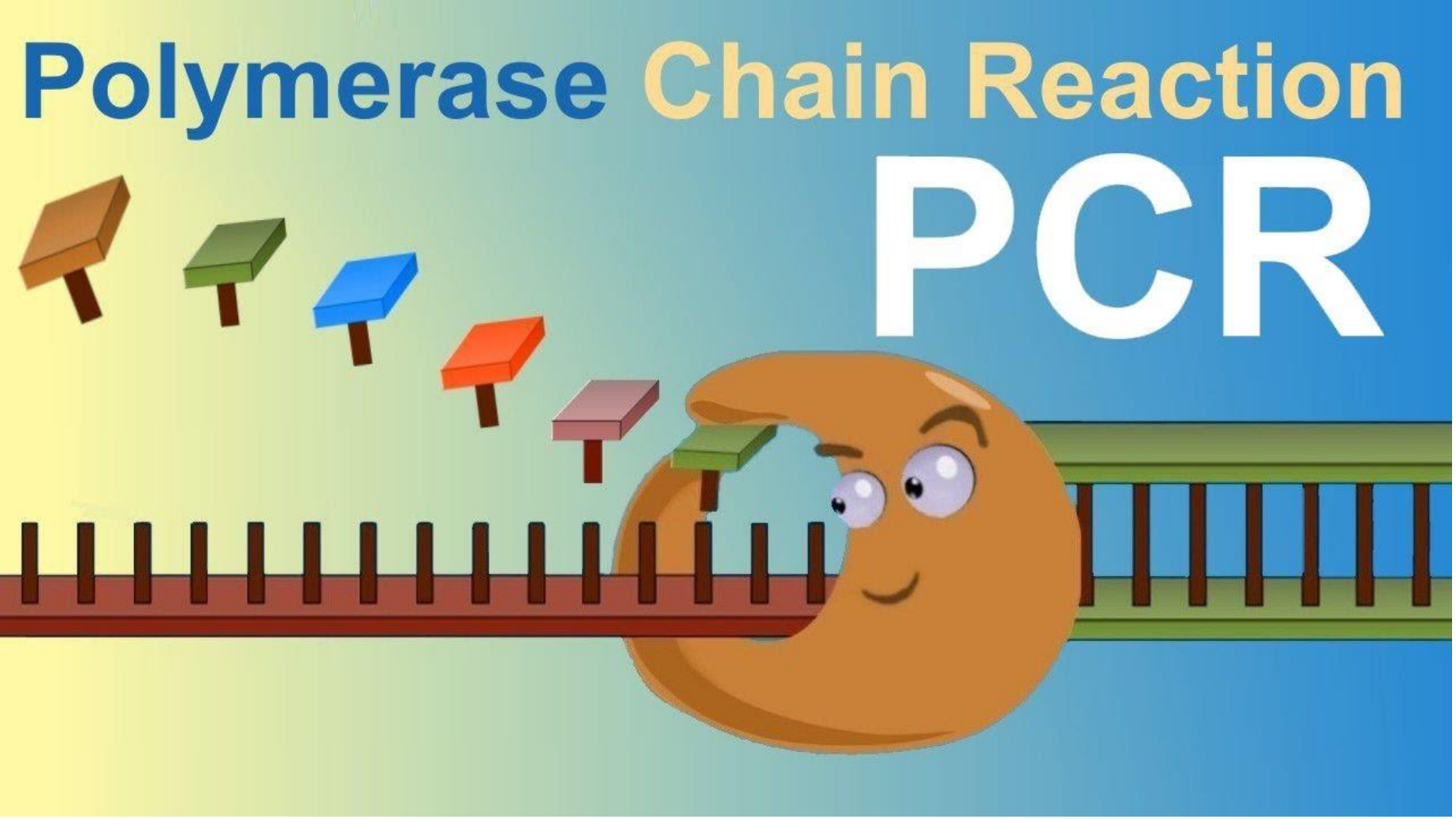

Polymerase Chain Reaction
PCR

Definitions of PCR

Definition-1

- A technique for amplifying DNA sequences in vitro by separating the DNA into two strands and incubating it with oligonucleotide primers and DNA polymerase.

Definition -2

- A biochemical technique used in Molecular Biology to amplify a specific fragment of target DNA.

Origin of Polymerase chain reaction (PCR)

In vitro DNA synthesis

- PCR was discovered by Kary B. Mulis in 1983 of Cetus Corporation, a Biotech company in California, USA.
- He won the Nobel Prize for Chemistry in 1993 for 'contributions to the developments of methods within DNA-based chemistry'.
- 'Taq polymerase' an enzyme used in PCR was described as 'molecule of the year' 1989.

Principle of PCR

- The PCR technique is based on the enzymatic replication of DNA
- In PCR, a short segment of DNA is amplified using primer mediated enzymes
- DNA Polymerase synthesises new strands of DNA complementary to the template DNA
- The DNA polymerase can add a nucleotide to the pre-existing 3’-OH group only. Therefore, a primer is required
- The DNA fragment to be amplified should not be greater than about 3 kb in length and ideally less than 1 kb

Requirements for PCR

- DNA template – DNA segment to be amplified.
- Two primers- a short segment of DNA (forward and reverse primers) about 20–25 bases long .
- Taq polymerase – an enzyme to synthesize DNA copies.
- Deoxynucleotide triphosphates – the building blocks for new DNA strand.
- Buffer solution – a suitable chemical environment.
- Divalent cations – Mg^{2+} ions
- Monovalent ions – Potassium ions
- PCR machine – a thermal cycler

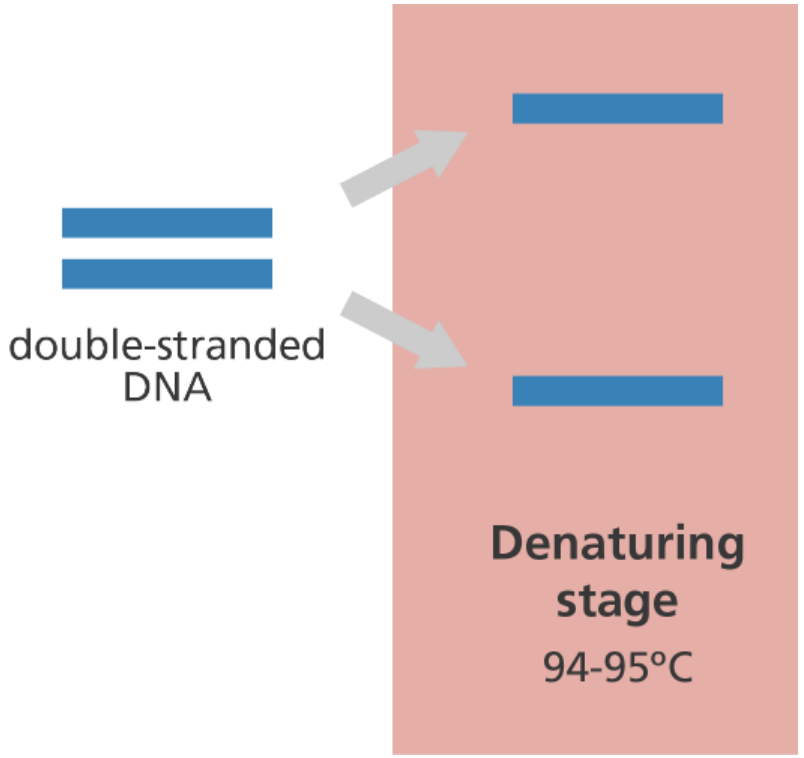
Components of PCR

- Components Of PCR constitutes the following:

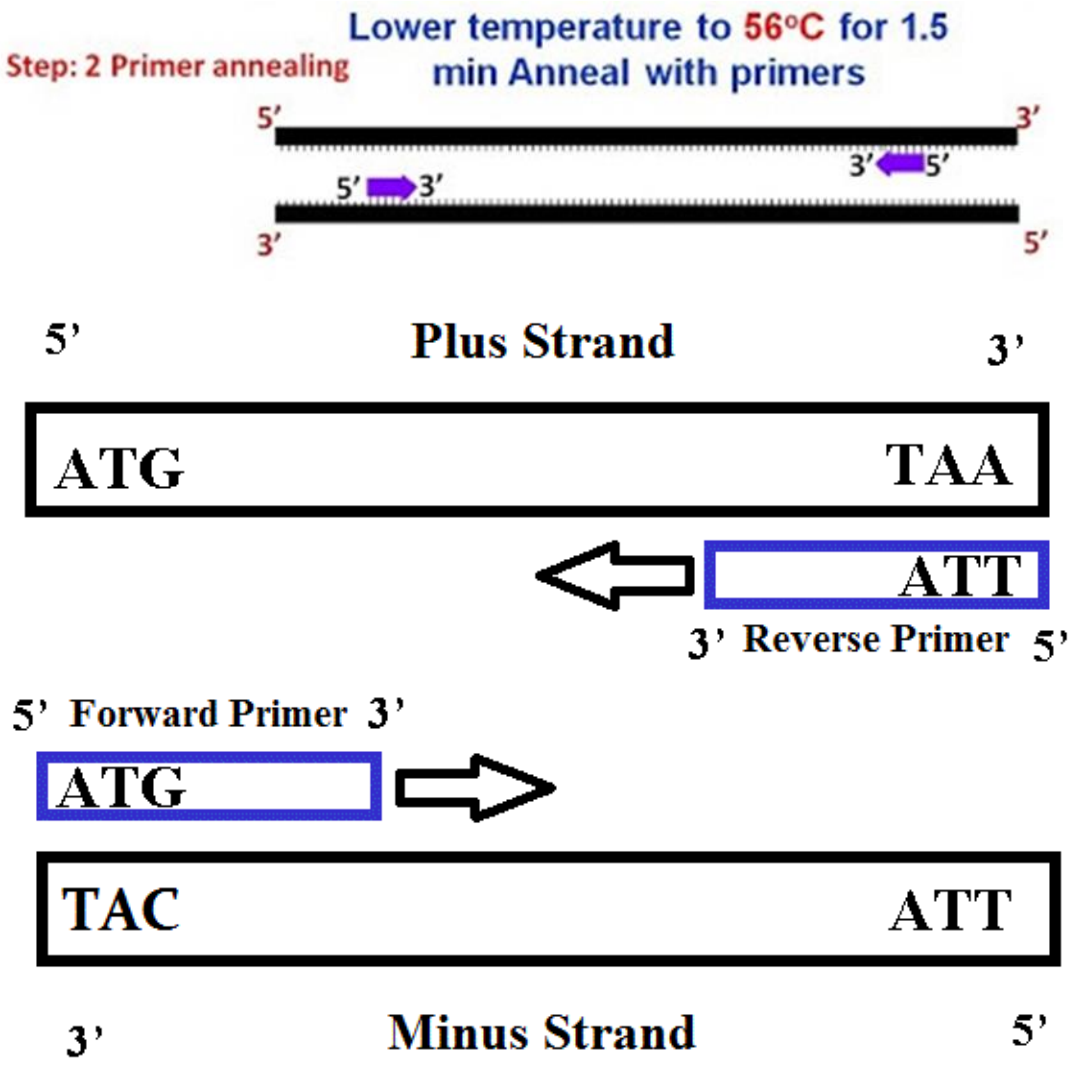
1. **DNA Template**– The DNA of interest from the sample
2. **DNA Polymerase**– Taq Polymerase is used. It is thermostable and does not denature at very high temperatures
3. **Oligonucleotide Primer**s- These are the short stretches of single-stranded DNA complementary to the 3’ ends of sense and anti-sense strands
4. **Deoxyribonucleotide triphosphate**–building blocks for the synthesis of DNA
5. **Buffer System**– Magnesium and Potassium provide optimum conditions for DNA denaturation and renaturation. It is also important for fidelity, polymerase activity, and stability.

PCR Steps: Denaturation

- Denaturation occurs when the reaction mixture is heated to 94℃ for about 0.5 to 2 minutes
- This breaks the hydrogen bonds between the two strands of DNA and converts it into a single-stranded DNA
- The single strands now act as a template for the production of new strands of DNA



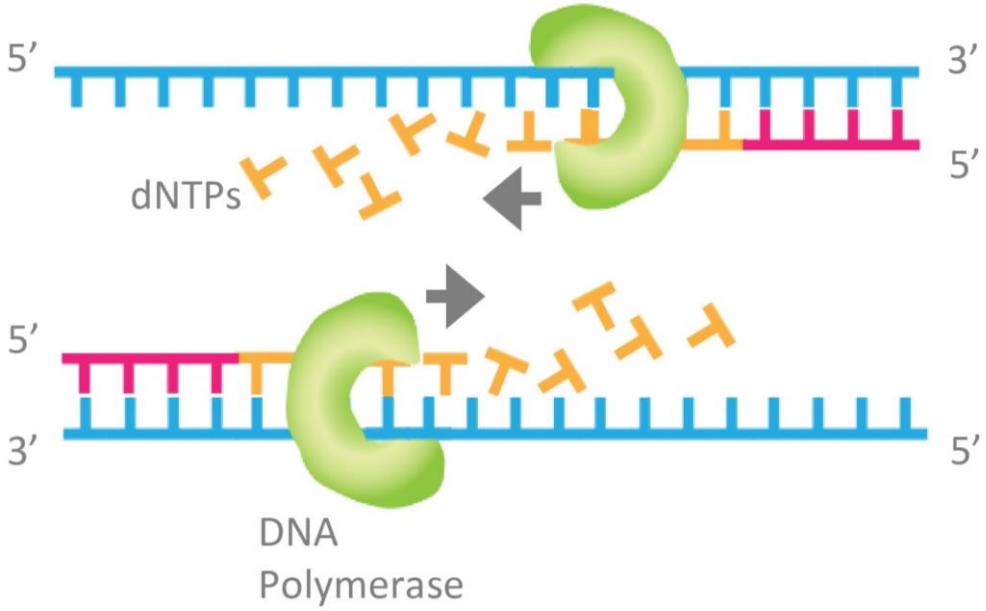
PCR Steps: Annealing



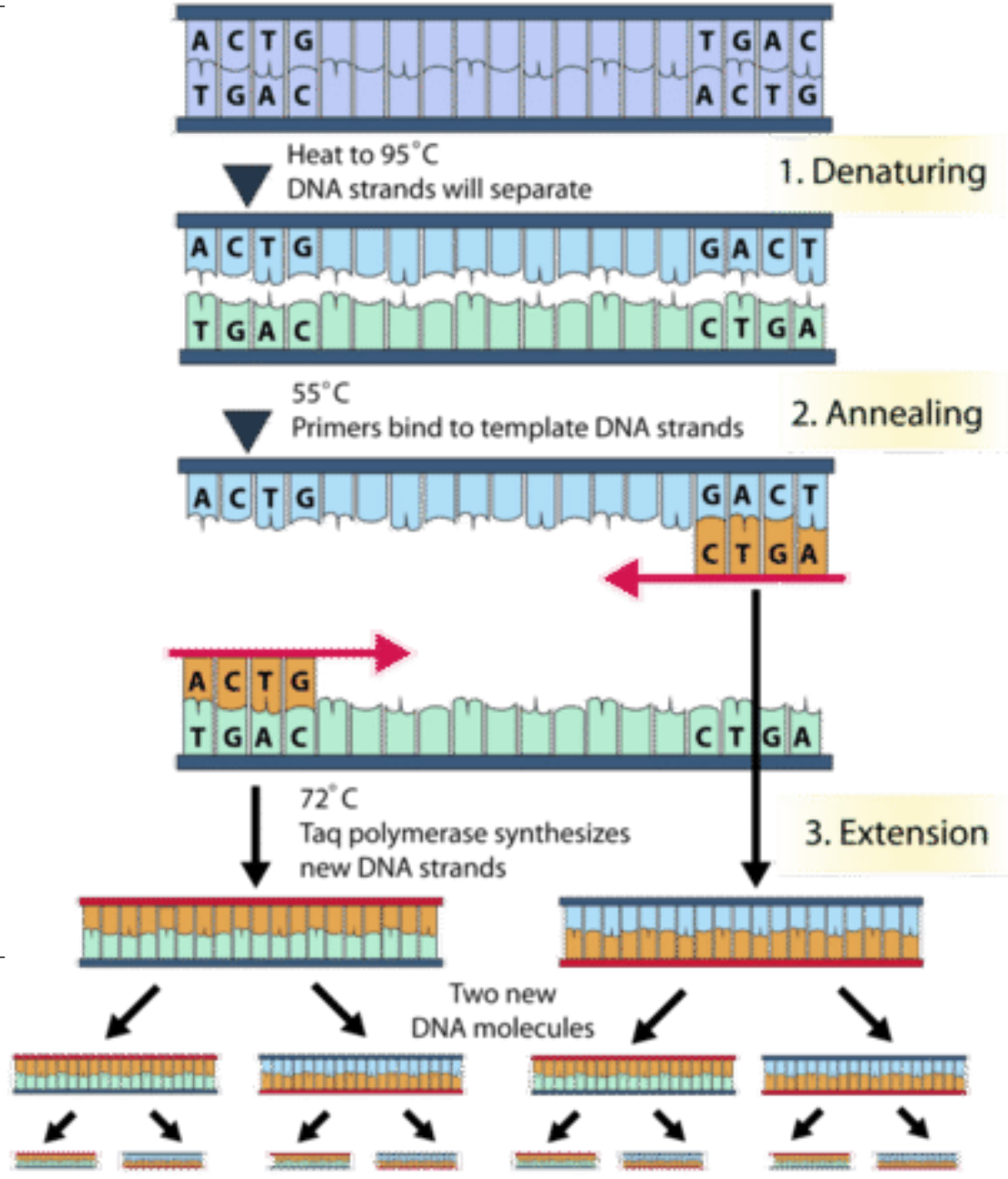
- The reaction temperature is lowered to 54-60℃ for around 20-40 seconds
- Here, the primers bind to their complementary sequences on the template DNA
- Primers serve as the starting point for the synthesis of DNA
- The two separated strands run in the opposite direction and consequently there are two primers- a forward primer and a reverse primer.

PCR Steps: Extension

- At this step, the temperature is raised to 72-80℃
- The bases are added to the 3’ end of the primer by the Taq polymerase enzyme
- This elongates the DNA in the 5’ to 3’ direction
- The DNA polymerase adds about 1000 bp/minute under optimum conditions
- Taq Polymerase can tolerate very high temperatures. It attaches to the primer and adds DNA bases to the single strand

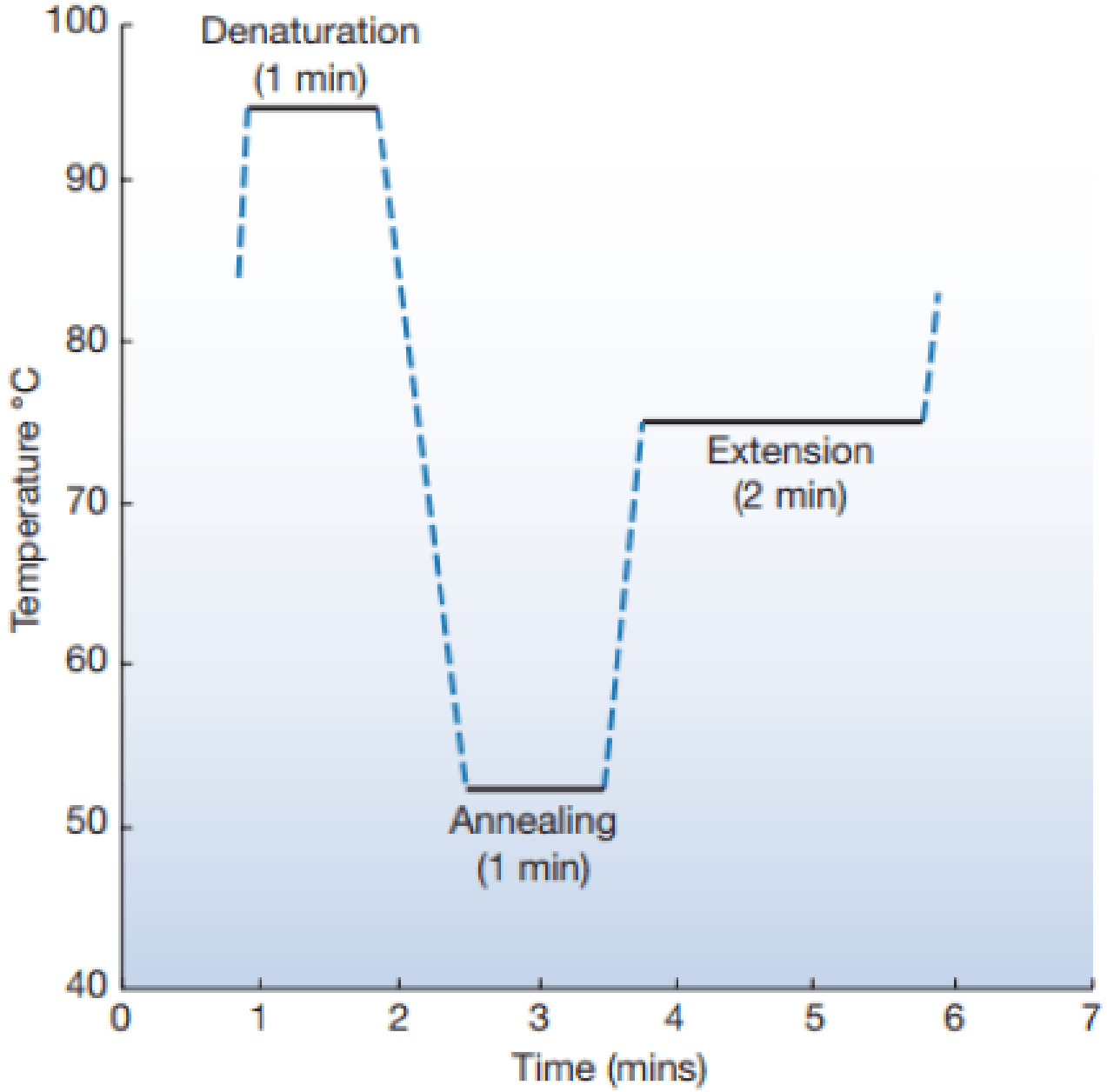


1st cycle- 2 copies of DNA (2^n, where n= no. of cycles)



2nd cycle- 4 DNAs (2^2)

3rd cycle- 8 DNAs (2^3)


Denaturation
(1 min)
Annealing
(1 min)
Extension
(2 min)
Temperature °C
100
90
80
70
60
50
40
0
1
2
3
4
5
6
7
Time (mins)

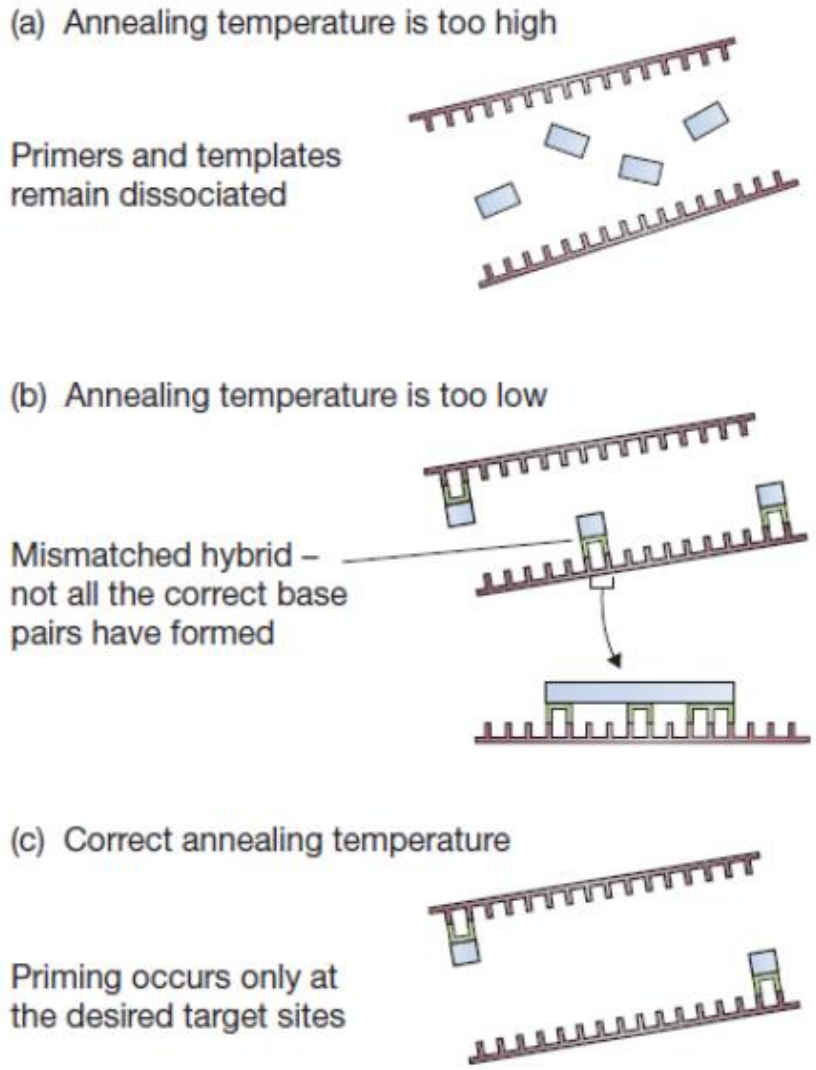
Factors affecting PCR

Annealing Temperature:

- Melting temperature (T_m): is the temp at which 2 strands of the duplex dissociate
- Melting temperature (T_m) between 55 and 65°C (usually corresponds to 45-55% G+C)
- The melting temperature of nucleic acid duplex increases both with its length, and with increasing (G+C) content
- A simple formula for calculation of the T_m is:

  $T_m = 4(G + C) + 2(A + T)°C$

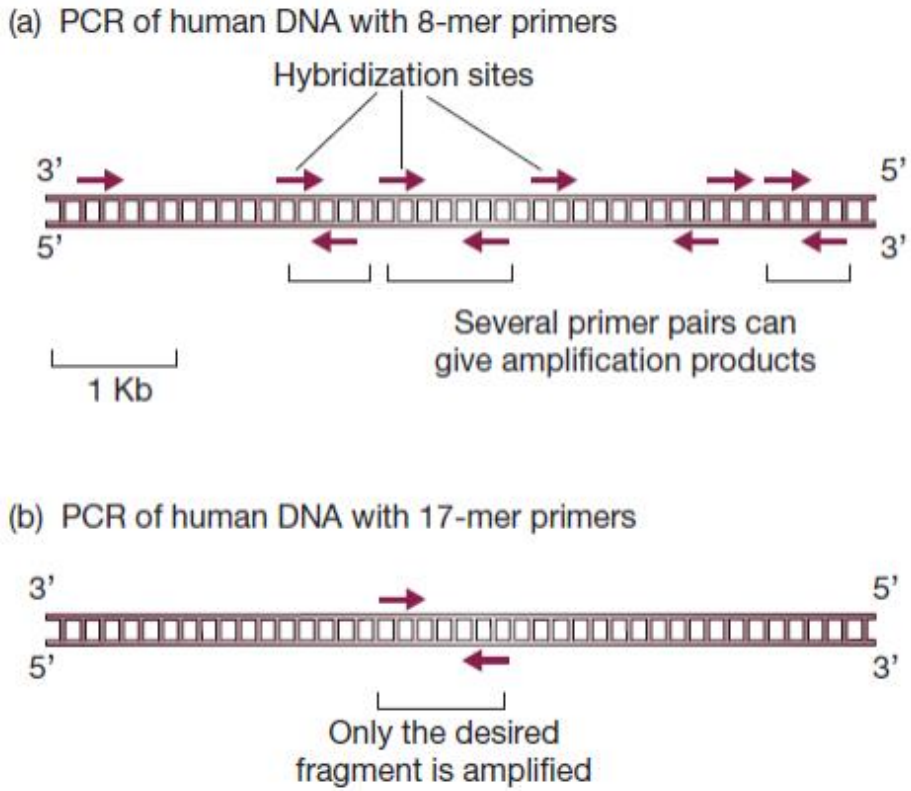
- Annealing temperature (T_a) is about 5°C below the T_m of the pair of primers (T_m - 5°C)
- Annealing temperature chosen for a PCR depends directly on length and composition of the primer(s)
- Two primers Tm's should be within 5°C of each other; the closer the better



Factors affecting PCR

Primer length:

- If the primers are too short they might hybridize to non-target sites and give undesired amplification products
- 4^8 = 65 536 bp, giving approximately 49 000 possible sites in the 3 200 000 kb of nucleotide sequence that makes up the human genome
- Sixteen base sequence will statistically be present only once in every 4^{16} bases (= 4 294 967 296 or 4 billion). This is about the size of the human genome
- Thus, the association of a greater-than- 17-base oligonucleotide with its target sequence is an extremely sequence-specific process
- Ideally, the length of the primer should range between 17-25 nts



Factors affecting PCR

Primer-Primer interactions:

- If a primer has a region of self-homology, partially double-stranded structures can occur which will interfere with annealing to the template
- Usually intraprimer homologies of 3 bp or more should be avoided

- Absence of significant hairpin formation (usually >3 bp)
- Primers should also not contain sequences of nucleotides that would allow one primer molecule to anneal to another primer molecule

Factors affecting PCR

Primer-Primer interactions:

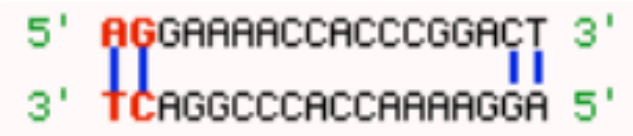
- Ideally, the smallest ΔG, the better but try to avoid primers annealing with ΔG values of –7 kcal/mol or higher

Here are other examples of interacting primers. The following are OK, but not great:

ΔG=-5 kcal/mole, OK but not great.

ΔG=-9.3 kcal/mole, pretty bad!!

The next is real nice! ΔG=-1.6 kcal/mole, good!!



Degenerate Primers

- In case where there is partial information of DNA is present, degenerate primers are used
- They are usually used when the protein sequence is known, which is translated backwards to derive the triplets (codon), and therefore the probable sequence
- Due to degenerate genetic code, there might be different possibilities for the third nucleotide in each triplet, that can be circumvented by the degenerate primers
- The degenerate primers are mix of primers, which are similar in their sequence, but differ in one or more positions

Degenerate Primers

aligning the **translated amino acid sequence**

Trp Asp Thr Ala Gly Gln Glu
5' TGG GAY ACN GCN GGN CAR GA 3'

This gives a mix of 256 different oligonucleotides.

Aspartic acid	D	GAC, GAT
Threonine	T	ACA, ACC, ACG, ACT
Alanine	A	GCA, GCC, GCG, GCT
Glycine	G	GGA, GGC, GGG, GGT
Glutamine	Q	CAA, CAG

Y = C + T
N = A + G + C + T
R = A + G

2*4*4*4*2 = 256

PCR Enzyme

Choosing enzyme

- Template (RNA or DNA?)
- Fidelity
- Temperature stability
- Processivity

- Template (RNA or DNA?)
- Tth DNA Polymerase from *Thermus thermophilus* reverse transcribes RNA in the presence of Mn^{2+}
  - Then dilute rxn & add Mg 2+ to do PCR
- Tfl DNA Polymerase from *Thermus flavus* has no RT activity: can mix with RNA & RT w/o activity then go directly to PCR after RT is done

- Fidelity
  - *Taq* from *Thermus aquaticus* has no proof-reading
  - goes faster, but error freq of 1 in 3000
  - Vent from *Thermococcus litoralis* has error frequency of 1 in 30,000
  - Pfu from *Pyrococcus furiosus* has error frequency of 1 in 400,000

- Temperature stability
  - E.coli DNA polymerase I denatures at 75° C
  - $T_{1/2}$ of Taq @ 95° C is 0.9 hours, < 0.1 hour @ 100° C
  - $T_{1/2}$ of Phusion @ 96° C is >6 hours, 2 hours @ 98° C
  - $T_{1/2}$ of Vent @ 95° C is 6.7 hours, 1.8 hours @ 100° C

- Processivity (how far does it go before falling off)

Thermo tolerant polymerases used for PCR (polymerase chain reactions) reactions

The total error rate of Taq polymerase has been variously reported between 1 x 10^{-4} to 2 x 10^{-5} errors per base pair.

Pfu polymerase appears to have the lowest error rate at roughly 1.5 x 10^{-6} error per base pair

Vent is intermediate between Taq and Pfu.

Polymerase	3'->5' Exonuclease	Source and Properties
Taq	No	From *Thermus aquaticus*. Halflife at 95C is 1.6 hours.
Pfu	Yes	From *Pyrococcus furiosus*. Appears to have the lowest error rate of known thermophilic DNA polymerases.
Vent	Yes	From *Thermococcus litoralis*; also known as Tli polymerase. Halflife at 95 C is approximately 7 hours.

Thermostable DNA polymerases used in PCR

Polymerase enzyme	5’→ 3’ exonuclease	3’→ 5’ exonuclease	Organism
Tli (Vent)	No	Yes	Thermococcus litoralis
Tfl	?	No	Thermus flavus
Pfu	?	Yes	Pyrococcus furiosus
Pwo	No	Yes	Pyrococcus woesei
Tth	?	No	Thermus thermophilus

PCR Product calculation

1→2→4→8→16 (exponential)

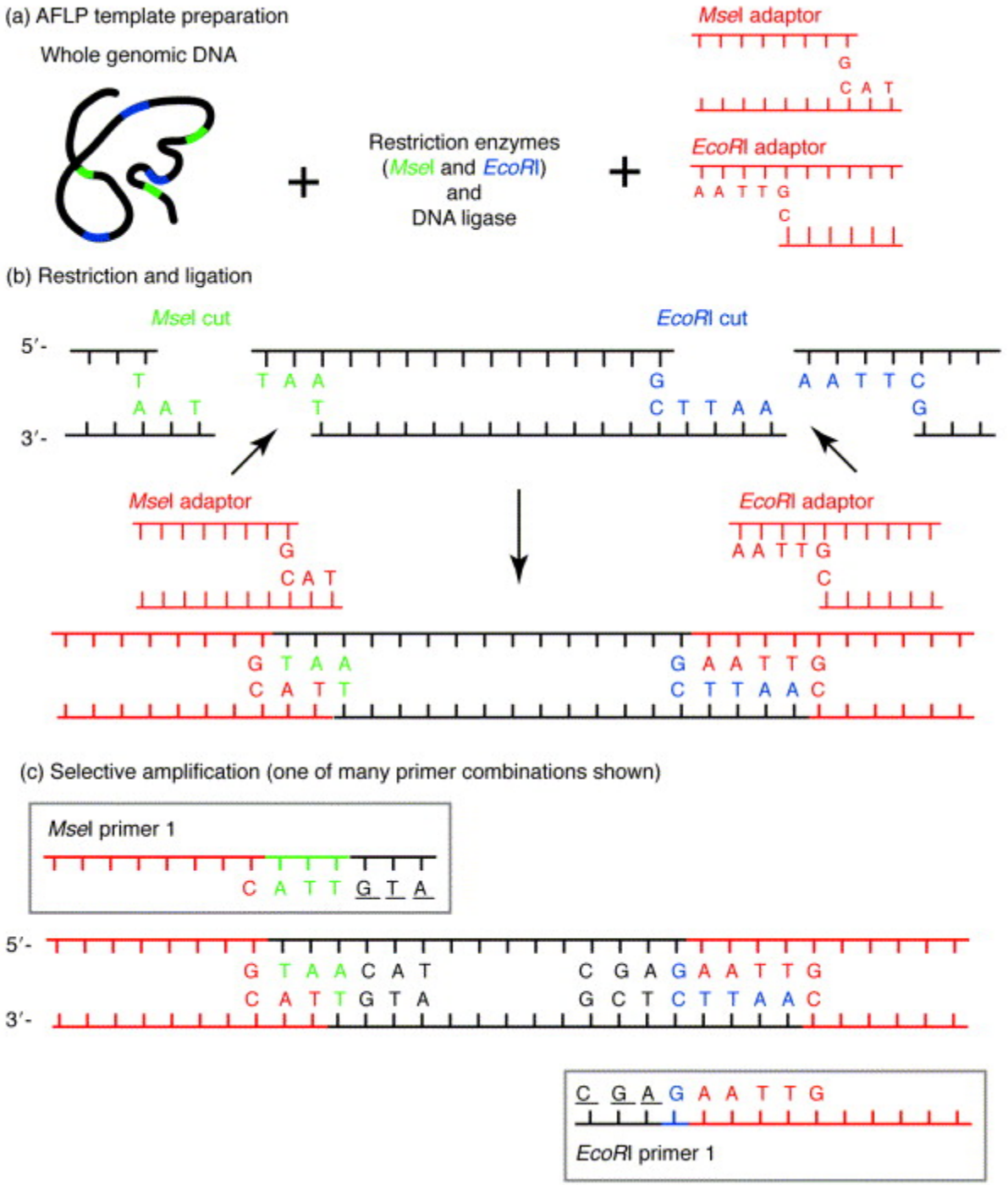
- **$N = N_0 \times 2^n$**

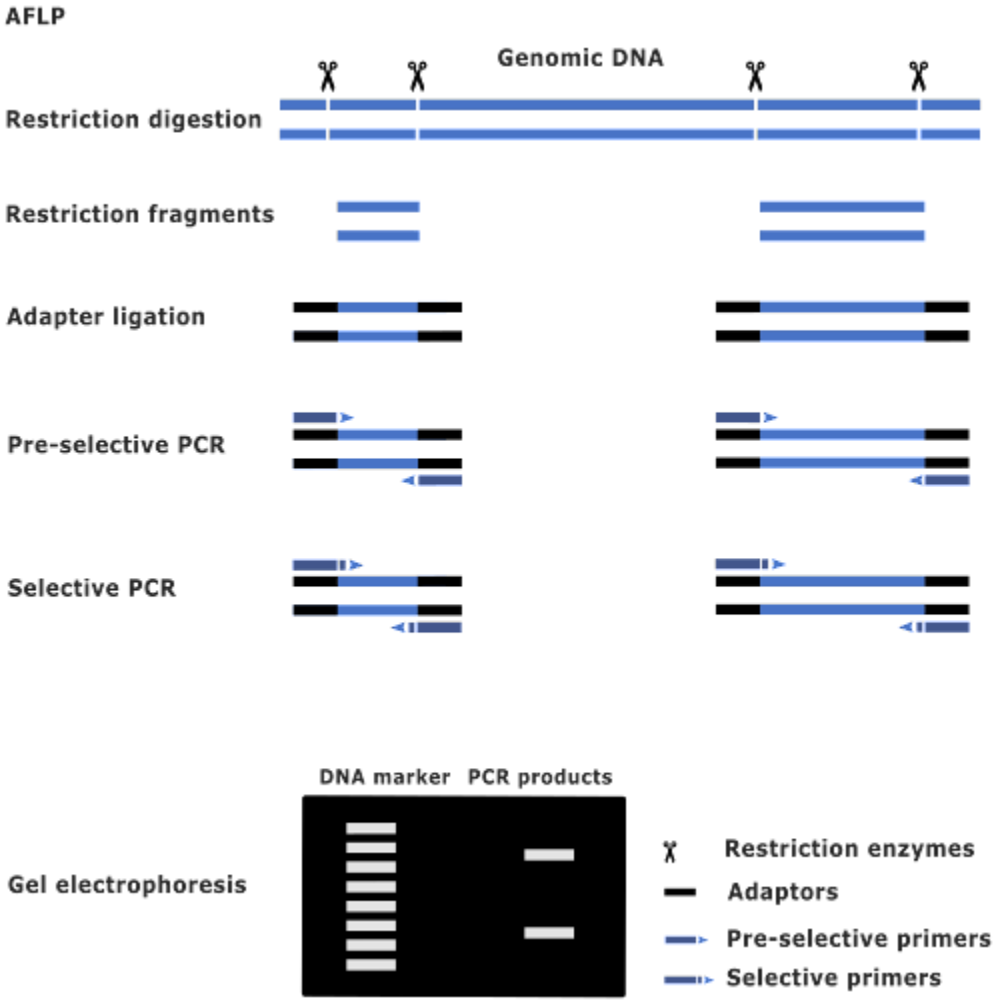
where N → Final number of templates, N_0 → Initial number of templates, n = no. of cycles

- If the values are given in concentrations of DNA:

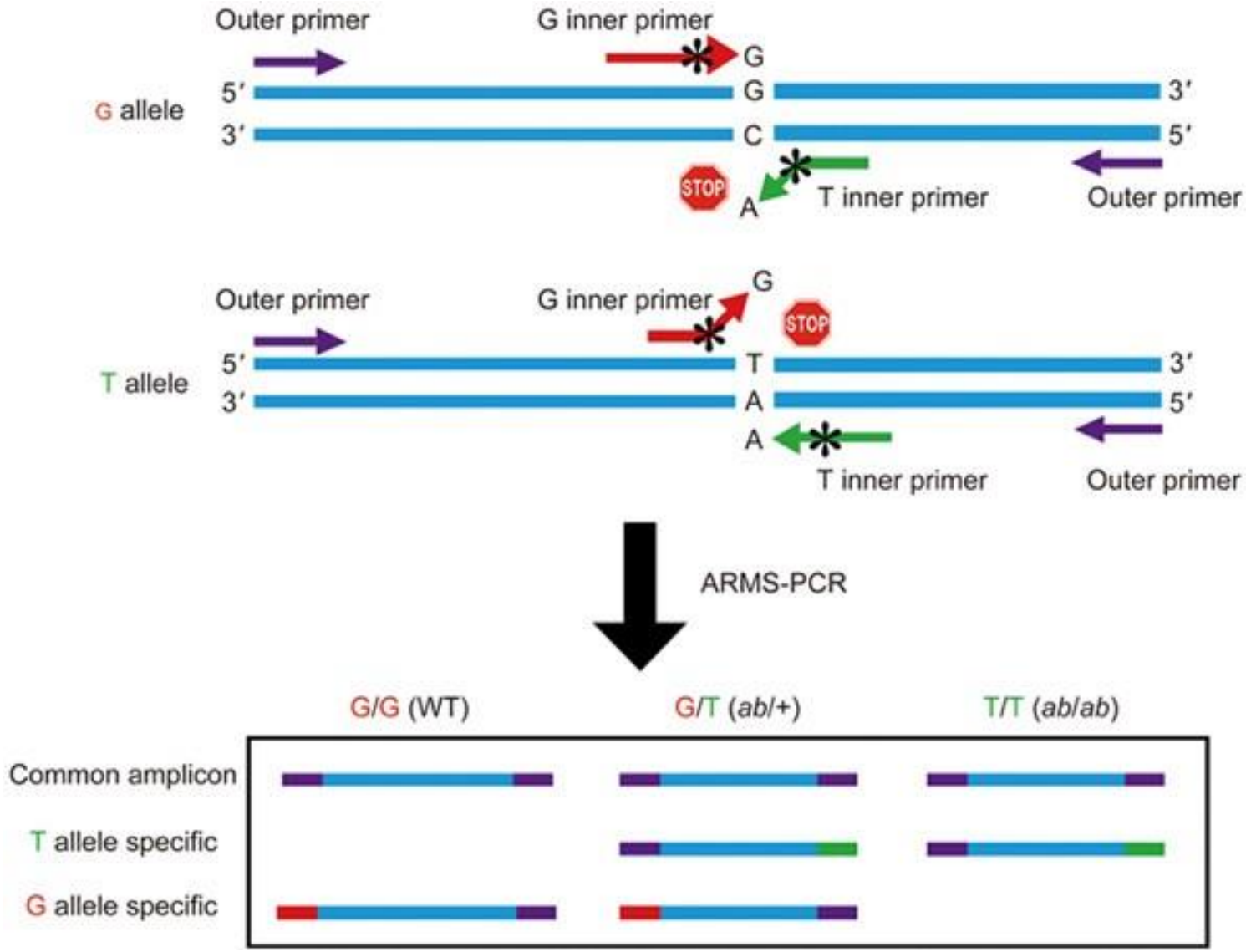
PCR product = initial amount X [1 + % efficiency]^n

Amplified fragment length polymorphism (AFLP)



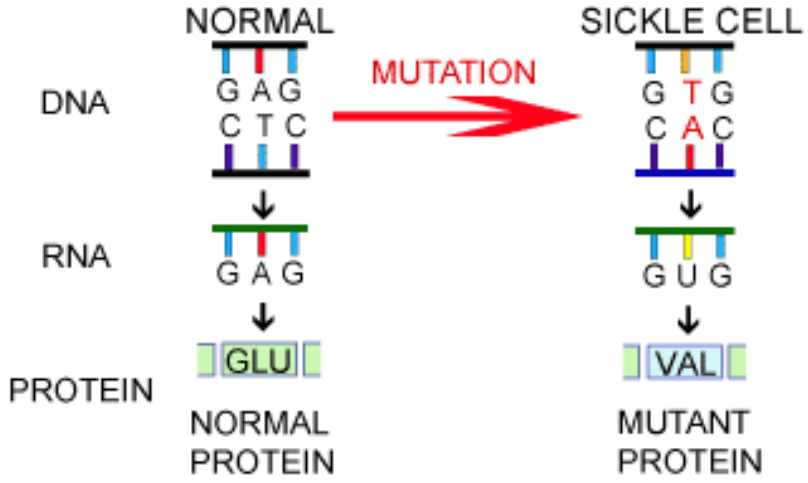


Allele-Specific PCR



- **GC- wild type (G allele)**
- **TA- mutant (T allele)**

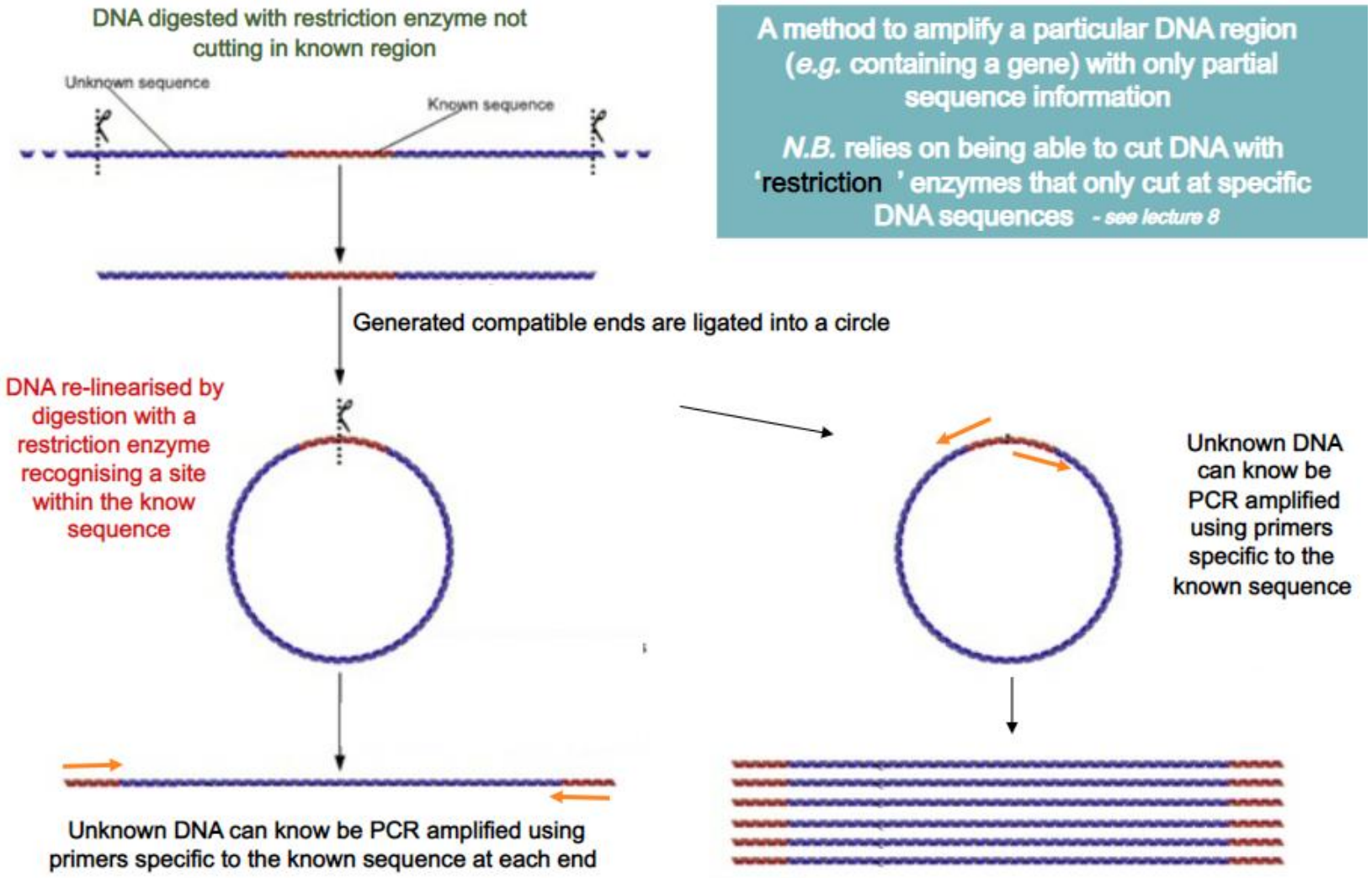
- **Outer primers- gene specific (forward and reverse)**
- **Inner primers- wild type and mutant type**



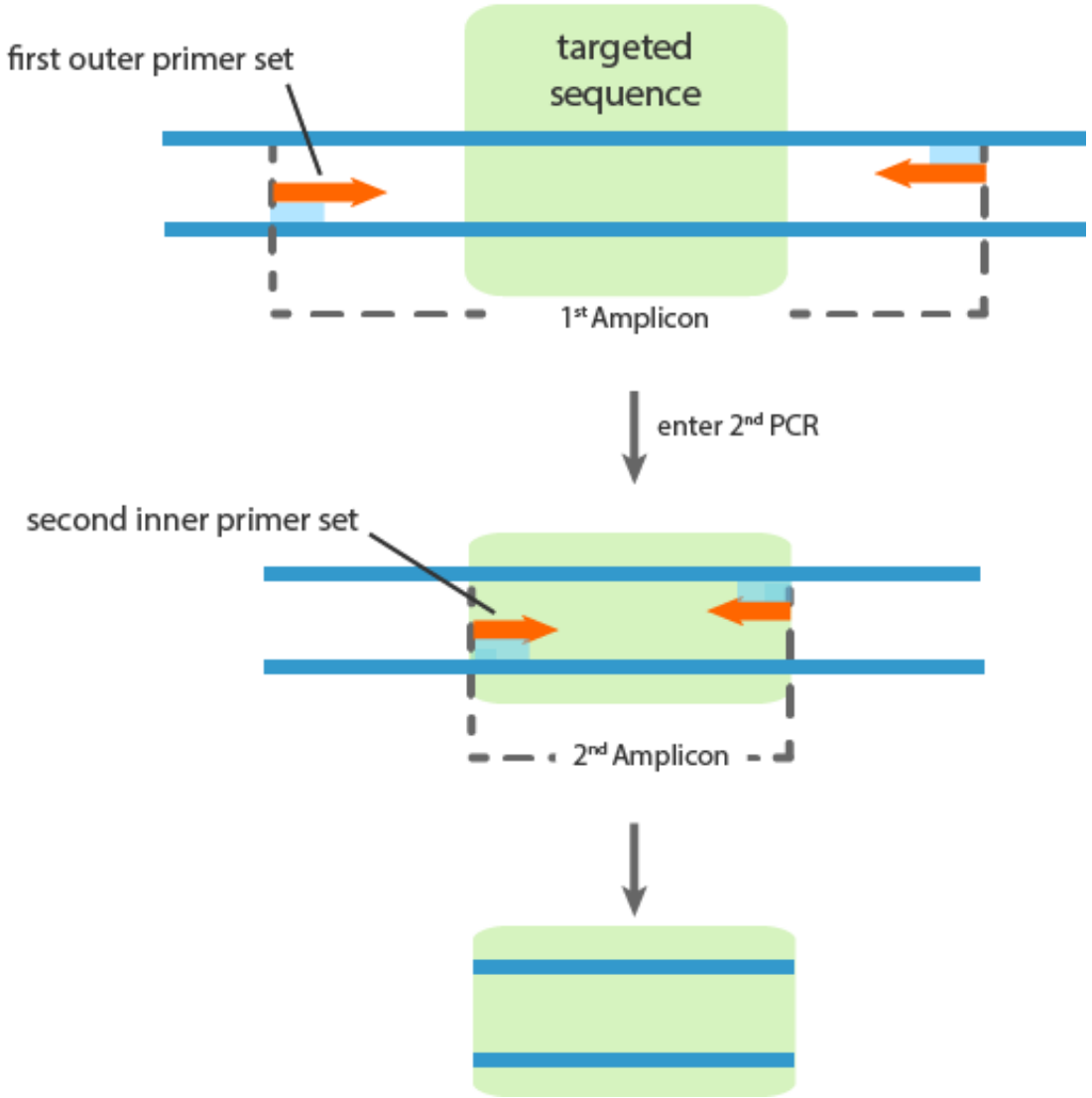
Inverse PCR

A method to amplify a particular DNA region (*e.g.* containing a gene) with only partial sequence information

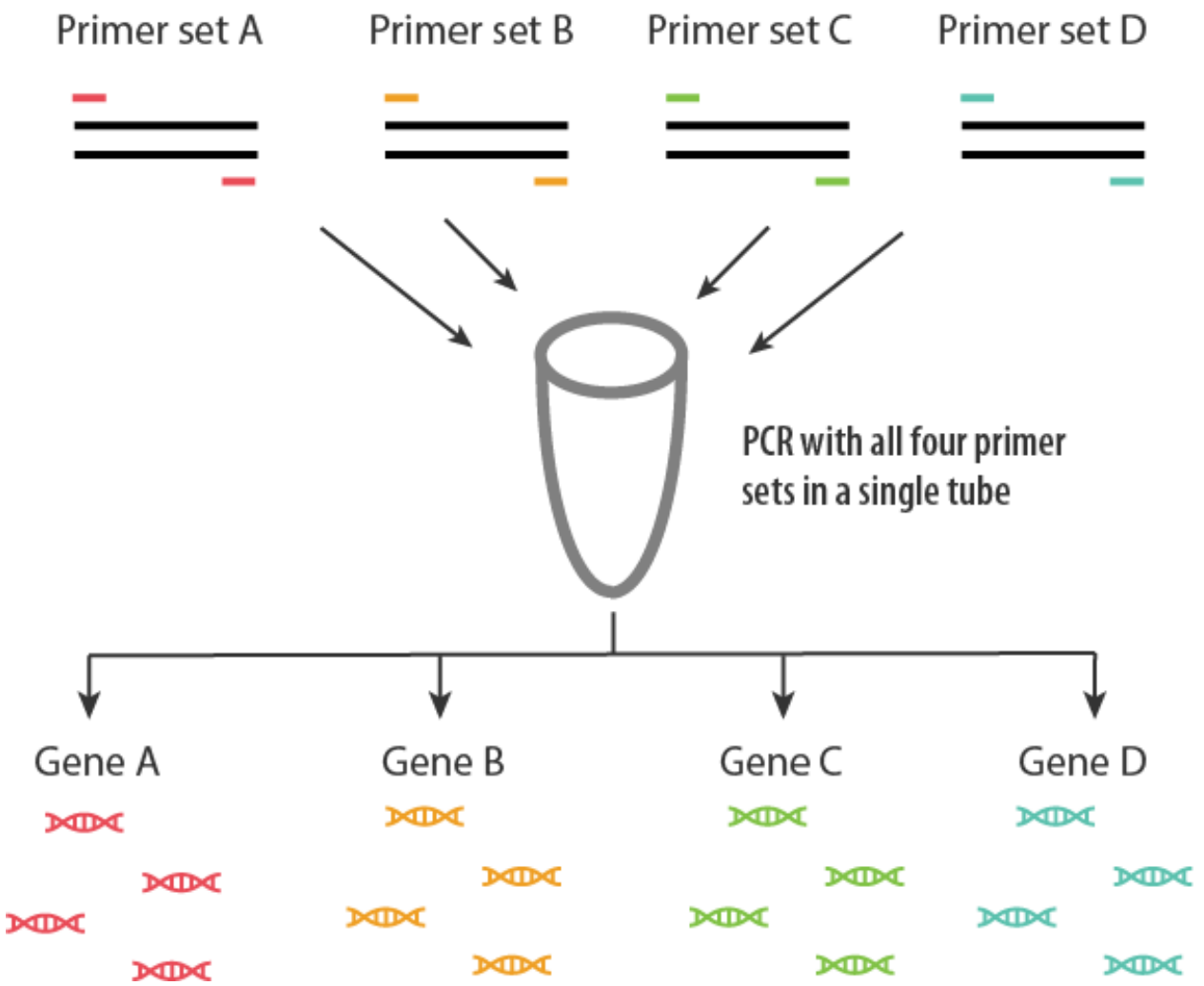
N.B. relies on being able to cut DNA with 'restriction ' enzymes that only cut at specific DNA sequences *- see lecture 8*



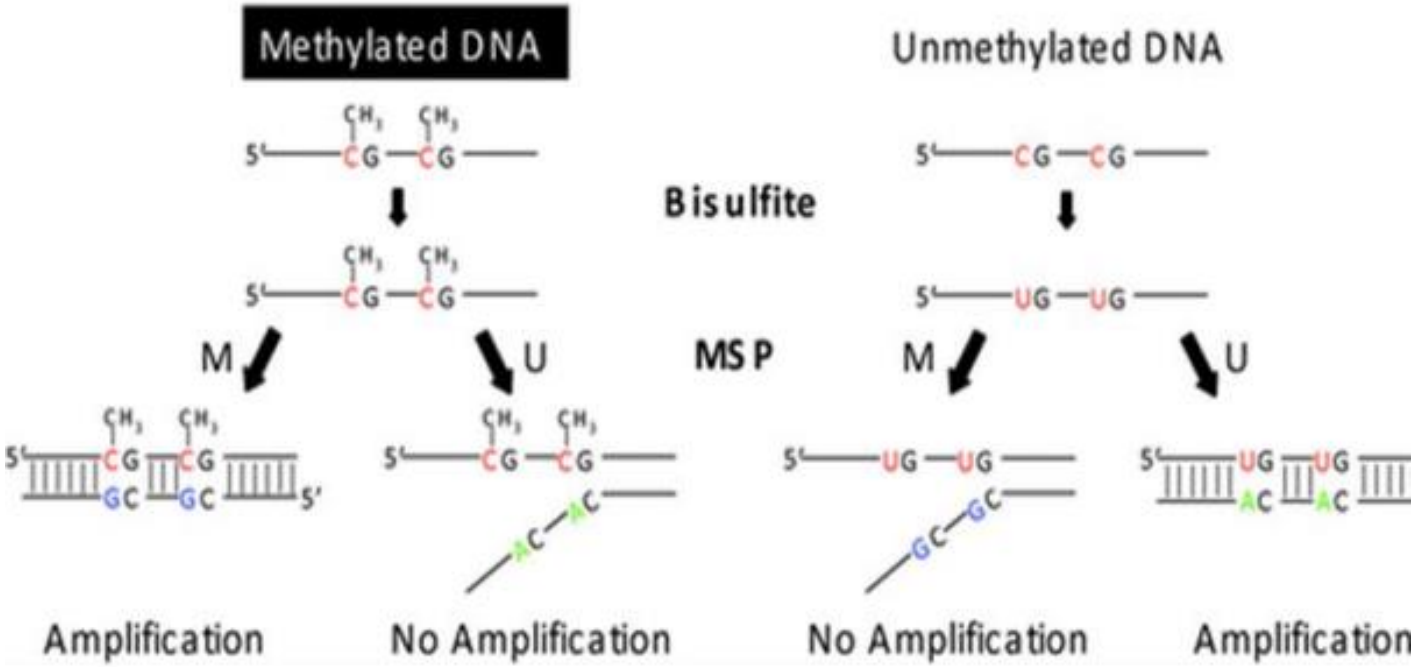
Nested PCR

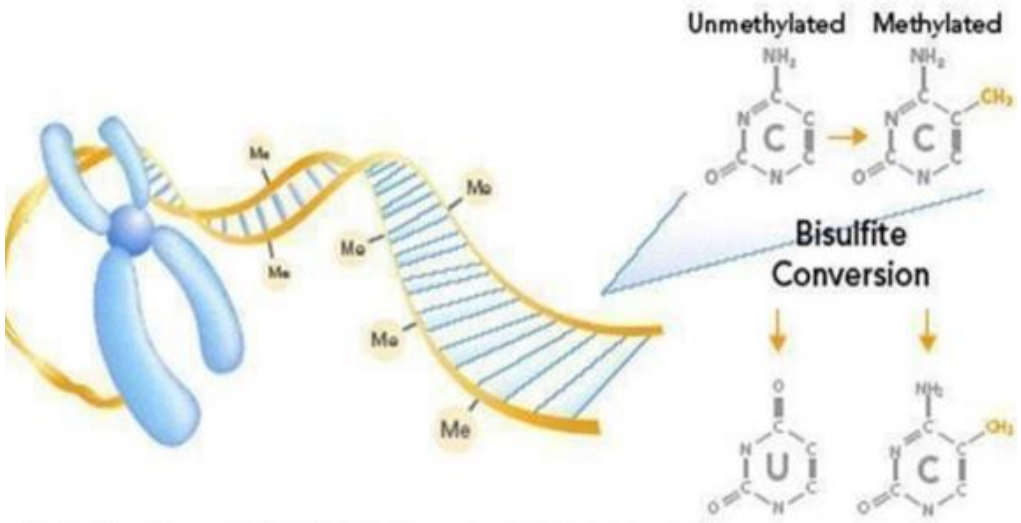

first outer primer set
targeted sequence
1st Amplicon
enter 2nd PCR
second inner primer set
2nd Amplicon

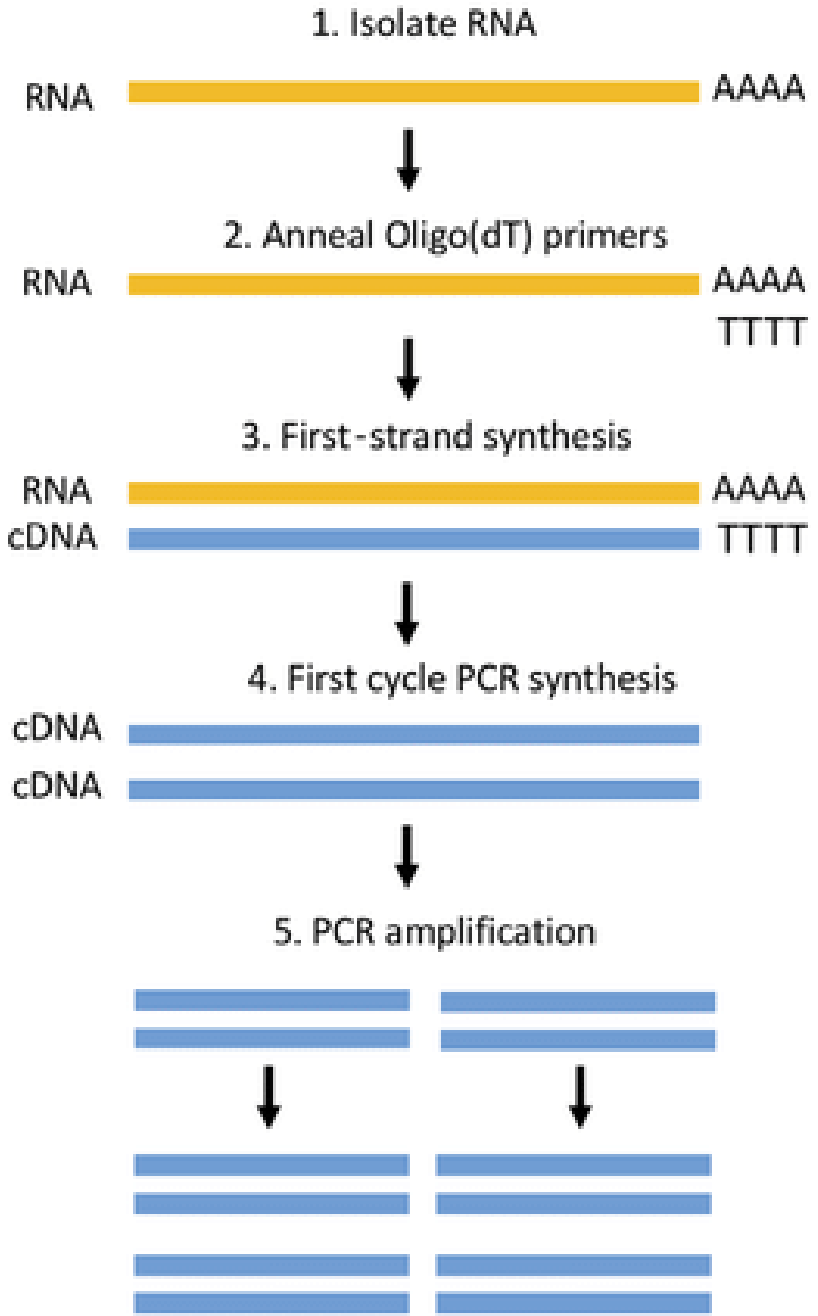
Multiplex PCR



Methylation-specific PCR (MSP)



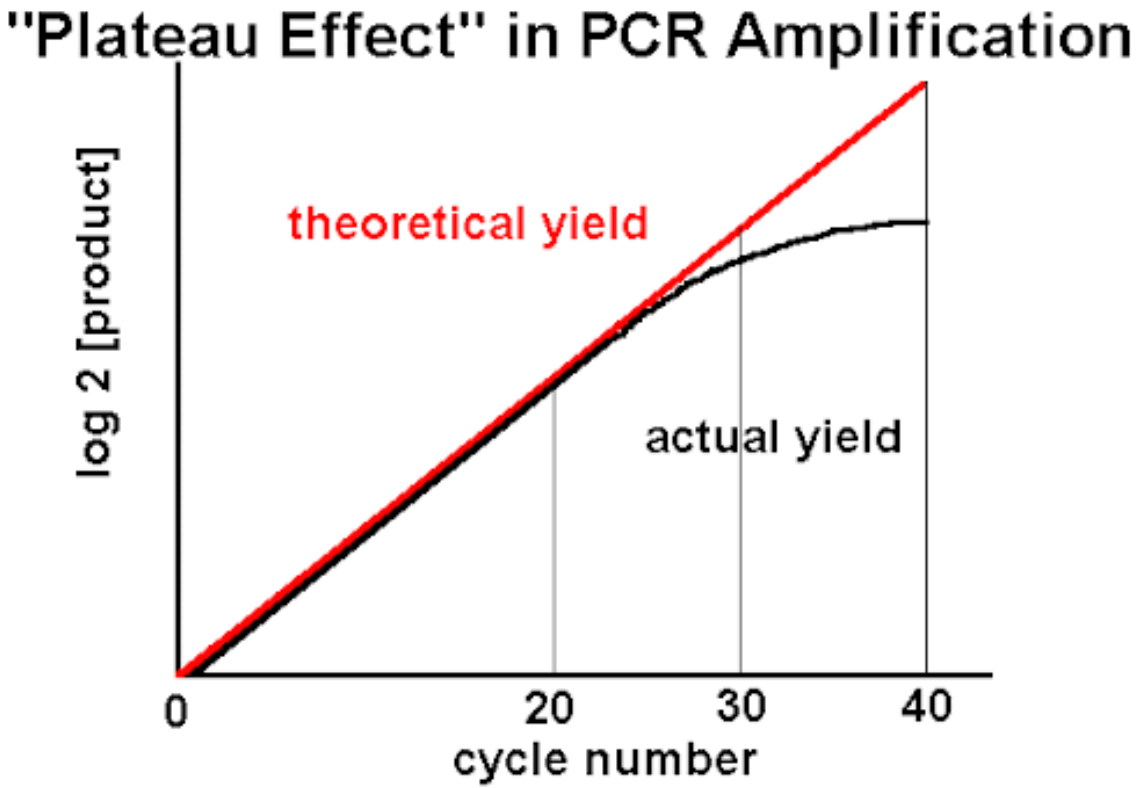




Reverse Transcriptase PCR

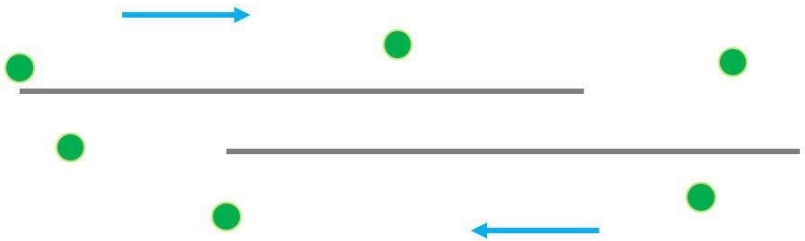
Plateau effect in PCR amplification

- **Plateau effect:** The point at which product accumulation is no longer exponential (saturation)
- Taq poly half life at 95℃ is 1.6 hrs
- Degradation of reagents
- Competition for reactants by non-specific products
- Competition for primer binding by re-annealing of concentrated products

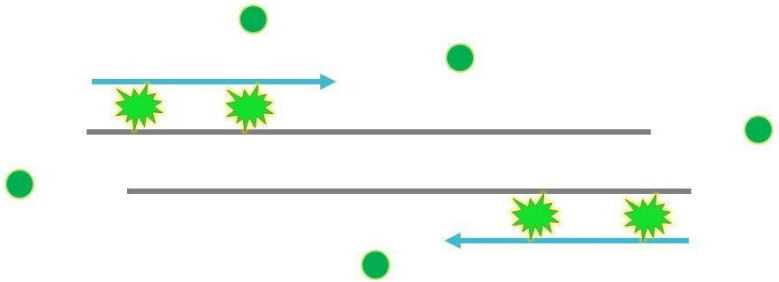


SYBR

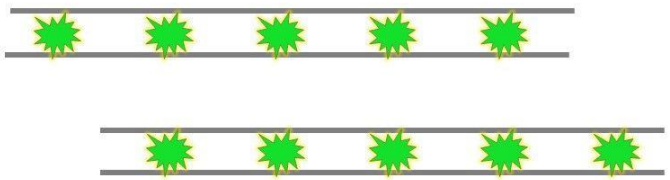
Denature

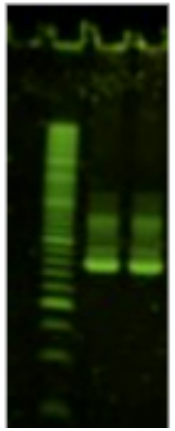


Polymerization



Signal detection (Polymerization completed)

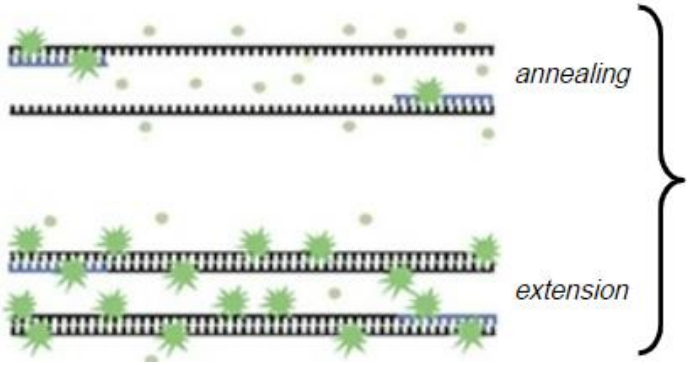




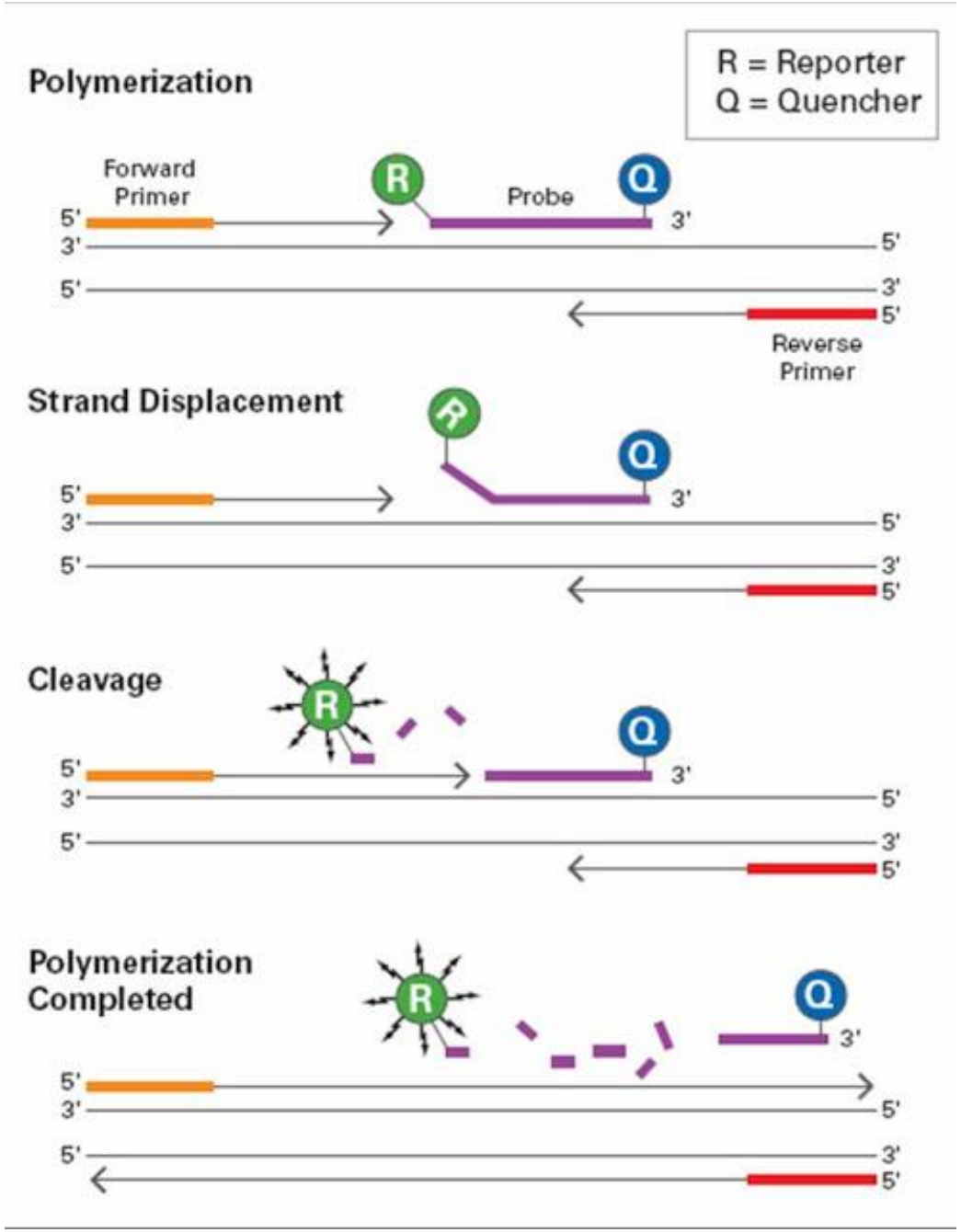
SYBR green-based Q-PCR assay

- ds DNA intercalating dye
- fluoresces green under blue light
- only emits fluorescence when bound to double stranded DNA

Real time PCR

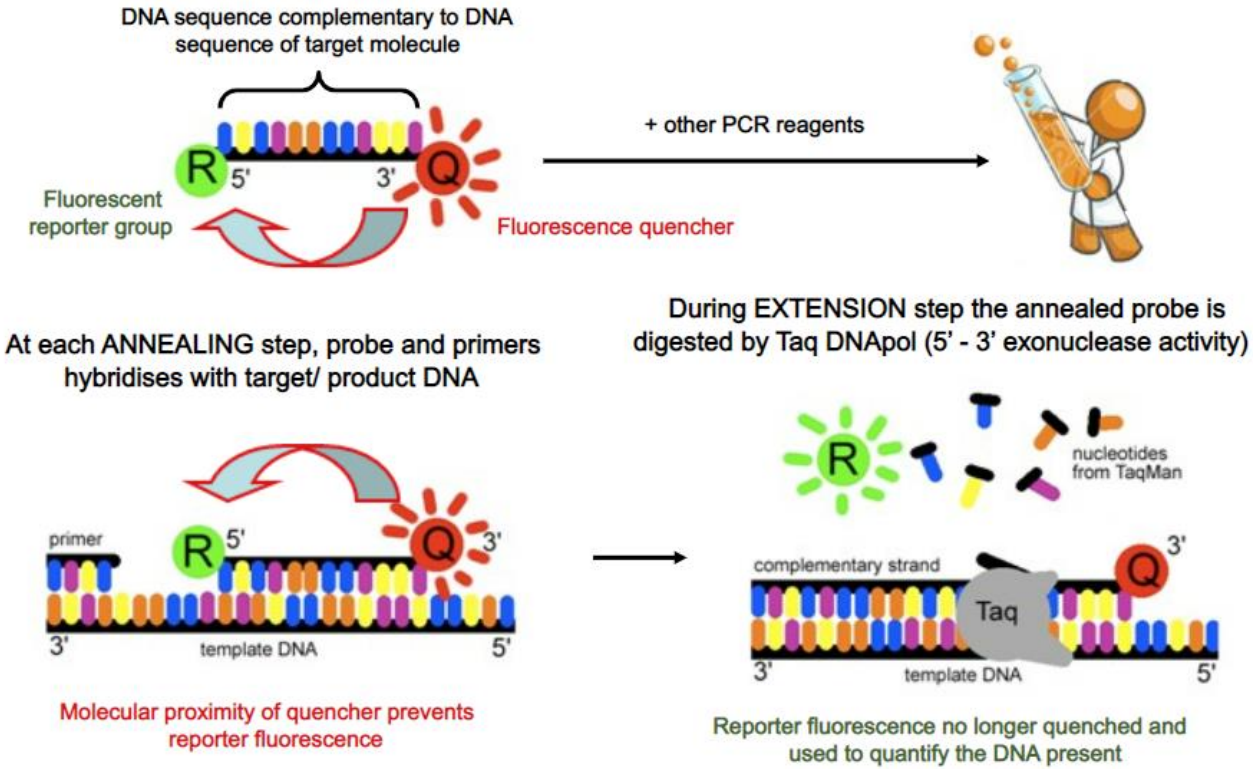


SYBR green fluorescence can be measured at the end of either the annealing* or extension steps after every PCR cycle and used to calculate the amount of DNA in the sample



Real-time PCR (Quantitative PCR or Q-PCR)

Fluorescent hybridisation probe based methods (*e.g.* TaqMan probes)



Real time PCR

Question 1

A single molecule of DNA is amplified by PCR for 25 cycles. Theoretically, how many molecules of amplicon will be produced?

($2^{10} = 1024$)

A. 64×10^5

B. 32×10^6

C. 8×10^6

D. 126×10^5

[$2^{10} = 1024 \approx 1000/10^3$]

$2^{25} = 2^{10} \times 2^{10} \times 2^5 =$ **32×10^6**

Question 2

A polymerase chain reaction was performed beginning with 400 template DNA molecules in 100 µL reaction. After 20 cycles of PCR, how many molecules of the amplified product will be present in 0.1 µL of reaction?

A. 2.19×10^4
B. 4.19×10^4
C. 2.19×10^5
D. 4.19×10^5

Solution 2

$N = N_0 \times 2^n$

$N = 400 \times 2^{20}$ **[$2^{10} = 1024 \approx 1000/10^3$]**

$N = 400 \times 10^3 \times 10^3 = 400 \times 10^6$ **[$400 \times 10^6 \rightarrow$ 100 µL]**

$400 \times 10^6 \rightarrow 100$ µL

x $\rightarrow$ 0.1 µL

x $= (400 \times 10^6 \times 0.1)/100 =$ **4×10^5**

Question 3

5. All the following are thermostable polymerases except

a) Taq polymerase

b) Vent polymerase

c) DNA polymerase III

d) pfu polymerase

Question 4

(8) Which of the following denotes the Accuracy of the DNA Polymerase?

1. Processivity
2. Rate of synthesis
3. Fidelity
4. Half-life

Question 5

(18) You will be given a Forward primer
5'GCGACGGTATTCGAACTTGT-3'
-calculate its Tm and Ta

1. 60 and 75
2. 70 and 30
3. 60 and 60
4. 60 and 55
5. 55 and 60

Solution 5

- **$T_m = 4(G + C) + 2(A + T)°C$**
- **$T_a = T_m - 5°C$**

5'GCGACGGTATTCGAACTTGT-3'

- $T_m = 4(6 + 4) + 2(4 + 6)°C$

  $T_m = 4(10) + 2(10)$

  $T_m = 40 + 20 =$ **60°C**

- $T_a = T_m - 5°C = 60 – 5 =$ **55°C**