1		


5 Assemble transfer RNA (tRNA)

An edited version of the tRNA is used. See our Transfer RNA model (MDNA-TRNA-228) for the complete sequence. Put the claret bases in this sequence:

5′ - CGCCA - 3′

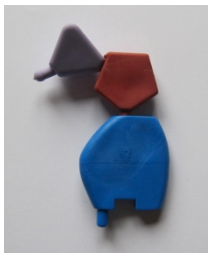


6 Assemble HIV RNA genome bases (claret sugars)

For HIV's RNA genome we use Uracil in place of Thymidine. Uracil is light blue.

Assemble the following using claret sugars:

Cytosine (Yellow)	x 5	Adenine (Dark Blue)	x 7
Guanine (Green)	x 11	Uracil (Light Blue)	x 11

After completion of the lesson

Separate the RNA bases (with claret sugars) and keep them intact in the storage boxes for use next time. Keep the DNA bases (red and black sugars) intact for the next time. Keep the HIV genome intact.

Preparation for subsequent uses (about 5 minutes)

Sort out the RNA bases (with claret sugars) by colour (yellow, green, dark blue or light blue). Give out DNA bases (red and black sugars). Then you are ready to go!

Key: Green = Guanine Yellow = Cytosine Dark Blue = Adenine Orange = Thymine
Light Blue = Uracil

HIV - a Genetic Hijacker – teacher's & student's notes

Lesson tips

Important - ensure students have bases correctly orientated so 5′ and 3′ are opposite each other when bases are paired (shown on page 2) & don't let your students take the bases apart!

Remember in base pairing the following match:

T/U pairs with A (orange/light blue with dark blue)
C pairs with G (yellow with green)

Lesson plan

- Students build HIV genome and host DNA to infect
- They copy HIV into DNA with reverse transcriptase
- Follow either short of long reverse transcriptase steps
- Model how integrase processes the HIV cDNA and cuts open host DNA
- Students integrate the HIV into the host genome

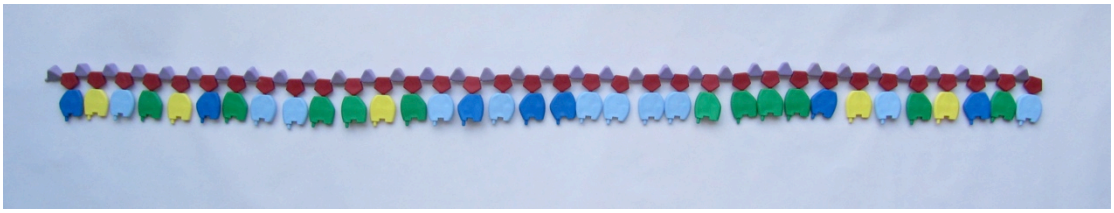
1 Make RNA HIV genome

The real HIV genome is 9749 bases long and is represented as follows:

- Integrase - first bases of integrase represent structural and enzymatic genes
- Long terminal repeats (LTRs) are important during viral integration.
- Primer binding site (PBS) where transfer RNA binds to prime first strand synthesis.
- TATA box is the start point of transcription.
- Polypurine tract (PPT) forms the primer for second strand synthesis.

Students should first make the HIV RNA genome sequence using claret bases as follows:

5′	ACUGCAGU	UGGCG	UAUAA	UUUU	GGGG	ACUGCAGU 3′
	5′LTR	PBS	TATA box	Integrase	PPT	3′LTR



2 Reverse transcriptase

After entry to host cell, HIV immediately copies its' RNA genome to a double stranded DNA copy (cDNA). This copying is carried out, during transport to the cell nucleus, by an amazing enzyme called reverse transcriptase. As the name suggests, it does the reverse of transcription (DNA copied to RNA) as it copies from RNA to DNA. Reverse transcriptase was discovered by Howard Temin and David Baltimore, who shared the Nobel Prize in Physiology or Medicine for their discovery in 1975.

Reverse transcriptase as a drug target
The first HIV medicine developed was AZT (azidothymidine). AZT is a DNA chain terminator like those used in DNA sequencing reactions. Chain terminators are DNA bases that lack a 3′ OH group so they halt DNA replication (or sequencing) when added by the DNA polymerase enzyme.

AZT does the same thing but to reverse transcriptase. Although not specific for HIV DNA replication, because HIV replicates more quickly than human cells, AZT preferentially blocks DNA replication by HIV.

Reverse transcriptase hijacks a host transfer RNA
As with DNA polymerase, reverse transcriptase needs a single stranded primer to start copying the RNA template. This primer is provided by a host transfer RNA (tRNA) molecule that

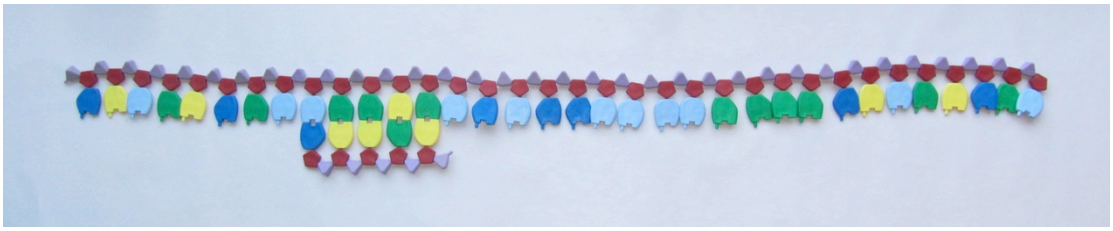
inadvertently helps HIV copy itself. It's interesting to think how this evolved as without this tRNA, HIV would not be able to carry out reverse transcription.

The tRNA is tRNA(Lys,3) which is for the anticodon UUU that encodes Lysine. The tRNA binds to the HIV genome as 18 of its' 3' bases are complementary to the primer binding site (PBS) on the viral DNA. This RNA region acts as a single stranded RNA primer for reverse transcriptase that then allows the viral RNA to be copied to the first strand of DNA.

3 Reverse transcriptase steps

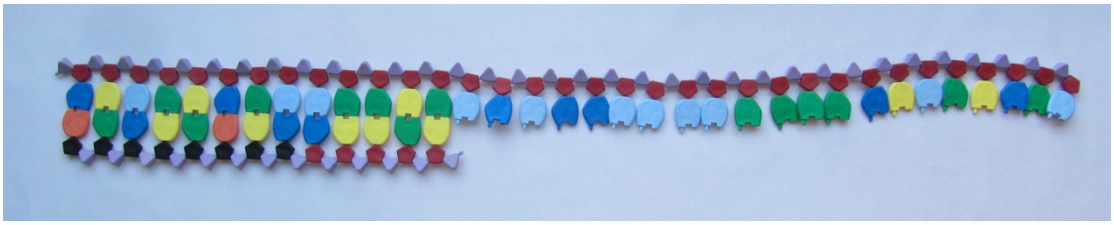
1 Students attach the tRNA to the HIV genome (vRNA) where it binds by complementary base pairing.

vRNA 5′ ACUGCAGU UGGCG UAUAA UUUU GGGG ACUGCAGU 3′
ACCGC---5′
tRNA = RNA primer



2 Reverse transcriptase copies in a 3-5 direction from the 3′ of the tRNA primer. Students build up the DNA copy using bases with black sugars (the cDNA is shown shaded):

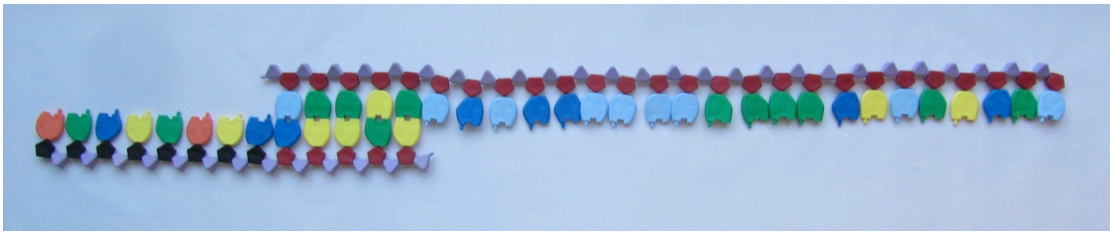
vRNA 5′ ACUGCAGU UGGCG UAUAA UUUU GGGG ACUGCAGU 3′
cDNA 3′ TGACGTCA ACCGC---5′
tRNA



Note - this gets quite complex so if you want to show how reverse transcriptase works, you can jump to step 13 and make the double stranded DNA genome shown.

3 After copying the first part of the viral genome, the vRNA that has been copied is destroyed by RNAse H. RNAse H is part of the HIV reverse transcriptase enzyme and is an important drug target due to its essential role in HIV replication. Students should remove the previously copied RNA:

vRNA 5′ UGGCG UAUAA UUUU GGGG ACUGCAGU 3′
cDNA 3′ TGACGTCA ACCGC---5′
tRNA

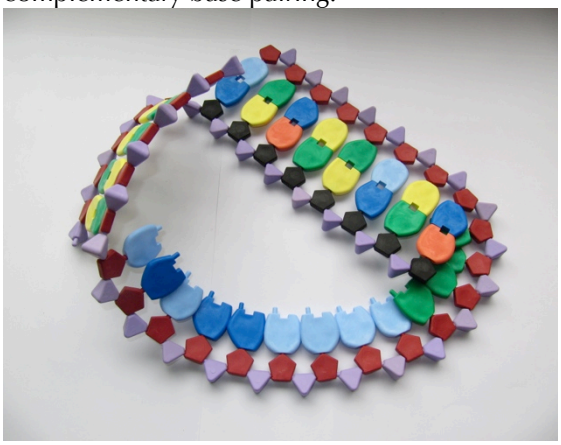


Key: Green = Guanine Yellow = Cytosine Dark Blue = Adenine Orange = Thymine
Light Blue = Uracil

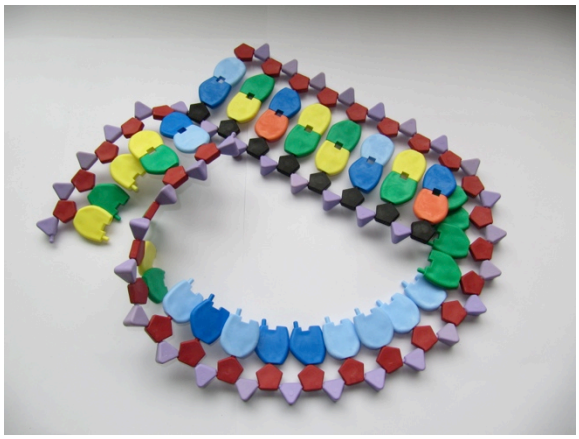
4 The tRNA and cDNA then jumps to the 3′ end of the viral genome and the first partial cDNA strand acts as a primer to copy the rest of the strand. It does this by forming a circle so the regions of homology in the LTRs meet.

4a Bring opposite ends together.

4b Homologous ends meet & bind by complementary base pairing.



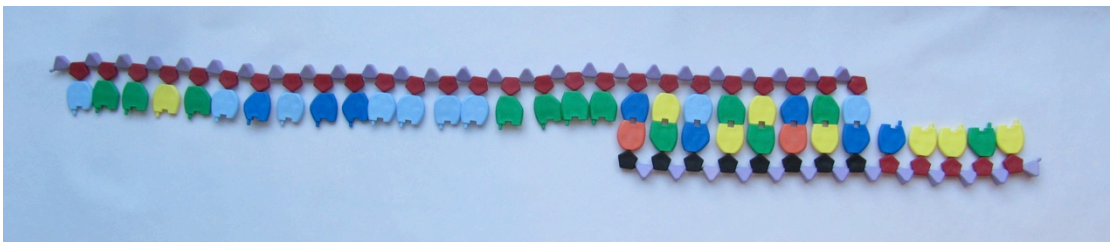
4c Separate the binding by tRNA primer.



4d Unwind and straighten the model.

5 This circularization has allowed the opposite strands to pair and causes the RNA primer and cDNA copy to swap to the opposite end of the strand:

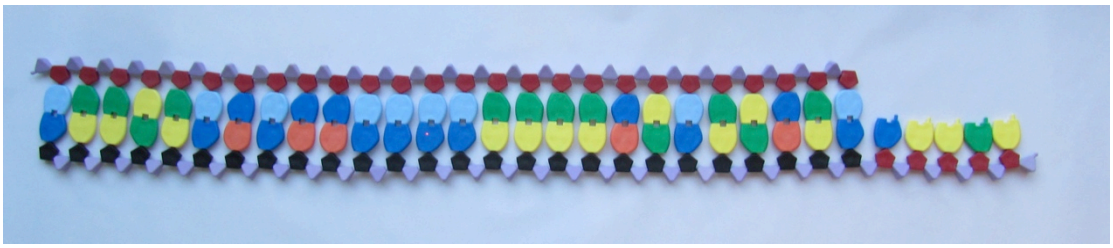
vRNA 5′ UGGCG UAUAA UUUU GGGG ACUGCAGU 3′§
cDNA 3′ TGACGTCA ACCGC---5′
tRNA



6 Reverse transcriptase fills in rest of first strand in 3-5 direction. Students should do this by adding black DNA sugars:

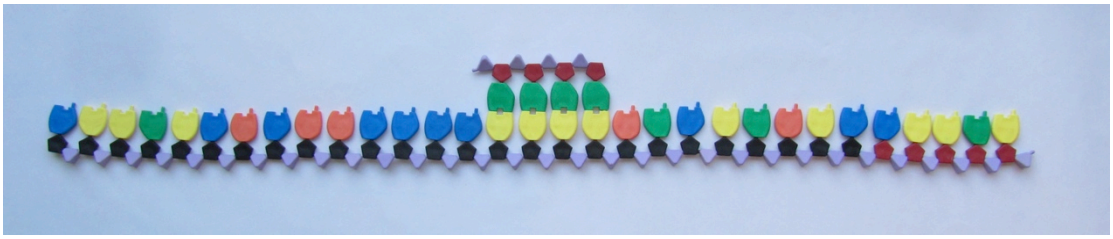
vRNA 5′ UGGCG UAUAA UUUU GGGG ACUGCAGU 3′
cDNA 3′ ACCGC ATATT AAAA CCCC TGACGTCA ACCGC---5′
tRNA

Key: Green = Guanine Yellow = Cytosine Dark Blue = Adenine Orange = Thymine
Light Blue = Uracil



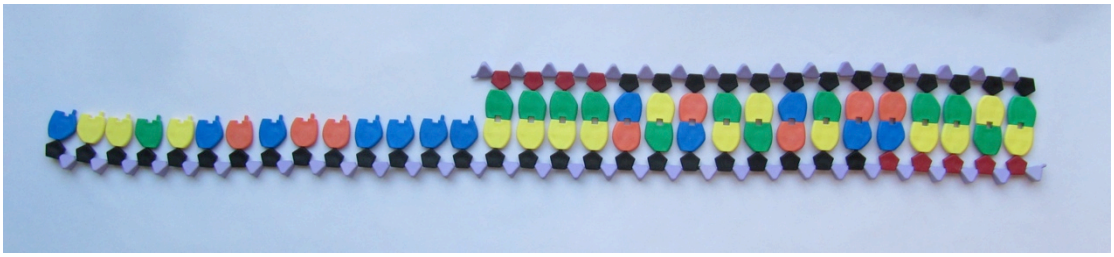
7 RNAse H degrades RNA that has been copied except for PPT which acts as primer for second strand. Students need to remove all claret bases apart from the PPT:

vRNA 5′ GGGG 3′
cDNA 3′ ACCGC ATATT AAAA CCCC TGACGTCA ACCGC---5′
tRNA



8 Second strand copied in 3-5 direction by reverse transcriptase. Students should fill in the first part of the strand with black DNA bases:

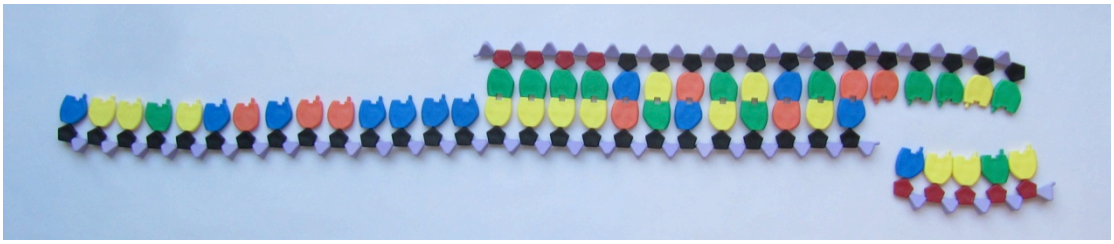
vRNA 5′ GGGG ACTGCAGT TGGCG 3′
cDNA 3′ ACCGC ATATT AAAA CCCC TGACGTCA ACCGC 5′
tRNA



9 The transfer RNA tRNA (Lys,3) is released:

vRNA 5′ GGGG ACTGCAGT TGGCG 3′
cDNA 3′ ACCGC ATATT AAAA CCCC TGACGTCA 5′

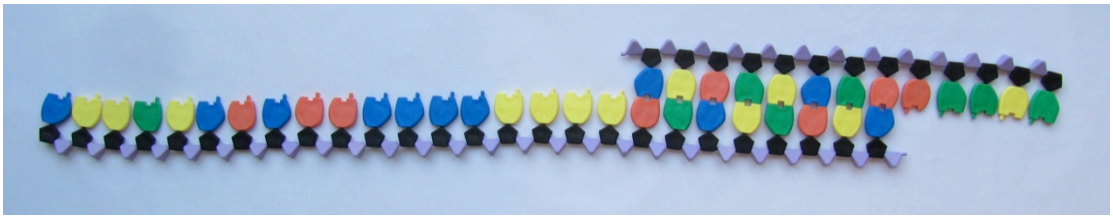
ACCGC---5′
tRNA released



Key: Green = Guanine Yellow = Cytosine Dark Blue = Adenine Orange = Thymine
Light Blue = Uracil

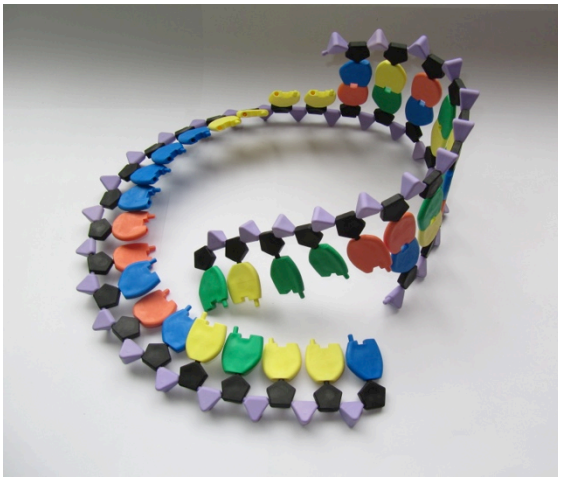
10 The last remaining part of the viral RNA genome PPT is degraded by RNAse H:

vRNA 5′ ACTGCAGT TGGCG 3′
cDNA 3′ ACCGC ATATT AAAA CCCC TGACGTCA 5′



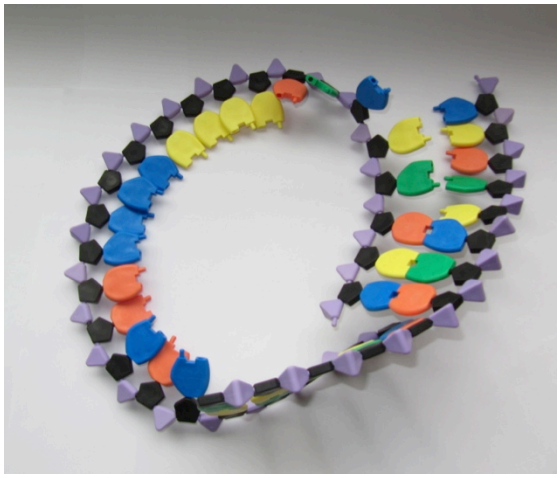
11 Once again, the copied DNA circularizes into a loop allowing the cDNA to jump to the other end of the virus. The LTRs form a homologous bond through complementary base pairing so the strands can swap.

11a Curl model so opposite ends meet:



11b Bind PBS to homologous end of virus:

11c Separate previously bound LTRs.

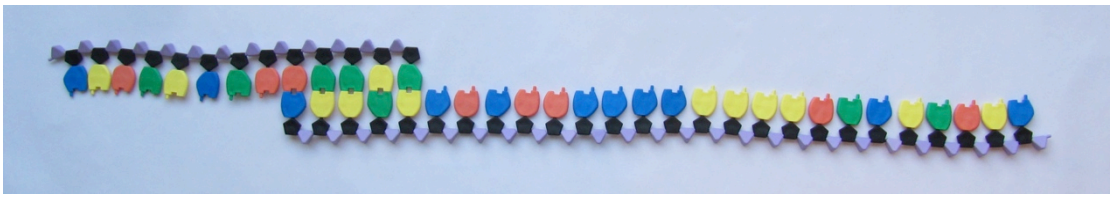


11d Now viral strands just joined at the PBS. Uncurl and straighten model.

12 The copied DNA should now be at the other end of the cDNA to act as a primer for the rest of the second strand synthesis by reverse transcriptase. If looks like this, well done! If not, try again.

cDNA 5′ ACTGCAGT TGGCG 3′
cDNA 3′ ACCGC ATATT AAAA CCCC TGACGTCA 5′

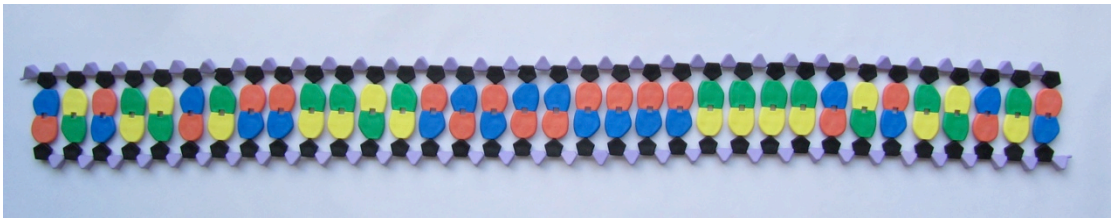
Key: Green = Guanine Yellow = Cytosine Dark Blue = Adenine Orange = Thymine
Light Blue = Uracil



Note - continue from here if you are just showing reverse transcriptase in action.

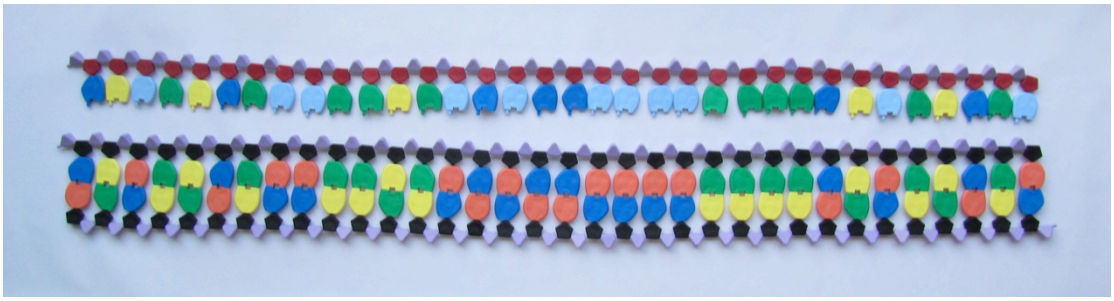
13 Reverse transcriptase acts as DNA dependent DNA polymerase and copies the DNA in 5'- 3' direction using the opposite strand as templates. Fill in the DNA using bases with black sugars:

cDNA 5′ ACTGCAGT TGGCG TATAA TTTT GGGG ACTGCAGT 3′
cDNA 3′ TGACGTCA ACCGC ATATT AAAA CCCC TGACGTCA 5′



14 Compare it to the original RNA genome and check if you got it right! Well done if you did! Have another go if you didn't.

5′ ACUGCAGU UGGCG UAUAA UUUU GGGG ACUGCAGU 3′
5′LTR PBS TATA box integrase PPT 3′LTR



3 3′ processing

This is the first part of how HIV attacks the host cell's DNA genome and forces its' way in. It is surprisingly violent.

First, integrase binds to the double stranded DNA (cDNA) in cytoplasm of the host cell and forms a large complex of viral and host proteins called the preintegration compex (PIC). The PIC is eventually transported through into the host cell nucleus through a nuclear pore complex.

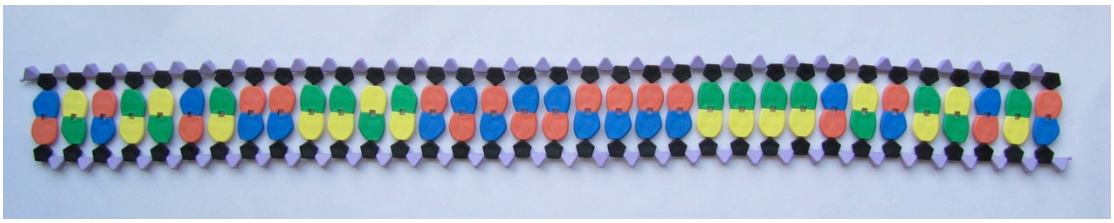
The ends of retroviruses are important for viral integration. To prepare it for an attack on host DNA, the viral cDNA undergoes 3′ processing whist it is being transported through the cytoplasm of the host cell on the way to the nucleus.

During 3′ processing, integrase cuts off two conserved GT bases from the 3′ blunt end of each cDNA strand so end up with what look like "sticky ends" as if cut with restriction enzyme. The remaining 3′ OH group is reactive and will attack a strand of host DNA.

1 Start with the cDNA from above:

5′ ACTGCAGT TGGCG TATAA TTTT GGGG ACTGCAGT 3′
3′ TGACGTCA ACCGC ATATT AAAA CCCC TGACGTCA 5′

Key: Green = Guanine Yellow = Cytosine Dark Blue = Adenine Orange = Thymine
Light Blue = Uracil

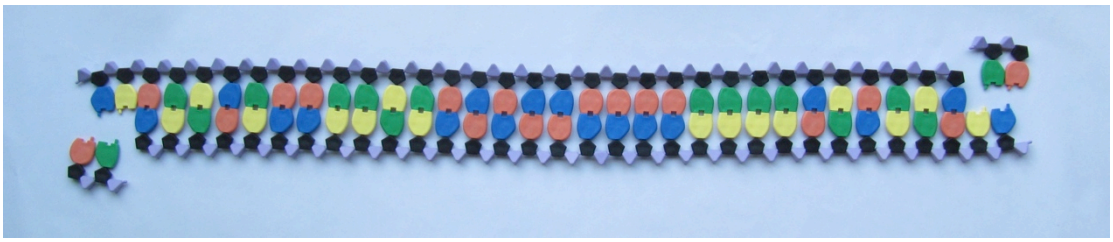


2 Remove the GT from the 3′ end of the HIV genome leaving an overhang:

5′-GT-3′

5′ ACTGCAGT TGGCG TATAA TTTT GGGG ACTGCA 3′
3′ ACGTCA ACCGC ATATT AAAA CCCC TGACGTCA 5′

3′-TG-5′

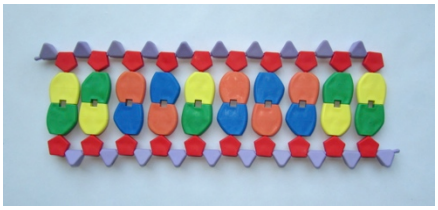


Next, the DNA strand transfer reaction occurs as the HIV cDNA is inserted into the host cell DNA.

4 DNA strand transfer reaction

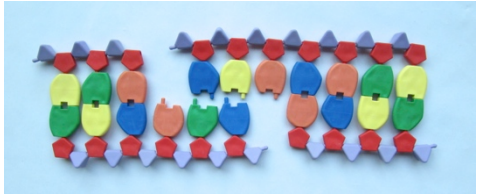
Inside the nucleus a host protein (LEDGF) is hijacked to attach the PIC to the host DNA. LEDGF plays an important role because it attaches HIV to an active transcription site – so HIV is likely to be highly expressed and thus makes lots more HIV particles. This is known because in mice that have been genetically engineered to have LEDGF knocked out, HIV expression is significantly reduced. Thus, LEDGF could be a target for medicine development.

1 Make the host DNA sequence with bases with red sugars:



2 Integrase cuts the host DNA in a staggered way so you get what look like overhanging sticky ends created by a restriction enzyme. However, unlike restriction enzymes, the sequence where integrase cuts is **not** sequence specific.

5′- CGT ACTATGC -3′
3′- GCATGA TACG-5′

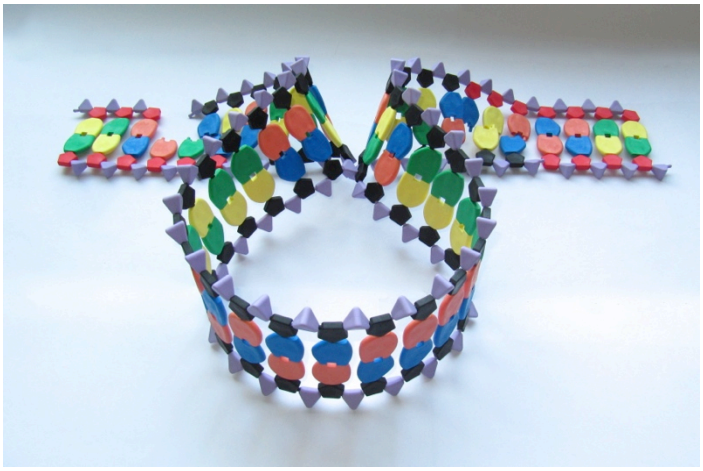


3 The DNA strand transfer happens when integrase launches an attack with the 3′ ends of cDNA on the host DNA. The 3′-OH groups of the viral DNA attach themselves to the host DNA where the cut has been made through a sugar-phosphate bond. **Thus, only the sugar-phosphate backbone is joined not the bases.**

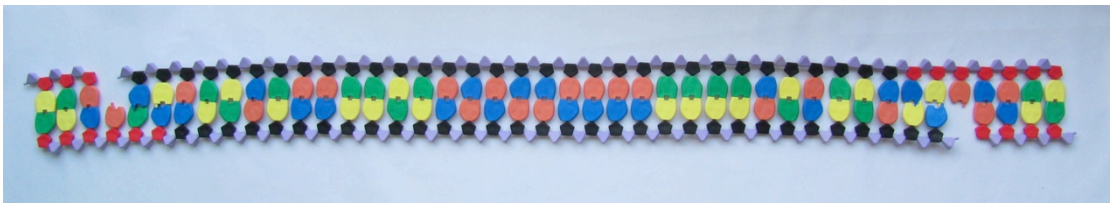
Key: Green = Guanine Yellow = Cytosine Dark Blue = Adenine Orange = Thymine
Light Blue = Uracil

5′- CGT ACTGCAGT--- ACTGCA- ACTATGC -3′
3′- GCATGA -ACGTCA--- TGACGTCA TACG -5′

4 The moment of attack!

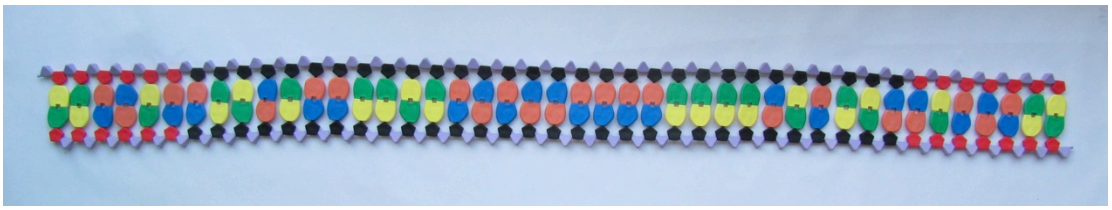


5 This leaves a one base pair gap and two bases unmatched with the opposite strands.



6 The cellular repair mechanism fills the gap and excises the mismatched bases and adds the correct host bases. Result is virus integrated (latent) into the host cell DNA. It is now called proviral DNA. Thus, HIV is now integrated into the genome:

5′- TTCGT ACTGCAGT--- ACTGCAGT ATGCT -3′
3′- AAGCA TGACGTCA--- TGACGTCA TACGA -5′



7 This is how your virus and host should look. The cell's DNA has now been successfully attacked.

Where does HIV integrate?

The human genome project and, more recently, next generation sequencing, have enabled a more detailed search for integration site sequences. They are found to be in active transcription units in genes but not in specific sequences. It may be where the genome has particular features that cannot be seen from DNA sequence alone – such where DNA is more exposed and where the host protein LEDGF binds the PIC to the host DNA.

The integration step is a major target for drug inhibitors as integrase is not an enzyme that occurs in human cells. For example, integrase inhibitors have been developed to compete for Mg2+ ions needed by integrase.

Viral genome expression

When the cDNA is transcribed into mRNA, the whole transcript is the HIV genome and various viral proteins are produced by shorter mRNAs. Viral proteins are translated and two copies of the RNA genome added to each viral particle. The complete mRNA is the HIV genome.