

## Oligomers design (primers and probes)

*Bioinformatics teachings*

*<http://bioinfomed.fr> - Olivier Croce*

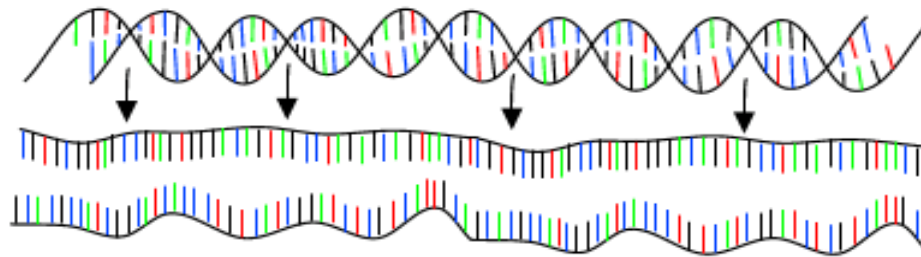
# Summary

- \* Use of PCR - Concept of primers and/or probes
- \* Sensibility, specificity and other thermodynamic constraints
- \* Examples of softwares used for oligomers design.
- \* Exercise : Searching of primers/probes for *Bacillus anthracis*, cya

# Use of PCR - Concept of primers and/or probes

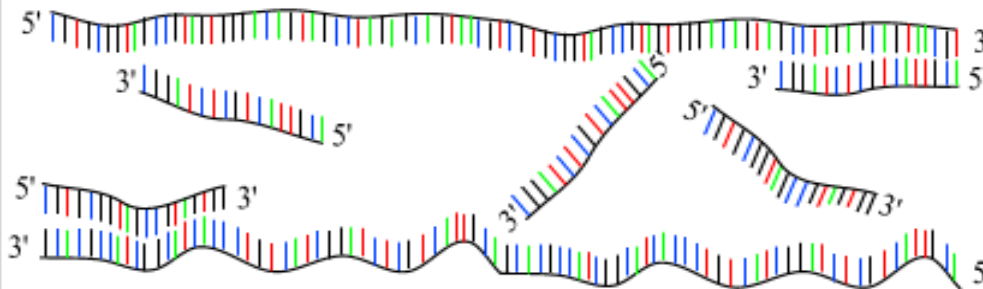
## PCR : Polymerase Chain Reaction

30 - 40 cycles of 3 steps :



**Step 1 : denaturation**

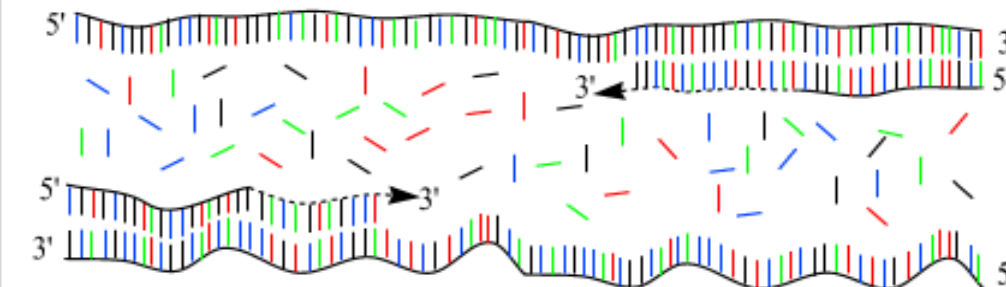
1 minut 94 °C



**Step 2 : annealing**

45 seconds 54 °C

**forward and reverse primers !!!**

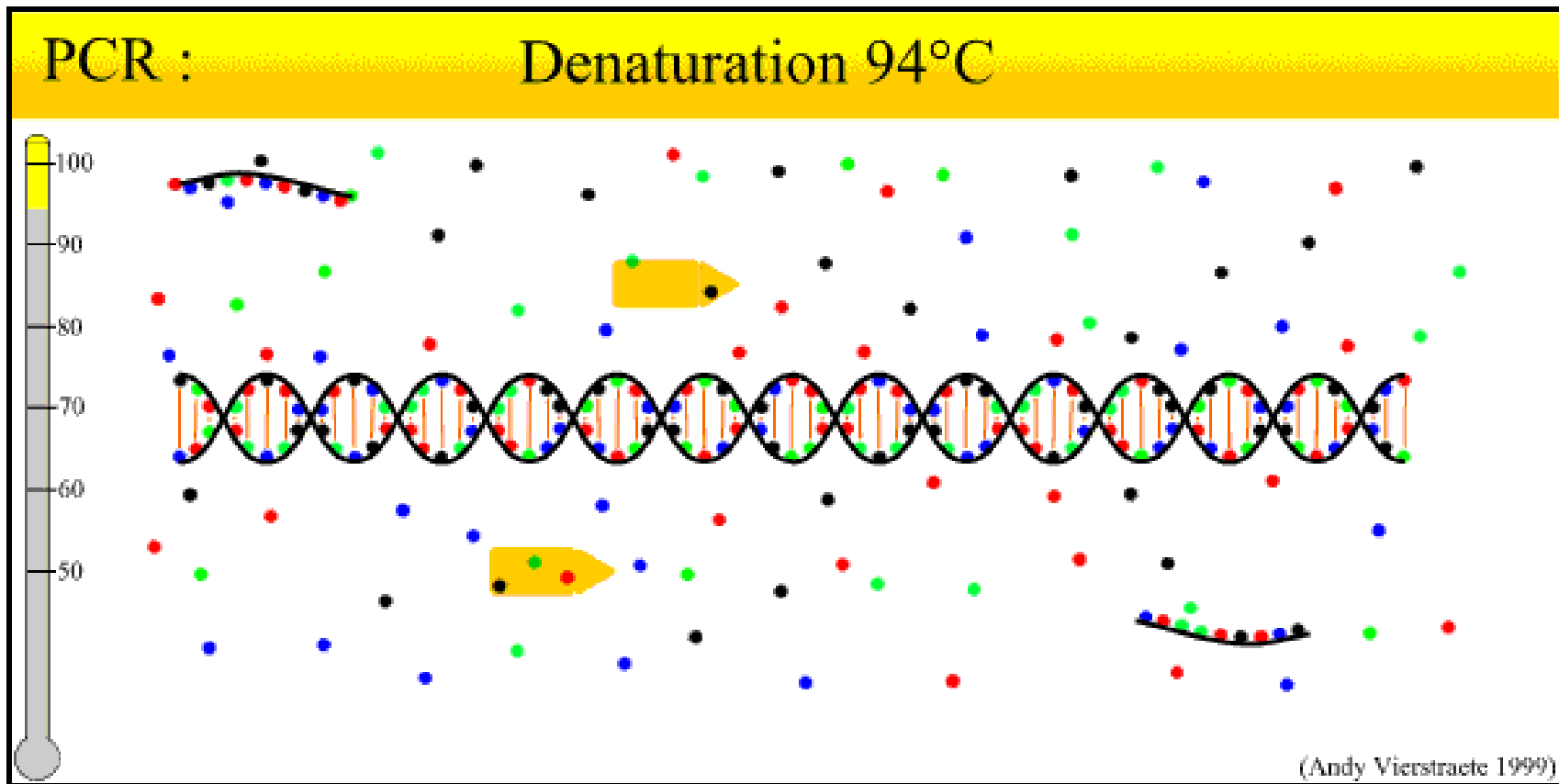


**Step 3 : extension**

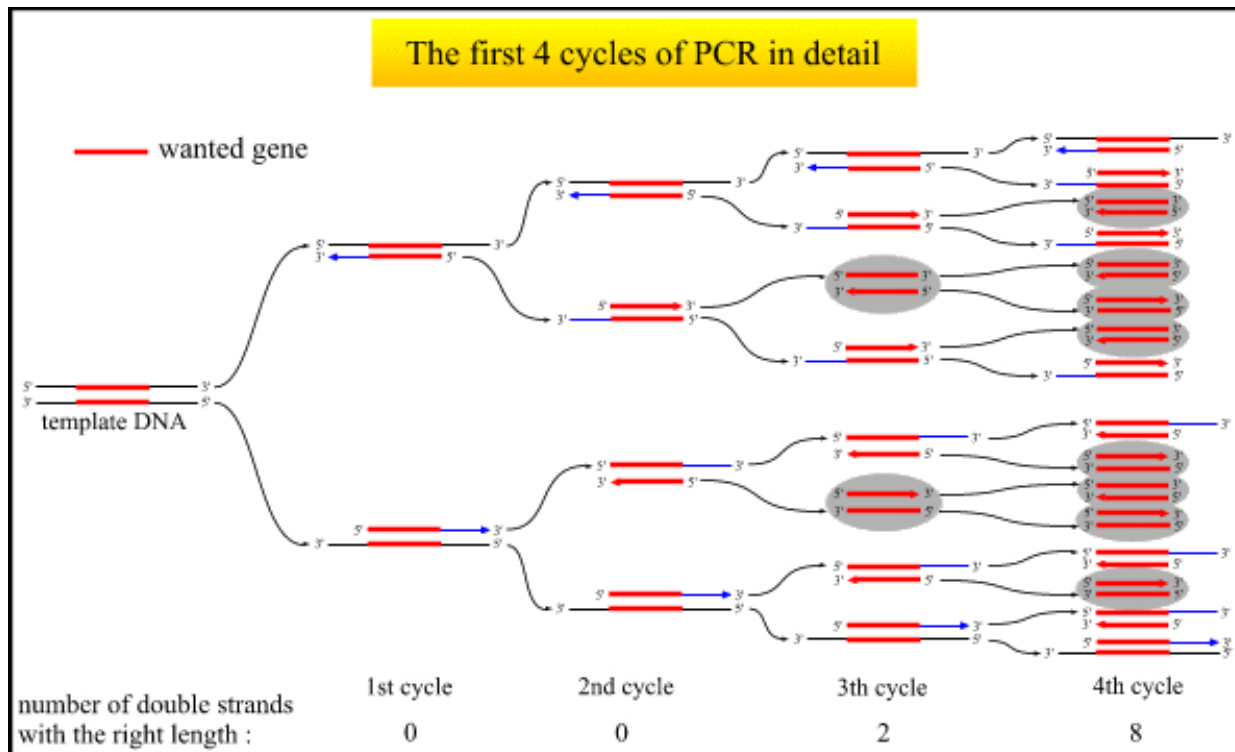
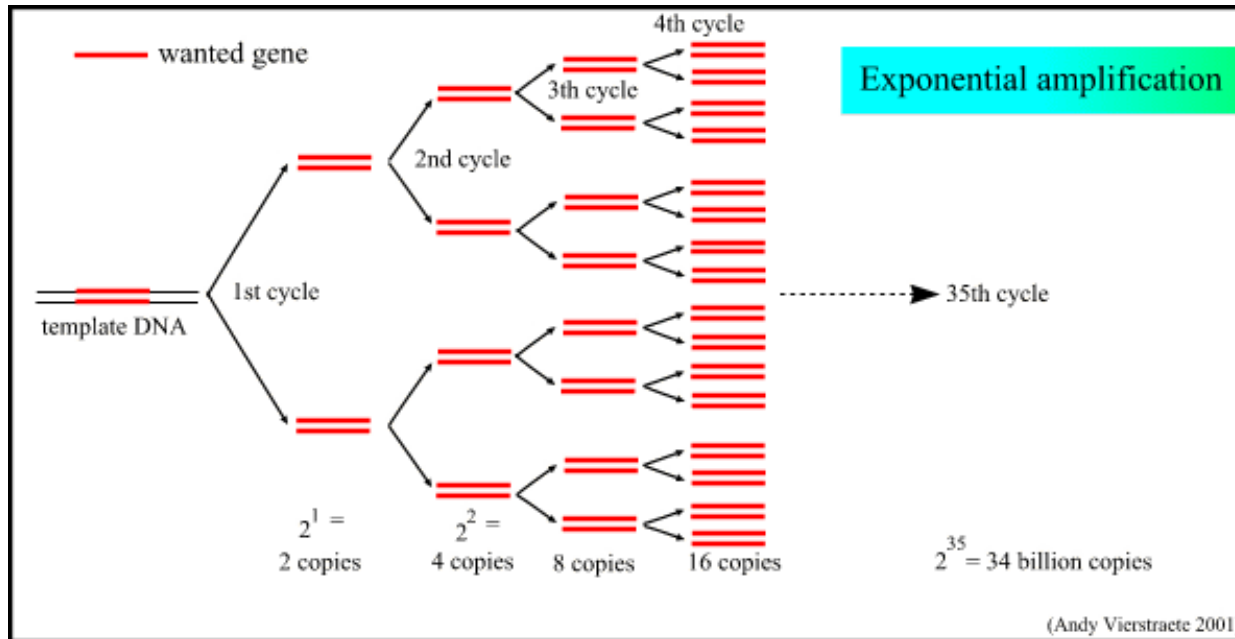
2 minutes 72 °C

**only dNTP's**

# Use of PCR - Concept of primers and/or probes



# Use of PCR - Concept of primers and/or probes



## Sensitivity

- A melting temperature ( $T_m$ ) in the range of 52 C to 65 C
- Absence of dimerization capability
- Absence of significant hairpin formation (>3 bp)
- Lack of secondary priming sites
- Low specific binding at the 3' end (ie. lower GC content to avoid mispriming)

## Specificity

- Must be **only one** target site in the template DNA where the primer binds  
=> the primer sequence shall be unique in the template DNA.
- No annealing site in possible contaminant sources, such as human, rat, mouse, unwanted bacteria, etc.

## Length

- Primer length has effects on uniqueness and melting/annealing temperature.
  - => the longer the primer, the more chance that it's unique
  - => the longer the primer, the higher melting/annealing temperature.
- Length of primer has to be at least 15 bases to ensure uniqueness.
- Usually, primers of 17-28 bases long.
  - => range varies based on if you can find unique primers with appropriate annealing temperature within this range

# Sensitivity, specificity and other thermodynamic constraints

## Base composition

- affects hybridization specificity and melting/annealing temperature.
- Random base composition is preferred.
- Avoid long (A+T) and (G+C) rich region if possible.
- Usually, average (G+C) content around 50-60% will give the right melting/annealing temperature for ordinary PCR reactions, and will give appropriate hybridization stability.

=> However, melting/annealing temperature and hybridization stability are affected by other factors (see later). Therefore, (G+C) content is allowed to change

## Annealing Temperature ( $T_{\text{anneal}}$ )

=> the temperature at which primers anneal to the template DNA. It can be calculated from  $T_m$ .

$$T_{\text{anneal}} = T_{m\_primer} - 4^{\circ}\text{C}$$



# Sensibility, specificity and other thermodynamic constraints

## Melting Temperature (Tm)

=> the temperature at which half the DNA strands are single stranded and half are double-stranded. Tm is characteristics of the DNA composition (ie. Higher G+C content DNA has a higher Tm due to more H bonds).

Basic : Marmur and Doty, 1962

$$Tm = 64.9 + 41.0 \times \left( \frac{yG + zC - 16.4}{wA + xT + yG + zC} \right)$$

Salt Adjusted : Howley et al., 1979

$$Tm = 100.5 + 41.0 \times \left( \frac{yG + zC - 16.4}{wA + xT + yG + zC} \right) - \left( \frac{820.0}{wA + xT + yG + zC} \right) + 16.6 \log([Na^+])$$

Nearest-neighbor thermodynamic : SantaLucia, 1998

$$Tm = \frac{\sum (\Delta H_d)_i + \Delta H_i}{\sum (\Delta S_d)_i + \Delta S_i + \Delta S_{self} + R \times \ln \frac{C_T}{b}} + C_{Na^+}$$

- Tm formula : <http://www.basic.northwestern.edu/biotoools/oligocalc.html>

- or see "DNAmate" (many servers available, => google)

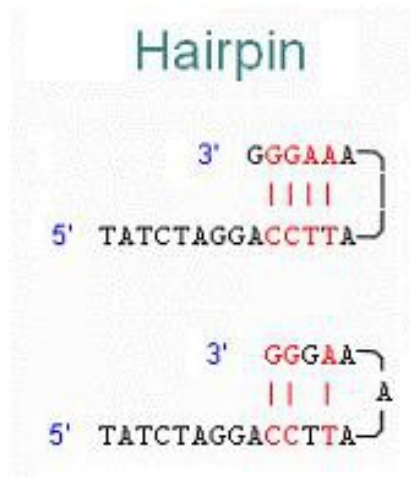
## Primer Pair Matching

- Primers work in pairs – forward primer and reverse primer. Since they are used in the same PCR reaction, it shall be ensured that the PCR condition is suitable for both of them.
- One critical feature is their annealing temperatures, which shall be compatible with each other.  
  
=> The maximum difference allowed is 3 °C. The closer their  $T_{\text{anneal}}$  are, the better.

# Sensitivity, specificity and other thermodynamic constraints

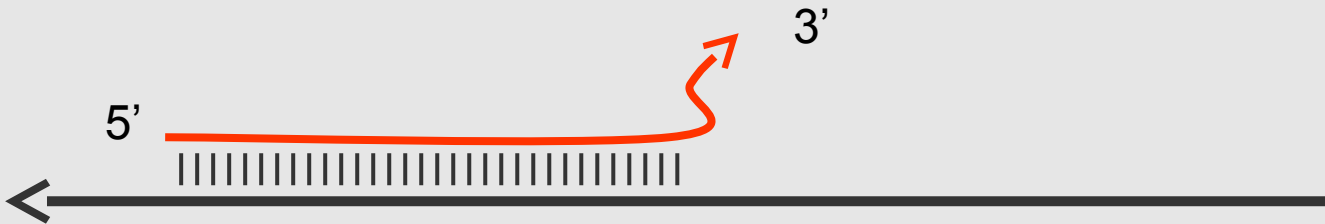
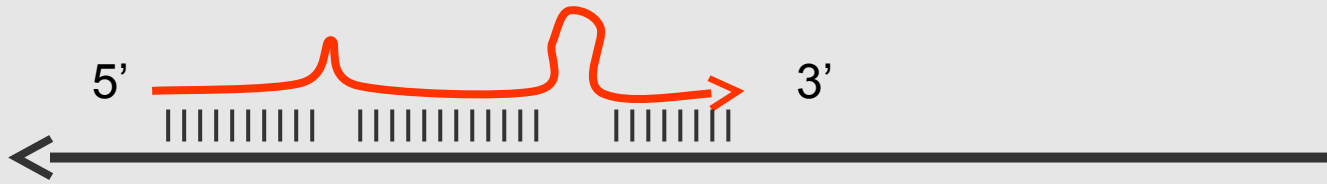
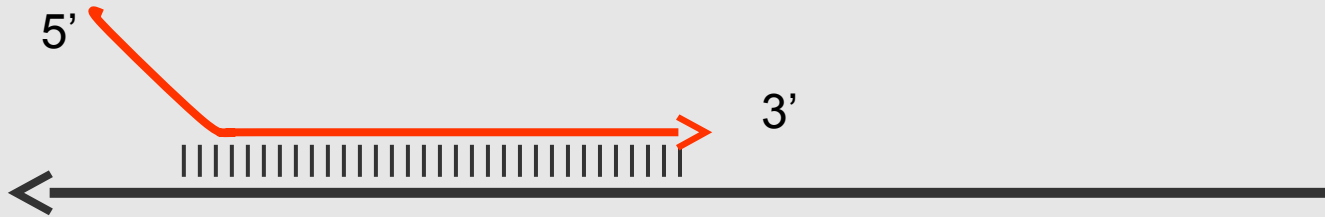
## Internal Structure (= secondary structure)

=> If primers can anneal to themselves, or anneal to each other rather than anneal to the template, the PCR efficiency will be decreased dramatically. They shall be avoided.



However, sometimes these 2° structures are harmless when the annealing temperature does not allow them to take form. For instance, some dimers or hairpins form at 30 °C while during PCR cycle, the lowest temperature only drops to 60 °C.

# Sensibility, specificity and other thermodynamic constraints



## Summary of required criteria

- Uniqueness: ensure correct priming site;
- Length: 17-28 bases. This range varies;
- Base composition: average (G+C) content around 50-60%; avoid long (A+T) and (G+C) rich region if possible;
- Optimize base pairing, in particular 3' extremity should have high stability
- Melting  $T_m$  between 55-80 °C are preferred;
- Assure that primers at a set have annealing  $T_m$  within 2 – 3 °C of each other.
- Minimize internal secondary structure: hairpins and dimmers shall be avoided.

## **Multiplex PCR**

- Multiple primer pairs can be added in the same tube to do the PCR
- Good for amplifying multiple sites
- Application example: genome identification/ finishing
- Design difficulty

Difficulties:

=> Melting temperatures should be similar

=> No dimer formulation

## Universal primers

Primers can be designed to amplify only one specific product.

Or, primers can also be designed to amplify multiple products.

=> “universal primers”. For example, design primers to amplify all 16S genes of a given clade (genera, families, etc.).

## Semi-universal primers

- Primers can be designed to amplify only a subset of template sequences from a large group of similar sequences.

For example, design primer to amplify HPV type 1 and type 6 gene, but not other types.

- Or, if a couple can not amplify all your set  
=> you may design > 2 couples to fulfill the set

# Sensibility, specificity and other thermodynamic constraints

## **Strategy for a small set of sequences:**

- Groups sequences into a set of "wanted sequences", and a set of "unwanted sequences"
- Align groups of sequences.
- Find the most conservative primers for "wanted sequences" and divergent for "unwanted sequences"
- Test forward and reverse primers to find the best pair in terms of thermodynamical constraints.
- Avoid hybridizations with unwanted sequences
  - compared alignment with your unwanted sequences set
  - or blast primers on a larger set of unwanted sequences that you may construct
  - or Blast on nt-like database, and check for unwanted (not always very efficient...)

## **Strategy for a big set of sequences (classes or phyla) :**

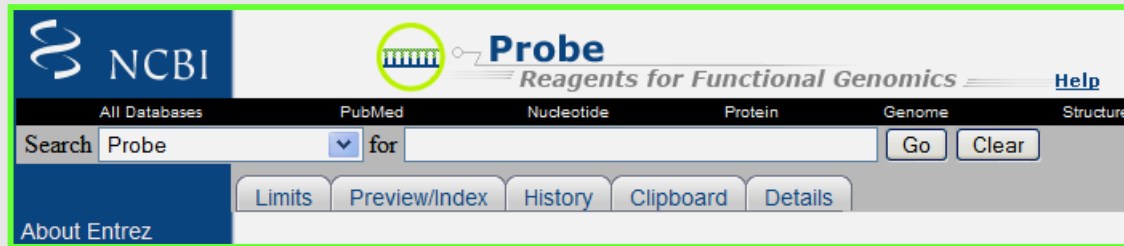
- Could be very problematic....



# Examples of softwares used for oligomers design

Before using software : do not reinvent the wheels !

- Search oligomers in publications
- Databases of primers



- NCBI Probe Database
- RTPrimerDB
- Primer Bank
- qPrimerDepot
- PCR-QPPD
- PerlPrimer

**qPrimerDepot** --- A quantitative real time PCR primer database



**RTPrimerDB**  
Real Time PCR Primer and Probe Database

> DNA Methylation analysis PCR Primer Database

Quick Search / Advanced Search

Gene   Substring  Exact phrase

Organism All   methPrimerDB ID

MASSACHUSETTS  
GENERAL HOSPITAL

CCIB  
Center for Computational & Integrative Biology

The Center for Computational  
and Integrative Biology

HARVARD  
MEDICAL SCHOOL

**Primer Bank**

PCR Primers for Gene Expression Detection and Quantification

## Examples of softwares used for oligomers design

- Primer/probes design by a skill human beings is far better done by automatic softwares !

### **Some primer design softwares:**

- Primer3 (<http://frodo.wi.mit.edu/>) - many servers existing through the web => google
- Primer3Plus (similar as Primer3)
- PrimerZ
- PerlPrimer
- Vector NTI Advantage 10
- PrimerX
- Oligo (Life Science Software)
- GCG:(Accelrys)
- Others: GeneFisher, Primer!, PrimaClase, Codhop, Web Primer, etc.

### **Some primer checker softwares and Tm calculation softwares:**

- PrimerBlast (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>)
- Oligoanalyzer (<http://eu.idtdna.com/analyzer/applications/oligoanalyzer/>)
- OHM (OligoHeatMap) (<http://bioinfo.unice.fr>)
- DNAMate (<http://melolab.org/dnaMATE/tm-pred.html>)
- BioMath (<http://www.promega.com/techserv/tools/biomath/calc11.htm>)

# Examples of softwares used for oligomers design

## Primer-Blast

**Primer-BLAST** A tool for finding specific primers

NCBI/Primer-BLAST: Finding primers specific to your PCR template (using Primer3 and BLAST). [More...](#) [Tips for finding specific primers](#)

[Reset page](#) [Save search parameters](#) [Retrieve recent results](#)

### PCR Template

Enter accession, gi, or FASTA sequence (A refseq record is preferred) [Clear](#)

Range

Forward primer  From  To  [Clear](#)

Reverse primer

Or, upload FASTA file  [Choisir...](#)

### Primer Parameters

Use my own forward primer (5'->3' on plus strand)  [Clear](#)

Use my own reverse primer (5'->3' on minus strand)  [Clear](#)

PCR product size

Min  70 Max  1000

# of primers to return  5

Primer melting temperatures (T<sub>m</sub>)

Min  57.0 Opt  60.0 Max  63.0 Max T<sub>m</sub> difference  3 [Clear](#)

### Exon/intron selection

A refseq mRNA sequence as PCR template input is required for options in the section [Clear](#)

Exon junction span  No preference [Clear](#)

Exon junction match

Exon at 5' side  7 Exon at 3' side  4

Minimal number of bases that must anneal to exons at the 5' or 3' side of the junction [Clear](#)

Intron inclusion  Primer pair must be separated by at least one intron on the corresponding genomic DNA [Clear](#)

Intron length range

Min  1000 Max  1000000 [Clear](#)

### Primer Pair Specificity Checking Parameters

Specificity check  Enable search for primer pairs specific to the intended PCR template [Clear](#)

Database  Refseq mRNA [Clear](#)

Organism  Homo sapiens

Enter an organism name, taxonomy id or select from the suggestion list as you type. [Clear](#)

[Add more organisms](#)

Exclusion (optional)  Exclude predicted Refseq transcripts (accession with XM, XR prefix)  Exclude uncultured/environmental sample sequences [Clear](#)

Entrez query (optional)  [Clear](#)

Primer specificity stringency

Primer must have at least  2 total mismatches to unintended targets, including

at least  2 mismatches within the last  5 bps at the 3' end. [Clear](#)

Ignore targets that have  6 or more mismatches to the primer. [Clear](#)

Misprimed product size deviation  4000 [Clear](#)

Splice variant handling  Allow primer to amplify mRNA splice variants (requires refseq mRNA sequence as PCR template input) [Clear](#)

[Get Primers](#)  Show results in a new window  Use new graphic view [Clear](#)

# Examples of softwares used for oligomers design

## Primer-Blast

**Advanced parameters**

**Primer Pair Specificity Checking Parameters**

Max number of Blast target sequences: 50000

Blast expect (E) value: 30000

Blast word size: 7

Max primer pairs to screen: 1000

**Primer Parameters**

PCR Product Tm	Min	Opt	Max
Primer Size	15	20	25
Primer GC content (%)	Min: 20.0	Max: 80.0	
GC clamp	0		
Max Poly-X	5		
Max 3' Stability	9		
Max GC in primer 3' end	5		

Secondary Structure Alignments:  Use Thermodynamic Secondary Structure Alignments  Use Old Secondary Structure Alignments

Warning: Thermodynamic Secondary Structure Alignments can be very slow.

Max Template Mispriming	Primer: 12.00	Pair: 24.00
Max Self Complementarity	Any: 8.00	3': 3.00
Max Pair Complementarity	Any: 8.00	3': 3.00

Excluded regions:

Concentration of monovalent cations: 50.0

Concentration of divalent cations: 1.5

Concentration of dNTPs: 0.6

Salt correction formula: SantaLucia 1998

Table of thermodynamic parameters: SantaLucia 1998

Annealing Oligo Concentration: 50.0

SNP handling:  Primer binding site may not contain known SNP

Repeat filter: Automatic

Avoid repeat region for primer selection by filtering with repeat database

Low complexity filter:  Avoid low complexity region for primer selection

**Internal hybridization oligo parameters**

Hybridization oligo:  Pick internal hybridization oligo

Hyb Oligo Size	Min: 18	Opt: 20	Max: 27
Hyb Oligo tm	Min: 57.0	Opt: 60.0	Max: 63.0
Hyb Oligo GC%	Min: 20.0	Opt: 50	Max: 80.0

## Exercice :

- **Detection of organism using PCR :**

Searching specific primers/probes for « *Bacillus anthracis* », gene « *cya* »

- **Main steps**

- Build your sets of sequences (SRS, Acnuc, NCBI, Blast)

- targets : homologues sequences of *Bacillus anthracis cya*

- non-targets : closed sequences of *Bacillus anthracis cya* (Blastn ...)

- Get primers :

- (1) existing primers : publications (using <http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed>), dedicated databases

- (2) design new ones (to be design using previous sequences sets).

- Check your primers (Blast, Oligo Heat Map, ...), and/or using alignments (Mega, ClustalW, Seaview...)