

Biochemistry 660

version 11/2016

DNASTAR'S Lasergene

Note: current lasergene version: 14

Selected webinars¹

Introducing Lasergene 14

Published on Oct 5, 2016

Techniques in biological research are constantly changing. These technological innovations fundamentally affect the depth and breadth of research scientists are able to accomplish. In this webinar, Tom Schwei will present an overview of Lasergene 14, including the introduction of three completely new applications (a genome browser, protein docking, and antibody structure prediction) as well as visualizing the power of combined data analysis from independent but linked genome assembly analysis approaches (e.g. RNA-Seq and ChIP-Seq) and other improvements in core molecular biology applications.



44:10

https://youtu.be/cm-qhvM3RL4

Older webinar: Introducing Lasergene 11: http://youtu.be/U_IexXTgdLE

Protein Sequence and Structure Analysis in Protean 3D

Presented by: Amanda Mitchell Originally Aired: November 14, 2012

Hear Amanda discuss the proteomics tools available in Lasergene 10.1 as she demonstrates some of the most powerful features in <u>Protean 3D</u>, including: epitope prediction, structural alignment, and macromolecular motion visualization.



52:07

http://youtu.be/t6vdZwktRXk

¹ http://www.dnastar.com/t-support-webinars.aspx

or

YouTube channel: http://www.youtube.com/user/DNASTARInc

Connect to Biochem. Network with VPN¹

All Biochemistry users can access the DNASTAR LASERGENE software by VPN connection when located outside of the Biochemistry buildings.

The Biochemistry Department Intranet lists all requirements and methods to access the network remotely:

https://biochem.wisc.edu/intranet/it/remote-access

Below are the steps necessary to connect from an iMac.

1. Requiements

- 1) The CISCO ANYCONNECT software² installed on the DMC computers.
- 2) Biochemistry username and password: obtain from Biochem IT or Class instructor



Connect to the Biochemistry network with the Cisco VPN software.

Locate and **double-click** on the **hard drive** Then **click on Applications** and then **Cisco**

Launch (double click) on the "**Cisco AnyConnect Secure Mobility Client.app**" which looks like this:



¹ VPN = Virtual Private Network – I allows you to "pretend" you are locally connected.

See https://kb.wisc.edu/page.php?id=27448

² Mac: https://kb.wisc.edu/page.php?id=27573

Win: https://kb.wisc.edu/page.php?id=27559

If the campus regions appear select the optic that reads:

UW Dept VPN (central campus)

If something else appears, then continue to next page.

If the campus regions do not appear by defa you will need to manually enter the server na for Central Campus:

dept-ra-cssc.vpn.wisc.ed

This choice will make **BIOCHEMISTRY** pear within a pull-down menu. Enter your Biochemistry username and password

Click Connect

| ions | ⊖ ○ Cisco AnyConnect Secure Mobility Client |
|-------|--|
| the | Connect to: UW Dept VPN (central campus) VUW Central DoIT VPN (legacy) UW Central DoIT VPN (legacy) UW Dept VPN (west campus) UW Dept VPN (central campus) UW Dept VPN (east campus) |
| | |
| - 1 | Version 3.0.11042 |
| ault, | ● ○ Cisco AnyConnect Secure Mobility Client |
| name | Connect to: dept-ra-cssc.vpn.wisc.edu v CISCO |
| | Please enter your username and password. |
| lu | Group: BIOCHEMISTRY \$ |
| | Username: |
| ap- | Password: |
| Ĩ | Connect |
| | |
| | Version 3.0.11042 |
| | O O Cisco AnyConnect Secure Mobility Client |
| | Welcome to the BIOCHEMISTRY VPN service! Unauthorized use is prohibited. While your computer is connected to this VPN service, it is subject to the UW-Madison Acceptable Use Policy (http:// www.cio.wisc.edu/policies/appropriateuse.aspx). |
| | Disconnect |
| | |
| enii | |

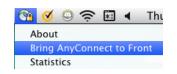
V

Click Accept

A top menu icon will appear on the Mac menu showing a locked (secure) connection.

3. Bring to Front

The top menu can be used to bring the running VPN window to the front.



*:

5

Thι

This will bring a "status" window showing the connection with information such as the IP number and how long the connection has been established.

This window can be used to disconnect from the network but it is even easier with the top menu option as shown below.

| | UW Dept VPN (central campus | ✓ cisco |
|----------------|---------------------------------------|------------|
| Connected to U | W Dept VPN (central campus). | |
| Connection Du | ration: 00:03:33 | Disconnect |
| Client Address | 10.12.22.24 | |
| Server Address | 144.92.42.108 (IPv6) Not Available | ~ |
| Bytes Sent | 27510 | |
| Bytes Received | 48512 | |

4. Disconnect

To disconnect from Departmental VPN, click on the AnyConnect VPN client icon in the top menu bar and select **Disconnect**.



<u>IMPORTANT</u>: YOU ARE LOGGED IN WITH YOUR NetID REMEMBER TO DISCONNECT AT THE END OF CLASS!

5. FYI: Windows VPN

Instructions for Windows are located here: https://kb.wisc.edu/page.php?id=27559

6. FYI: Other Campus locations

Generic instructions for the Campus are located here: https://kb.wisc.edu/page.php?id=27589

The server names for all the Departmental VPN campus regions are as follows:

East Campus: West Campus: Central Campus: dept-ra-432nm.vpn.wisc.edu dept-ra-animal.vpn.wisc.edu dept-ra-cssc.vpn.wisc.edu





VPN Connect - page - 207

Location of the Lasergene software on iMac computers

| 00 | Applications | | | |
|----------------|--------------------------------|-----------------------------------|---------|------|
| | | Q | | |
| AVORITES | Name | Date Modified | Size | Kind |
| 📃 All My Files | Dictionary.app | Mar 6, 2013 9:21 AM | 7.1 MB | Appl |
| AirDrop | 🔻 🚞 DNASTAR | Today 4:20 PM | | Fold |
| | DNASTAR Lasergene Install Log | Today 11:55 AM | 4 KB | Simp |
| Desktop | 🔻 🚞 Lasergene 11 | Today 11:55 AM | | Fold |
| 😭 dmc | CommuterLicenseManagerLG11.app | Today 11:55 AM | 11 MB | App |
| Applications | DM11 | Sep 12, 2013 3:06 PM | 4.6 MB | Unix |
| Documents | 🔠 EditSeq.app | Today 11:55 AM | 7.4 MB | App |
| Documents | 😭 Enzymes.ezd | Mar 4, 2013 12:30 PM | 27 KB | Enzy |
| ARED | 뎙 EnzymeSelectors.sel | Mar 4, 2013 1:30 PM | 33 KB | Enzy |
| EVICES | 💹 GeneQuest.app | Today 11:55 AM | 10 MB | App |
| | Q GenVision Utility.app | Today 11:54 AM | 2.4 MB | App |
| BOUTCAMP | 📕 GenVision.aip | Today 11:54 AM | 2.5 MB | Ado |
| | 🐲 LicenseManagerLG11.app | Today 11:55 AM | 15.5 MB | App |
| | < MegAlign Pro.app | Today 11:54 AM | 74.5 MB | App |
| | 🧱 MegAlign.app | Today 11:55 AM | 13.3 MB | App |
| | PrimerSelect.app | Today 11:55 AM | 9.4 MB | App |
| | 🞯 Protean 3D.app | Today 11:54 AM | 85.6 MB | App |
| | Protean.app | Today 11:54 AM | 10.6 MB | Арр |
| | SeqBuilder.app | Today 11:54 AM | 25.1 MB | App |
| | 🔲 SeqMan.app | Today 11:54 AM | 17.8 MB | App |
| | 😿 SeqNinja.app | Today 11:54 AM | 78.8 MB | Appl |
| | 🕨 🚞 Lasergene 11 Data | Today 11:54 AM | | Fold |
| | 📄 Lasergene 11 Install Log | Today 11:55 AM | 1.4 MB | Simp |

All Lasergene software is located in the directory /Applications/DNASTAR



Book 2: Bioinformatics with DNASTAR Lasergene





adison-based D N A S T A R (www.dnastar.com) is an international company developing software since 1984. The company now defines itself as "... a pioneer in the development and sale of software used to increase life scientists' productivity using their desktop computer or on the Amazon Cloud. DNASTAR's comprehensive software suite, Lasergene, supports molecular biologists, geneticists, and structural biologists in meeting virtu-

ally all of their DNA, RNA, and protein sequence needs, including Sanger and next- generation sequence assembly and analysis, protein sequence and structure analysis, and protein structure prediction with easy to use, affordable, flexible computer software."

The following computer hands-on lab modules are self-paced exercises demonstrating some *Lasergene* applications on sequence editing (**EditSeq**), plasmid mapping (**SeqBuilder**) and protein structure (**Protean** and **Protean 3D**.)

We are grateful to D N A S T A R for providing academic teaching licenses when necessary, and to the CALS computer lab personnel for their help in installing the software onto the classroom computers.

UW Biochemistry students can access D N A S T A R software on the teaching computers (licensed via the Biochemistry Department) in room 301 on the 3rd floor of the Biochemistry Laboratories building. **Biochemistry Students** can install the *Lasergene* package following instructions on the Biochemistry <u>Intranet</u> (*NetID login required*):

https://biochem.wisc.edu/intranet/it/lasergene

Other users interested in installing the trial version of *Lasergene* can request a free trial. Simply click on the blue button **Free Trial** on their web home page (www.dnastar.com).

All materials contained herein are intended for educational purposes only.



The standard genetic code.

The table shows the 64 codons and the amino acid for each. The mRNA is 5' to 3'

| nor | npolar p | polar basic | acidio | c (stop codo | n) | | | | |
|------|--------------------|------------------------------|--------|----------------|-----|------------------------------|-----|---|------|
| 1st | | | | | 2nc | l base | | | 3rd |
| base | • | U | | С | | Α | | G | base |
| | UUU | (Phe/F) | UCU | | UAU | (Tyr/Y) | UGU | (Cys/C) Cysteine | U |
| U | UUC | <u>Phenylalanine</u> | UCC | (Ser/S) | UAC | <u>Tyrosine</u> | UGC | (Cys/C) <u>Cysteme</u> | С |
| U | UUA | | UCA | Serine | UAA | Stop (Ochre) | UGA | <u>Stop</u> (Opal) | Α |
| | UUG | | UCG | | UAG | <u>Stop</u> (Amber) | UGG | (Trp/W) <u>Tryptophan</u> | G |
| | CUU | (Leu/L) | CCU | | CAU | (II:a /II) II:atidina | CGU | | U |
| С | CUC | Leucine | CCC | (Pro/P) | CAC | (His/H) <u>Histidine</u> | CGC | $(\Lambda_{\rm HC}/{\rm B})$ $\Lambda_{\rm HC}$ in in (| С |
| C | CUA | | CCA | <u>Proline</u> | CAA | (Gln/Q) Glutamine | CGA | (Arg/R) <u>Arginine</u> | Α |
| | CUG | | CCG | | CAG | (Gin/Q) <u>Giutannie</u> | CGG | | G |
| | AUU | (T1 /T) | ACU | | AAU | (Asp /NI) Asparazion | AGU | (Ser/S) Serine | U |
| | AUC | (Ile/I) Isoleucine | ACC | (Thr/T) | AAC | (Asn/N) <u>Asparagine</u> | AGC | (3ci/3) <u>3ciiiic</u> | С |
| Α | AUA | <u>130iedenie</u> | ACA | Threonine | AAA | | AGA | | Α |
| | AUG ^[A] | (Met/M) <u>Methionine</u> | ACG | | AAG | (Lys/K) <u>Lysine</u> | AGG | (Arg/R) <u>Arginine</u> | G |
| | GUU | | GCU | | GAU | (A (D) A | GGU | | U |
| G | GUC | (Val/V) | GCC | (Ala/A) | GAC | (Asp/D) <u>Aspartic acid</u> | GGC | (C_{1}) (C_{1}) (C_{1}) | С |
| G | GUA | Valine | GCA | Alanine | GAA | (Glu/E) Glutamic acid | GGA | (Gly/G) <u>Glycine</u> | Α |
| | GUG | | GCG | | GAG | (Giu/E) Glutamic acid | GGG | | G |

^A The codon AUG both codes for methionine and serves as an initiation site: the first AUG in an <u>mRNA</u>'s coding region is where translation into protein begins. [Nakamoto T (March 2009). "Evolution and the universality of the mechanism of initiation of protein synthesis". Gene. 432 (1-2): 1-6. doi:10.1016/j.gene.2008.11.001.PMID 19056476.]

| Ala/A | GCU, GCC, GCA, GCG | Leu/L | UUA, UUG, CUU, CUC, CUA, CUG |
|-------|------------------------------|-------|------------------------------|
| Arg/R | CGU, CGC, CGA, CGG, AGA, AGG | Lys/K | AAA, AAG |
| Asn/N | AAU, AAC | Met/M | AUG |
| Asp/D | GAU, GAC | Phe/F | UUU, UUC |
| Cys/C | UGU, UGC | Pro/P | CCU, CCC, CCA, CCG |
| Gln/Q | CAA, CAG | Ser/S | UCU, UCC, UCA, UCG, AGU, AGC |
| Glu/E | GAA, GAG | Thr/T | ACU, ACC, ACA, ACG |
| Gly/G | GGU, GGC, GGA, GGG | Trp/W | UGG |
| His/H | CAU, CAC | Tyr/Y | UAU, UAC |
| lle/l | AUU, AUC, AUA | Val/V | GUU, GUC, GUA, GUG |
| START | AUG | STOP | UAG, UGA, UAA |

Source: http /en.wikipedia.org/wiki/Genetic

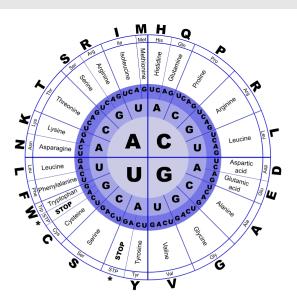
L02: Lasergene part 1

Editseq / SeqBuilder / Protean / Protean3D

1. Table of Contents

| Book 2: Bioinformatics with DNASTAR Lasergene 2 | :09 |
|--|-----|
| The standard genetic code 2 | 10 |
| L02: Lasergene part 1 2 | 11 |
| Editseq / SeqBuilder / Protean / Protean3D 2 | 11 |
| 1. Setup:2 | 13 |
| 2. Lasergene modules | 13 |
| 3. EditSeq and NCBI Sequences2 | 14 |
| 3.1 Launching EditSeq2 | |
| 3.2 Importing sequence from NCBI2 | |
| 4. Sending sequence to SeqBuilder2 | |
| 4.1 Sending a sequence from EditSeq to SeqBuilder | |
| 4.2 Create and display an enzyme restriction map2 | |
| 4.3 Restriction enzymes: linear map2 | |
| 4.4 Restriction enzymes: circular map2 | |
| 4.5 Restriction enzymes: choosing individual enzymes2 | |
| 4.6 Minimap: a linear outlook of restriction sites2 | |
| 4.7 File saving from SeqBuilder | |
| 5. SeqBuilder Features Plasmid Maps2 | |
| 5.1 Creating a new Protein Translation Feature2 | |
| 5.2 Creating a new Graphical Feature2 | |
| 5.3 Editing the Graphical Features of a demo file | |
| 5.4 Quit SeqBuilder | |
| 6. Protean – Protein Analysis | |
| 6.1 Opening an Entrez protein sequence2 | |
| 6.2 Protean Default View and Methods2 | |
| 6.3 Enzyme and chemical digestion and SDS PAGE Gel simulations | 232 |
| 6.4 Composition | 235 |
| 6.5 Titration Simulation | |
| 6.6 Model Structure Simulations: Helical Wheel2 | |
| 6.7 Patterns Search: Prosite | 237 |
| 6.8 Help - Patterns Search: user-specified descriptor (Ariadne File) | |
| 6.9 Quit Protean | |
| 7. Protean: Secondary Structure Predictions2 | |
| 7.1 Open the Protean software2 | |
| 7.2 Open a protein sequence and display predictions | |
| 7.3 Recalculating graphics (changing defaults)2 | |
| 7.4 Recalculating options from the methods panel (left) | 243 |
| 7.5 Reorganizing the panel plots2 | |
| 7.6 More methods | 244 |

| 7.7 Changing colors and patterns of plots | |
|--|--|
| 7.8 Tabular Data | |
| 7.9 Reality check | |
| 8. Protean 3D | |
| 8.1 Try Protean3D | |
| 8.2 Analysis panel | |
| 8.3 Learn more from DNASTAR tutorials and videos | |
| 9. Quit All Programs | |
| L02 – Supplemental: Primer Design | |
| 1. L02 - Setup: | |
| DNASTAR PrimerSelect | |
| 2.1 Launch PrimerSelect | |
| 2.1.1 PrimerSelect Help | |
| 2.2 Importing a sequence | |
| 2.3 Limiting the Annealing Region | |
| 2.4 Calculating primer pairs: | |
| 2.5 Evaluating primer pairs: menu "Report" | |
| 2.6 Locating Primers | |
| 2.7 Upper and lower primers | |
| Designing Primers in SeqBuilder | |
| 3.1 Creating and Modifying Primer Pairs for a Region of Interest | |
| 4. Primer3, free web-based application | |
| 4.1 Launch web browser | |
| 4.2 Enter data | |
| DNASTAR Resources and Support | |
| 6. Ending session | |



http://all-free-download.com/free-vector/vector-clip-art/genetic_code_rna_54619.html

1. Setup:



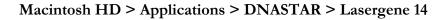
Create a folder named LO2 on the desktop. This will be used to save files created during this session.

| | File | Edit | View | Go | Window | F |
|----|------------|----------|------|----|--------|---|
| | Nev | w Find | ж | N | | |
| it | Nev | w Fold | er | | ひ 第1 | N |
| | A 1 | <u> </u> | | | 37.001 | |

2. Lasergene modules

Lasergene is composed of modules. Most modules can share sequence files simultaneously. The suffix .app may or may not appear on your system depending on system preferences.

The modules are located within the folder:



| EditSeq.app | GeneQuest.app | GenVision Pro.app | GenVision Utility.app | |
|---------------------|----------------|-------------------|--------------------------|-----------------------------------|
| MegAlign Pro.app | MegAlign.app | PrimerSelect.app | Protean 3D.app | Lasergene modules |
| Protean.app | SeqBuilder.app | SeqMan Pro.app | SeqNinja.app | |
| SeqMan Pro | | Seqi | uence Assembly | and Contig Management |
| GeneQuest | | Gen | e Discovery, Šec | nuence Annotation, Publication |
| Protean | | Proi | ein Structure D | iscovery, Annotation, Publication |
| Protean3D | | New | v module for exp | loring 3D PDB files |
| MegAlign | | Mu | ltiple and Pairw | ise Sequence Alignment |
| PrimerSelect | | Olig | onucleotide Desi | ign and Analysis |
| SeqBuilder | | Edi | t and format. M | aps and Plasmid maps. |
| EditSeq | | Imp | orting and Edit. | ing Primary Sequence Data |
| GeneVision | Pro | geno | mic visualizatio | n application |

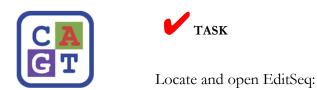
Note: some Lasergene modules only run in Windows (e.g. ArrayStar)

3. EditSeq and NCBI Sequences

The National Center for Biotechnology Information (NCBI) is the main repository for sequences. Lasergene modules can retrieve sequences directly from the NCBI server called "Entrez" (which is French for "Come in!")

EditSeq as well as SeqBuilder, PrimerSelect, and GeneQuest can import NCBI /Entrez sequences directly.

3.1 Launching EditSeq



Macintosh D > Applications > DNASTAR > Lasergene 12 > EditSeq

This will open an "Untitled Seq #1: SEQUENCE" blank window with 2 panes, the bottom pane showing the time of creation and the top pane, now blank, will contain the sequence.

| 📹 EditSeq | File | Edit | Search | Speech | Features | Goodies | Net Search | Window | Help |
|---------------------|-----------|-------|----------|-------------|--------------|-----------|------------|--------|------|
| 00 | | | _ | Untitled Se | eq #1 : SEQU | IENCE | | _ | |
| Position | 1:1 | | 20 | 10 | Fo | <i>co</i> | | 00 | O bp |
| 10 Ll. | 20 | | 30 | 40 | 50 | 60 | 70 80 | 90 | |
| | _ | | | | | | | | 0 |
| - | Se | 01 | lenc | eu | sill a | DDea | r here | 0 - | |
| | | 8 | | | | 7-00 | | | |
| | | | | | | | | | _ |
| Created: Tuesday | , Sept | ember | 13, 2011 | 10:24 AM | | | | | |
| | 0 | | | 1 - | .11 | | | | |
| | $C \circ$ | m | nent | cs u | | ppel | v her | e - | |
| | | | | | | | | | |
| 📲 🏵 🔍 🖉 Unspecified | Search | | | | | | - | | |

PRACTICAL: To create a new sequence file from your own sequence data you could simply paste the relevant sequence within the top panel and save the file for later use.

3.2 Importing sequence from NCBI

V TASK

- Within the File menu select the option
 Open Entrez Sequence... or use the keyboard menu shortcut **%**R
- Within the new window **enter** the sequence accession code **M77811**
- **Change** the default look-up area from genome to **nucleotide**
- Click OK
- To change the default download area click on the "Where" pull-down menu and select L02 on the Desktop.

| New | • | |
|------------------------------------|--------|------------|
| Open | жo | |
| Import | | |
| Open Entrez Sequence | ЖR | |
| | | |
| Entrez Server: | | |
| http://www.ncbi.nlgov/entrez/eutil | s/ | ۶ |
| | | |
| Enter Sequence ID or Locus Name | | |
| M77811 | | |
| ✓ genome | | |
| nucleotide | | |
| protein cel |) (-0) | \bigcirc |
| | | |
| | | |
| Save | | _ |
| | 6- | |

File Edit Search Speech Fe

| Save As | : M77811.seq | |
|---------|---------------------------|------|
| Where | ✓ 🙆 Downloads | |
| | 🔜 Macintosh HD 🔮 iDisk | Save |

Note: The default location for downloads on a modern Macintosh is the "downloads" directory (located within the DMC username.)

Above we created a directory called L02 to contain files for these exercises.

If L02 is not visible on the pull-down menu, click on the top right triangle within the blue square to expand the view (circled in above image.)

This will always allow you to navigate to the correct location.

• **Click Save** to save the file within the L02 directory. It will be named M77811.seq by default.

🗸 READ

- The original file will now be parsed into 3 separate panels:
 - <u>Top</u> pane: Actual sequence in the same case as on NCBI (lower case.)
 - <u>Middle</u> pane: definitions from the header.
 - <u>Bottom</u> pane: special notes and references that can be edited.
- Each panel has a vertical and a horizontal scroll bar.

🗸 INFO

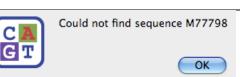
• The computer can read the sequence aloud for proofreading. Of course this is most useful when verifying short sequences.

The first 3 lines of the middle pane read:

```
LOCUSSYNBLUEV2746 bpDNAcircular SYN 16-MAR-2000DEFINITIONBlueScribe cloning vector.ACCESSIONM77811 M77798VERSIONM77811.1 GI:208035
```

This sequence is that of a cloning vector, last revised in the year 2000.

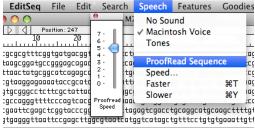
 Both accession number M77811 & M77798 are listed, but M77798 is obsolete and replaced by M77811. NCBI keeps version tracks.



This cloning vector was engineered from different sequences, listed in the "source" section within the middle and bottom panes.

DO NOT CLOSE the window, we will use this sequence in the next exercise!

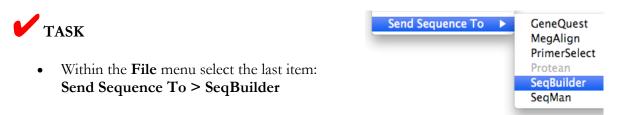




4. Sending sequence to SeqBuilder

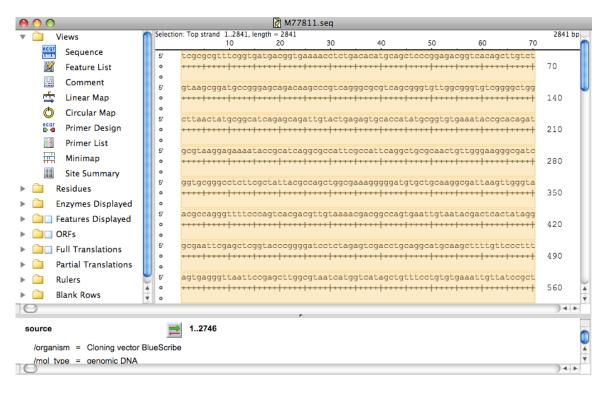
Modules can send sequences to each other without the need to save the file and reopening it within the other module.

4.1 Sending a sequence from EditSeq to SeqBuilder



If SeqBuilder is not yet running, it will open automatically and load the new sequence.

When SeqBuilder opens, the default view is that of the sequence within a window with a side window showing viewing options.



SeqBuilder can be used to edit and display sequences. Many optional views can be seen simultaneously: Sequence view, Feature List view, Comment view, Linear map, Circular map, Primer Design view, Primer List view, Minimap, and Site Summary.

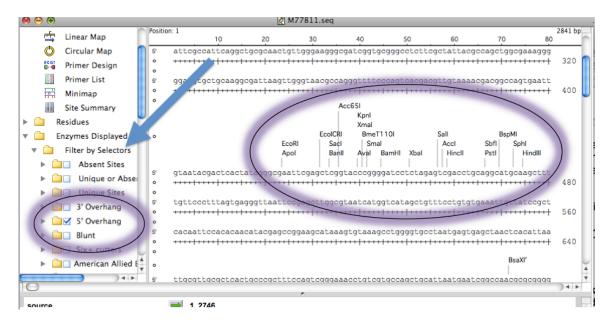
We will use the display capabilities to create an enzyme restriction map.

4.2 Create and display an enzyme restriction map

TASK

- Within the left panel **click on** the small triangle to expand both directories: **Enzymes Displayed** and **Filter by Selectors**
- Click on 5'Overhang
- If necessary scroll down the left panel.
- When the 5' Overhang is clicked, scroll down on the right panel if the window is too small. The area between bases 400-500 is the polylinker where many enzymes are unique cutters.
- The polylinker enzymes are listed within the header of the NCBI file:

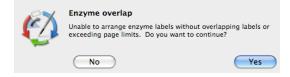
POLYLINKER HindIII-SphI-PstI-SalI-XbaI-BamHI-SmaI-KpnI-SacI-EcoRI



4.3 Restriction enzymes: linear map

V TASK

- On the left panel click on Linear Map
- Click Yes after the warning if any.



• Click on the magnifying glass to zoom in the desired area around 400.

| | • | | | |
|------------------|----------------------------|--------------------|---------------|----------------|
| 0 0 | | ● 🕅 M77811.seq | | |
| Views | Position: 1 Magnificati | an: 1.0x | | |
| Sequence | 20011 Out Mode | Kasl ARCHEIDI 101 | | |
| 📓 🛛 Feature List | | Tarl Bishillish MI | BsaXI' | |
| Comment | Ndel Pfol BstA | | BsaXI Sapl | Pcil Afilli |
| 📥 🛛 Linear Map | | | | |
| O Circular Map | 200 | 400 | 600 800 | |

A magnification of around 24 to 36 should provide a better view:

| | | Rositic Magni | on: 1 fication: 2 | 6x | | | | | | | | 841 bp e 1 of 1 |
|---|-----|------------------|----------------------|----------|-------|------|-----|--------|------|-------|---------|--------------------|
| l | | | Acc65I | BmeT | 110 | | S | all | Pstl | | | |
| | | EcolCRI | | Kpnl Sma | al | | | Accl | | BspMI | | |
| | | Sac | | Xmal | BamHI | | | Hincll | | | Sphl | |
| | | Ban | I | Aval | | Xbal | | | Sbfl | | HindIII | |
| ŀ | 430 |)) | 435 | 440 | 445 | 450 | 455 | 460 | 465 | 470 | 475 | 480 |
| | | | | | | | | | | | | |

4.4 Restriction enzymes: circular map

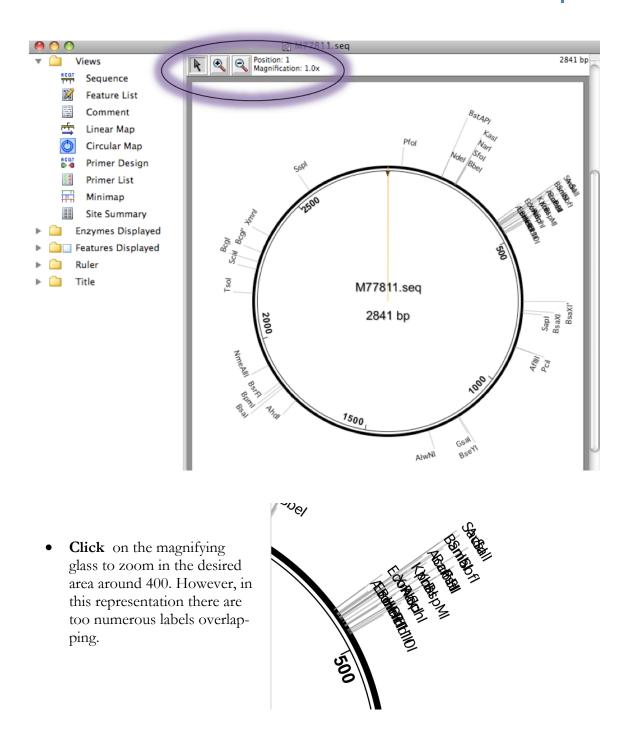


- On the left panel click on **Circular Map**
- Click Yes if warning appears

On the left panel, under Enzymes Displayed / Filter by Selectors:

Click "Unique Sites"





4.5 Restriction enzymes: choosing individual enzymes

Enzymes can be chosen individually from an alphabetical list.





- <u>Un</u>select the 5' Overhang and/or Unique Sites
- Close the Filter by Selectors tab
- Open tab All Enzymes Alphabetical



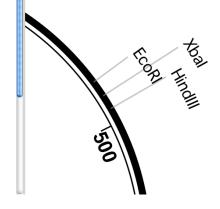
The next 2 tabs are AarI..FaII and Fall'..ZraI containing all enzymes in alphabetical order.

For the sake of example we will select 3 enzymes present in the linker to be displayed: Eco-RI, HindIII and XbaI

🔻 🚞 🗹 EcoNI..Fall

- **Click** the tabs to uncover the 3 enzymes alphabetically.
- This image shows the final result as well as the left panel with 2 of the enzymes visibly selected.





4.6 Minimap: a linear outlook of restriction sites

Another interesting way to visualize and find cutting enzymes is via the minimap view.



- From the Enzymes Displayed > Filter by Selectors button on the left panel
- Select Unique Sites
- Click Yes if warning appears.



- On the left panel **Click on Minimap**
- Regardless of the magnification within the previous view, it will be reset to a magnification of 1.0x within the Minimap view

| ц. | Linear Map |
|------------|---------------|
| \bigcirc | Circular Map |
| RCGT | Primer Design |
| | Primer List |
| | Minimap) |
| | |

The default Minimap view will show the list of cutting enzymes in alphabetical order, followed by the cutting frequency and an horizontal line symbolizing the sequence location with a blue vertical line locating the enzyme cutting position.

| 00 | | | | R | M77811.s | eq | | | |
|-------|------------------|--------|----------|----------------|----------|-------|------|-------|---------|
| 🔻 🚞 V | /iews | R 🔍 | | Magnification: | 1.0x | | | | 2841 bp |
| RCGT | Sequence | Name | Freq | | 1500 | 11000 | 1500 | 12000 | 12500 |
| 2 | Feature List | Acc65I | 1 | | | | | | |
| | Comment | | <u> </u> | 1 | | | | | |
| ц. | Linear Map | Accl | 1 | | | | | | |
| Ó | Circular Map | Aflii | 1 | | | | | | |
| RCGT | Primer Design | Ahdl | 1 | | | | | | |
| | Primer List | AlwNI | 1 | | | | | | |
| | Minimap | Apol | 1 | | | | | | |
| | Site Summary | Aval | 1 | | | | | | |
| 🕨 🧰 E | nzymes Displayed | | <u> </u> | | | | | | |
| | | BamHI | 1 | | | | | | |
| | | Banll | 1 | | | | | | |
| _ |) 4 + | Bbel | 1 | | | | | | A V |
| 10 | | | | | <u>,</u> | | | |) 4 ► |

The list order can be changed by clicking within the table header. For example, to see the cutting as a function of the sequence position:

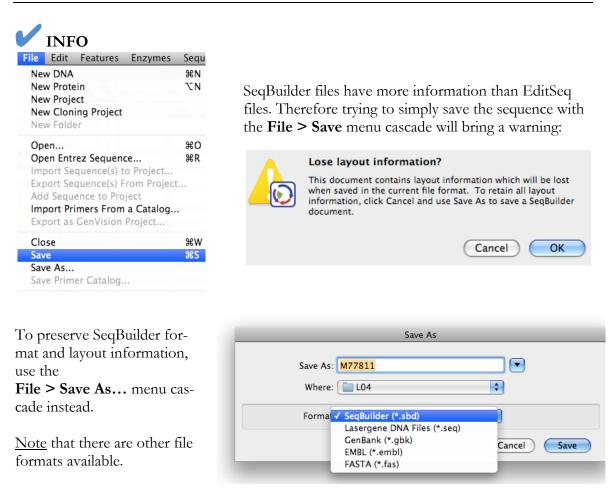


- Click on the sequence number line at the top
- Scroll down to find the polylinker area (between 4 and 500)

The view will now be be reset to the new settings.

| | / | Magnification: | | | | | |
|----------|-----|----------------|------|------|------|-------|------|
| Name | req | | 1500 | 1000 | 1500 | 12000 | 2500 |
| Sspl | 1 | | | | | | |
| Apol | 1 | | | | | | |
| EcoRI | 1 | | | | | | |
| EcolCRI | 1 | | | | | | |
| Banll | 1 | | | | | | |
| Sacl | 1 | | | | | | |
| Acc65I | 1 | | | | | | |
| Aval | 1 | | | | | | |
| Kpnl | 1 | | | | | | |
| Xmal | 1 | | | | | | |
| BmeT110I | 1 | | | | | | |
| Smal | 1 | | | | | | |
| BamHI | 1 | | | | | | |

4.7 File saving from SeqBuilder



5. SeqBuilder Features Plasmid Maps

5.1 Creating a new Protein Translation Feature

This exercise section uses the same Bluescribe vector file as in the previous exercise: M77811.seq.

The Bluescribe vector contains an open reading frame on the negative strand, coding for ß-Galactosidase (LacZ) which is used for selecting clones.

In the following exercise we will tell SeqBuilder where the sequence is and create a graphical feature to highlight the protein on the circular plasmid map.

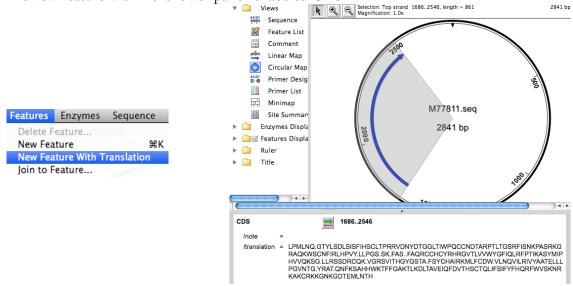


- Open SeqBuilder
- File > Open... (or **#O**) and open file M77811.seq saved previously (*Alternatively use the File > Open Entrez Sequence... as we have seen in a previous exercise.*)
- Click on Circular Map on the left panel.
- Select Edit > Go To Position and enter the range 1686..2546 (the corresponding region will become highlighted within the circular map. Note the 2 dots between the numbers.)

| Edit Features | Enzymes | Sequen |
|-------------------------------|------------|-----------|
| Undo Hide Fea Redo Show En | | 策Z 公策Z |
| Cut | zynie cuts | #X |
| Сору | | #C |
| Copy as Pictur | e | |
| Copy As | | ► |
| Paste | | жv |
| Delete Select All | | ЖA |
| Rename | | 5074 |
| Find | | ЖF |
| Find Again | | ЖG |
| Go to Position | | жJ |

• Select **Features > New Feature With Translation** and note how a new, blue arrow will appear at the same time as a new feature entry within the bottom panel together with a new translation. However, the translation is currently wrong because the coding sequence is in fact on the *opposite strand* and therefore we will need to make a few changes.

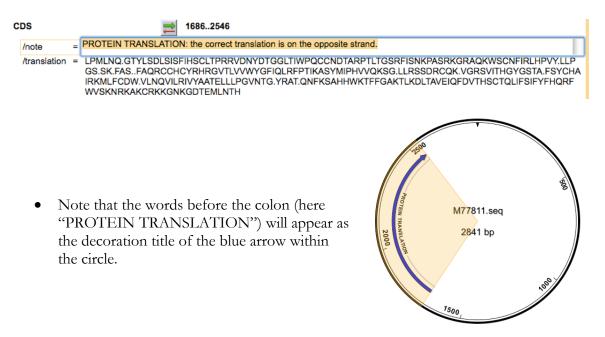
The new feature within the lower panel is labeled CDS.



TASK

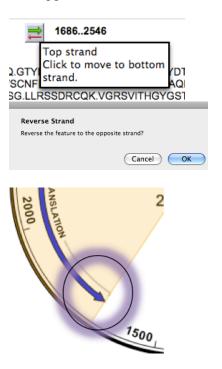
• Within the lower panel click after the = sign on the /note keyword and enter a reminded that the translation is on the other strand: e.g.

"PROTEIN TRANSLATION: the correct translation is on the opposite strand."



- In the bottom panel: Hover the mouse on the grey square with the green arrow next to CDS and <u>wait long enough</u> for the instructions to appear.
- **Click on the square** to change to the opposite strand.
- The red arrow will become larger and the translation text will change from black to red.
- **Click OK** in the warning box

• <u>Note</u>: at the same time the blue arrow within the circle will point in the opposite direction



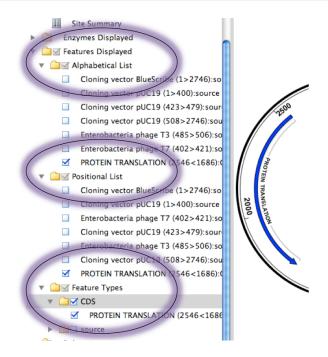
• Select menu **Features > New Translation** and a new translation, in black color will appear. This new translation is from the opposite strand and is now the correct protein sequence!

| Features Enzymes Sequence | CDS | 16862546 |
|--|--------------|---|
| Delete Feature | /note | = PROTEIN TRANSLATION: the correct translation is on the opposite strand. |
| New Feature % New Feature With Translation Join to Feature | /translation | EVPMLNQ.GTYLSDLSISFIHSCLTPRRVDNYDTGGLTIWPQCCNDTARPTLTGSRFISNK RAQKWSCNFIRLHPVYLLPGS.SK.FAS.FAQRCCHCYRHRGVTLVVWYGFIQLRFPTI HVVQKSG.LLRSSDRCQK.VGRSVITHGYGSTA.FSYCHAIRKMLFCDW.VLNQVILRIV* PGVNTG.YRAT.QNFKSAHHWKTFFGAKTLKDLTAVEIQFDVTHSCTQLIFSIFYFHQRF* KAKCRKKGNKGDTEMLNTH |
| New Note New Translation | /translation | MSIQHFRVALIPFFAAFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRP MSTFKVLLCGAVLSRIDAGQEQLGRRIHYSQNDLVEYSPVTEKHLTDGMTVRELCSAA TAANLLLTTIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVAMATTLF LLTLASRQQLIDWMEADKVAGPLLRSALPAGWFIADKSGAGERGSRGIIAALGPDGKP TGSQATMDERNRQIAEIGASLIKHW |

INFO - Note that the protein translation feature (the blue arrow and its named decoration) appears as a highlighted gray or yellow arc within the circle while it is selected. Similarly, it is possible to graphically select this feature by clicking on either the blue arrow or the decoration's name "PROTEIN TRANSLATION" within the circle.

In addition the new feature "CDS" was added within the "Features Displayed" list on the left side panel in all 3 categories: Alphabetical List, Positional List, and Feature Types.

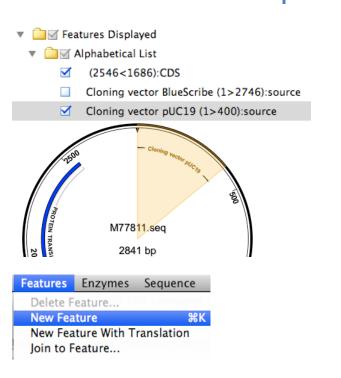
Since this vector was created from various sources, these list indicate the origin of the sequences and can be toggled on and off.



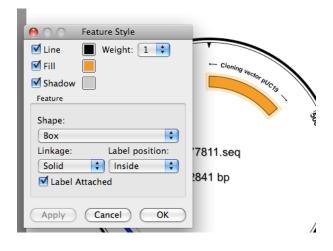
5.2 Creating a new Graphical Feature

New graphical features can be added for features listed within the Genbank file. As an example we will add an arc-circle for the sequence segment corresponding to the pUC19 cloning vector:

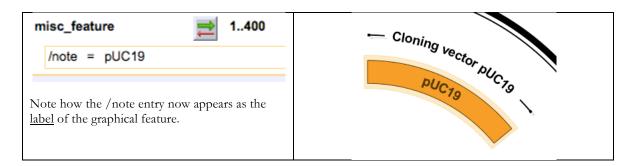
- Within either section of the "Features Displayed" (*e.g* under "Alphabetical List") click on "Cloning vector pUC19 (1>400):source"
- Click on the new description within the circle to select this area. The arc corresponding to this area will become highlighted.
- Within the menu bar select Features > New Feature or use the %K shortcut.



- A new thick arc will appear within the circle as the graphical representation of this area in the default position: below the text, as an orange arc.
- Select the menu cascade: Features > Edit Feature Style
- Unselect Shadow
- Click Apply
- Click OK



- Select menu cascade Features > Edit Feature Description
- This will bring the bottom panel section into focus with a default /misc_feature keyword and a /note keyword entry as well.
- Within the /note section type "pUC19"



5.3 Editing the Graphical Features of a demo file

SeqBuilder files end with the .sbd filename extention. We will look briefly at 2 files located within the demo data of the Lasergene 8 package.

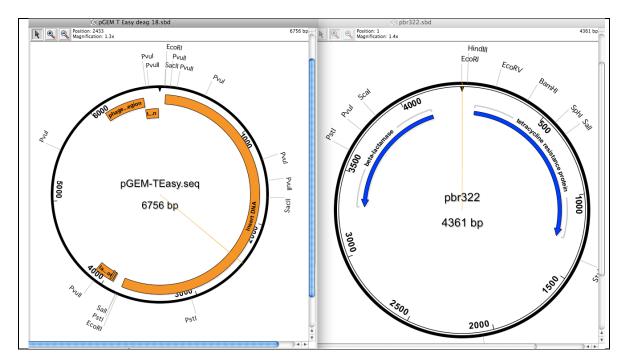
SeqBuilder should be running from the previous exercise segment. If not, launch SeqBuilder.

TASK Using the **File > Open** menu cascade open successively the following 2 files:

pGEM T Easy deag 18.sbd

pbr322.sbd

both located within /Applications/ DNASTAR/Lasergene 12 Data/Demo Data/

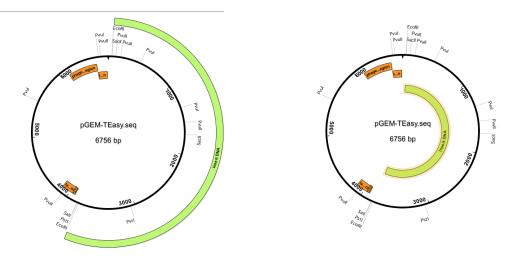


As an exercise we will change the color of the "insert DNA" feature as well as its position within the circle:

- Double-click on the "insert DNA" feature within the circle
- Select menu Features > Edit Feature Style
- Click on the Fill color square
- Select a new color e.g. light green (as shown here)
- Click OK (within the "colors" window)

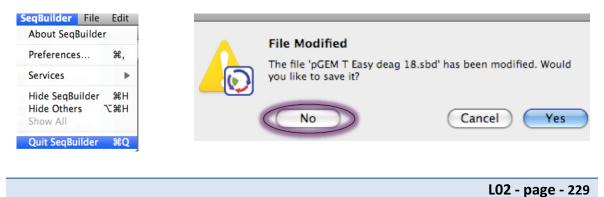
| r Design r L O O Colors hat I Hat | 6000 phage egion | |
|---|--|------------|
| un Q / | O O O Feature Style | |
| st ;D | ✓ Line Weight: 1 ↓ ✓ Fill Shadow Feature | |
| | Shape: | |
| | Box | Insert DNA |
| | Linkage: Label position: | Lieg |
| | Solid 🗘 Inside | |
| I | ⊡ Label Attached | |
| | Apply Cancel OK | |
| Cancel OK | | |

- Click Apply and then Click OK (within the "Feature Style" window)
- **Click and hold** the selected **graphical feature segment** and while holding the mouse button down **slide the feature either** to the outside or the inside of the circle to obtain *either* of these representations:



5.4 Quit SeqBuilder

TASK Quit the SeqBuilder program: **SeqBuilder > Quit SeqBuilder** or use the **#Q** shortcut. Do not save the changes made to the demo files (Answer NO when asked if you want to save changes)



6. Protean – Protein Analysis

TASK Open the Protean module located within /**Applications/Classes/DNASTAR/Lasergene 12**/

A logo will flash and seemingly not much else. However, note that the top menu has switched from the finder menu to the Protean menu:

| d Protean File Edit Analysis Sites & Features Options Net Search Window | н |
|---|---|
|---|---|

6.1 Opening an Entrez protein sequence

Protein sequence can be either imported from Entrez, or opened from a file. Proteins sequences files end with .pro in Lasergene.

File

New...

Edit

| Т | ASK |
|---|--|
| ٠ | Select the menu File > Open Entrez Protein |
| | or the BR shortcut |

• Request protein sequence with accession code **P62152** (formerly P07181).

<u>Note</u>: The entry name of this calcium-binding calmodulin is CALM_DROME and details of the protein can also be found at http://us.expasy.org/uniprot/P62152 or

http://www.uniprot.org/uniprot/P62152

- Save file within L02 directory created on the desktop earlier.
- Accept the proposed .pro extension
- Click Save

| Open | жо |
|--------------------------------------|-----|
| Enter Entrez Protein | ₩R |
| | |
| | |
| Entrez Server: | _ |
| http://www.ncbi.nlgov/entrez/eutils/ | ′ 🕞 |
| | |
| Enter Sequence ID or Locus Name | |
| P62152 | |
| protein | |
| ? Cancel | ОК |
| | |
| | |
| | |
| 6 -111 | _ |
| Save | |

Analysis

Sites &

ЖN

| Save As: | P62152 | .pro | | |
|----------|--------|------|--------|----------|
| Where: (| 🚞 L04 | | | ÷ |
| | | | Cancel | Save |

<u>Note</u>: in the future when you have your own sequences, if these are not available on Entrez, simply use EditSeq to create the .pro file containing the sequence you want to analyze.

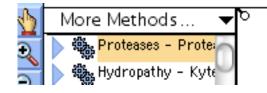
6.2 Protean Default View and Methods

The default view is that of standard secondary structure prediction with different methods that will be reviewed later.

| P62152 .pro |
|--|
| m.w. 16810.62, ch 23.87, pl 3.93 [149 AA] Page 1 of 2 |
| Pige 1 of 2 Disciple Dis |
| 10 20 30 40 50 60 70 80 90 100 110 120 130 140 |
| |

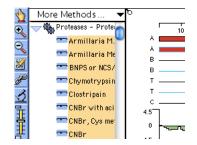


- At the top left click on "More Metods" and select Proteases – Protease Map
- The new analysis method will appear at the top left of the list of method currently displayed:



• Click on the small triangle to see the list of available enzymes and chemical treatments.

| More Methods | |
|-----------------------------|--|
| Title | 0 21 |
| Ruler | |
| Sequence | |
| Proteases - Protease Map | |
| Patterns | • |
| Charge Density - Charge | |
| Secondary Structure | Image: A state of the state |
| Hydropathy | • |
| Antigenicity | اله مع |
| Amphilicity – Eisenberg | 4 <u>8</u> 7 |
| Surface Probability - Emin | i |
| Flexibility - Karplus-Schul | z |
| | |



INFO

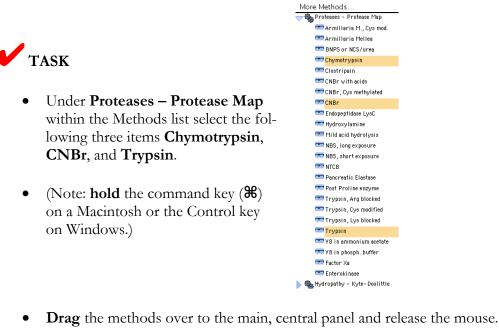
The same information is provided within a list by choosing Sites & Features > Show Protease List as shown in the next image:

231

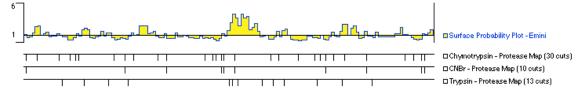
| 0 0 | | | Proteases |
|-------------------------|-------------------------|-------------|--|
| Name | Sites | Exceptions | Remarks |
| Armillaria M., Cys mod. | Х1С,Х1К | | Cys is reversibly aminoethylated. |
| Armillaria Mellea | X^K | | See the variant of this method with reversible modification of Cys residues. |
| BNPS or NCS/urea | W^X | | Trp is modified. Problem: with NCS Met can be modified into a sulfone. |
| Chymotrypsin | WnX,FnX,YnX,MnX,LnX,HnX | V^P | Not very specific, can also sometimes cut after Asn, Gln, Lys and Thr. |
| Clostripain | B^X . | | Same specificity can be obtained with Trypsin after blockage of Lys residues. |
| CNBr with acids | M°X,W°X | | Met is modified to Homoserine Lactone and Trp is also modified and blocked on its C-Terminal. |
| CNBr, Cys methylated | Mrx,xrc | | Cys is modified to Serine. Met is modified to Homoserine Lactone. |
| NBr | M°X | | Met is modified to Homoserine Lactone.Problem: Cys could be oxydized (but it is reversible). |
| Endopeptidase LysC | K^X | | Same specificity can be obtained with Trypsin after blockage of Arg residues. |
| lydroxylamine | N°G,N°L,N°A,N°M | | Cuts rarely: Asn-Leu, Asn-Ala and Asn-Met. |
| Ald acid hydrolysis | D^P | | Not a very efficient method. |
| NBS, long exposure | WrX,YrX,HrX | | Trp, Tvr and His are oxydized. A shorter exposure will not cleave after His (see corresponding method). |
| NBS, short exposure | W^X,Y^X | | These conditions are unfavorable for histidine cleavage. Trp, Tyr and His could be ineversibly oxydized. |
| NTCB | xnc i | | Cys is 5-cyanylated and can be reduced to Ala with Ni Raney. Problem: Cys can cyclise irreversibly. |
| Pancreatic Elastase | A^X,G^X,S^X,Y^X | | Not very specific. |
| Post Proline enzyme | P^X | P^P | Cleave slowly Ala-Xaa. Work only with proteins or fragments less then 50 residues long. |
| Frypsin, Arg blocked | K^X | K^P | One of the variants of the classical Trypsin method. Same specificity can be obtained with Endopeptidase LysC. |
| Trypsin, Cys modified | R^X,K^X,C^X | R^P,K^P,C^P | One of the variants of the classical Trypsin method. |
| rypsin, Lys blocked | B*X | B^P | One of the variants of the classical Trypsin method. Same specificity can be obtained with Clostripain. |
| rypsin | B^X,K^X | R^P,K^P | Classical method: see the variants of this method using reversible modifications of either Arg, Cys and Lys1. |
| /8 in ammonium acetate | E^X | | See also the method with Y8 in phosphate buffer which allows to cut Asp-Xaa. |
| /8 in phosph. buffer | D^X,E^X | | Slow to cut when Xaa is Phe, Val or Leu. See also the method with V8 in ammonium buffer which will not cut after Asp |
| actor Xa | IEGR^X | | Restriction Protease |
| Enterokinase | DDDDK^X | | Restriction Protease |

(Note: after viewing close this window)

6.3 Enzyme and chemical digestion and SDS PAGE Gel simulations



- The new methods will be shown a the bottom of the pane below the last displayed method (here the surface probability).
- Click in a white are to deselect them for better viewing.
- Recognition sites are shown as vertical bars



• Holding the Shift key, Click on each of these 3 lines to select them again

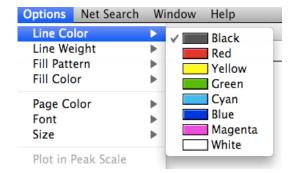
| | Chymotrypsin - Protease Map (30 cuts CNBr - Protease Map (10 cuts) Trypsin - Protease Map (13 cuts) |
|---|---|
| Select menu Sites & Features > SDS PAGE Gel Simulation. An image of the simulated separation will appear within a new window. | Sites & Features Options Net S New Feature %K Edit Feature %K Join to Feature Generate Features Sort Features Image: Sort Features New Protease Show Protease List SDS PAGE Gel Simulation Protease Fragment Summary Show Prosite Database Show Pattern Description |

• The background color default can be changed from black to white with the Options menu items.

| \varTheta 🔿 SDS-PAGE Simulation of P62152 | Options Net Search Windo | w Help O SDS-PAGE Simulation of P62152 |
|---|-----------------------------|--|
| BRL High+Low MW Marker | Line Color 🔹 🕨 | BRL High+Low MW Marker |
| BioRad NEB CHYM CNBR TRYT BRL | Line Weight | |
| 250000 | Fill Color | 250000 |
| 100000 100000 | Page Color 🕨 🕨 | Black 100000 |
| 5000050000 | Font 🕨 🗸 V | Vhite 50000 |
| 2500025000 | Size 🕨 | 25000 |
| 10000 | | 10000 |
| 5000 | | 5000 |
| 2500 | | 2500 |
| 1000 - | | |
| 500500 | Options > Page Co | 10r > 500 |
| 250250 | White | 250250 |
| 100 - 100 | | 100100 |
| 50 - | | 50 - |
| | | |



- Changing the page color from black to white automatically changes the line colors to the inverse color as well, here black.
- There are other colors to choose from as well.



<u>Note</u>: these colors apply ONLY to the lines selected within the Main View and NOT the SDS gel.



To see a summary of the simulated peptides select the following menu cascade: Sites & Features > Protease Fragment Summary. This opens a new window with the summary of the proteolytic fragments.

V INFO

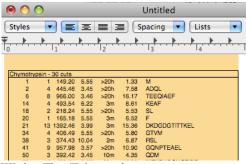
A right-click of the mouse would bring an option to **Copy** the contents and paste it elsewhere. Depending on the nature of the software this is pasted into, the result could be either an image or a plain text list.

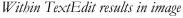
TASK

- **Right-click** (or Ctrl-Click) and select Copy
- **Open TextEdit** Application
- Paste
- Try to move: it's a picture
- TextEdit: File > New

(or use **\mathcal{B}N** shortcut)

- TextEdit: Format > Make Plain Tex .
- Paste •
- Now it is selectable plain text •





| | | | | - The course of | minuo | n neip | |
|-----|-------------|-----------|--------|---|------------|--------------------|---|
| 'ex | t | | | Font | | • | |
| | | | | Text | | • | |
| | | | | Make | Plain Tex | ct | |
| | | | | Prever | nt Editing | 3 | |
| | | | | | - | | |
| 0 | 00 | | | Unt | titled 2 | | |
| Pos | ition | Length | Weight | рI | 1/2 lif | e HPLC rt Fragment | : |
| Chy | motrypsin - | - 30 cuts | | | | | |
| 1 | 1 | 149.20 | 5.55 | >20h | 1.33 | M | |
| 2 | 4 | 445.46 | 3.45 | >20h | 7.58 | ADQL | |
| 6 | 8 | 966.00 | 3.46 | >20h | 16.17 | TEEQIAEF | |
| 14 | 4 | 493.54 | 6.22 | Зm | 8.61 | KEAF | |
| 18 | 2 | 218.24 | 5.55 | >20h | 5.53 | SL | |
| 20 | 1 | 165.18 | 5.55 | Зm | 6.52 | F | |
| 21 | 13 | 1392.46 | 3.99 | Зm | 15.36 | DKDGDGTITTKEL | |
| 34 | 4 | 406.49 | 5.55 | >20h | 5.80 | GTVM | |
| 38 | 3 | 374.43 | 10.04 | 2m | 6.87 | RSL | |
| 41 | 9 | 957.98 | 3.57 | >20h | 10.90 | GQNPTEAEL | |
| 50 | 3 | 392.42 | | 10m | 4.35 | QDM | |
| | | | | | | | |

17.80

30m

TextEdit File Edit Format

Open...

INEVDADGNGTIDF

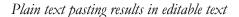
New

Format Window Help

Windo

ЖN

жΟ



1479.51 3.07

<u>Note</u>: Sequence is one of the Methods and can be added for clarity within the main panel.

53

14

| 1 2 6 14 18 20 21 34 | 8 9 4 4 2 2 1 1 13 13 | 66.00 93.54 18.24 65.18 92.46 | 3.46 6.22 5.55 5.55 3.99 | >20h 3m >20h 3m 3m | 1.33 7.58 16.17 8.61 5.53 6.52 15.36 5.80 | KEAF SL F DKDGDGTITTKEL |
|---|-----------------------------------|---|--------------------------------------|--------------------------------|--|---|
| 38 | 3 3 | 74.43 | 10.04 | 2m | 5.80 6.87 | RSL |
| 41 | 99 | 57.98 | 3.57 | >20h | 10.90 | GONPTEAEL |
| 50 | 3 3 | 92.42 | 3.45 | 10m | 4.35 | GONPTEAEL QDM INEYDADGNGTIDF |
| 53 | 14 14 | 79.51 | 3.07 | 30m | 17.80 | INEYDADGNGTIDF |
| 67 | 3 2 | - | | | | |
| 70 | 1 1 | | ut | | | |
| 71 | 2 2 | | Сору | | | |
| 73 | 1 1 4 5 13 15 | | .opy | | | |
| 74 | 13 15 | P | aste | | | |
| 91 | 3 4 | 0 | loar | | | |
| 91 | 7 7 | C | lear elect | | | and the second se |
| 101 | 66 | S | elect | IIA · | | |
| 107 | 2 3 | 5 | | | | |
| 109 | 2 2 2 2 3 3 | | | | _ | |
| 111 | 3 3 | F | ragn | ient | Sumr | nary Options |
| 114 | | 40.00 | 0.44 | 22011 | 1.55 | GENL |
| 118 | 8.9 | 66.96 | 3.12 | >20h | 9.44 | TDEEVDEM |
| 126 | 14 15 | 64.62 | 3.58 | 30m | 18.79 | IREADIDGDGQYNY |
| 140 | 3 4 | 23.41 | 3.57 | 30m | 8.99 | EEF |
| 143 | 3 3 | 49.44 | 5.55 | >20h | 4.58 | IREADIDGDGQYNY EEF YTM M |
| 146 | 1 1 | 49.20 | 5.55 | >20h | 1.33 | м |
| 147 | 3 3 | 34.35 | 9.00 | >20h | 0.00 | TSK |
| | | | | | | |
| CNBr - 10 c | | | | | | |
| 1 | 1 1 | 49.20 | 5.55 | >20h | 1.33 | м |
| 2 | 36 39 | 79.37 | 3.84 | >20h | 71.35 | ADQLTEEQITTKELGTYM |
| 38 | 15 16 | 88.82 | 3.96 | 2m | 54.38 | RSLGQNPTEAELQDM INEYDADGNGTIDFPEFLTM |
| 53 | 20 21 | 98.38 | 3.04 | 30m | 63.77 | INEYDADGNGTIDFPEFLTM |
| 73 | 1 1 | 49.20 | 5.55 | >20h | 1.33 4.23 | M |
| 74 | 4 5 | 04.64 | 11.05 | >20h | 4.23 | ARKM |
| 78 | 33 38 | 13.10 | 4.48 | 300 | 07.20 E4.26 | KDTDSEEEISAAELRHYM |
| 111 | | | | | | TNLGEKLTDEEYDEM |
| 126 146 | 20 23 | 40.20 | 3.48 E EE | 30M | 1.00 | IREADIDGDGQYNVEEFYTM |
| 140 | 2 2 | 99.20 | 0.00 | >20h | 1.33 0.00 | M |
| H | 5 5 | 04.00 | 9.00 | 72011 | 0.00 | TON |
| Trypsin - 13 | euts. | | | | | |
| 1 | 14 16 | 52.83 | 3.83 | >20h | 21.81 | MADQLTEEQIAEFK |
| | | | | 0.0 | 40.40 | the second state with the second state |
| 1 15 | 8 9 | 56.05 | | | | EAFSLFDK |
| 15 | 89 | 56.05 06.93 | 4.18 | 30m 3m | 9.72 | EAFSLFDK DGDGTITTK |
| 15 23 32 | 99 | 06.93 | 3.92 6.24 | 3m 30m | 9.72 13.23 | DGDGTITTK ELGTYMR |
| 23 | 99 | 06.93 | 3.92 6.24 | 3m 30m | 9.72 13.23 | DGDGTITTK ELGTYMR |
| 23 32 | 99 78 3740 | 06.93 04.95 71.49 46.17 | 3.92 6.24 3.35 9.00 | 3m 30m ≽20h 3m | 9.72 13.23 70.07 0.00 | DGDGTITTK ELGTYMR |
| 23 32 39 76 | 9 9 7 8 37 40 1 1 | 06.93 04.95 71.49 46.17 | 3.92 6.24 3.35 9.00 | 3m 30m >20h 3m | 9.72 13.23 70.07 0.00 | DGDGTITTK ELGTYMR SLGQNPTEAPEFLTMMAR K |
| 23 32 39 76 | 9 9 7 8 37 40 1 1 | 06.93 04.95 71.49 46.17 | 3.92 6.24 3.35 9.00 | 3m 30m >20h 3m | 9.72 13.23 70.07 0.00 | DGDGTITTK ELGTYMR SLGQNPTEAPEFLTMMAR K |
| 23 32 39 76 | 9 9 7 8 37 40 1 1 | 06.93 04.95 71.49 46.17 | 3.92 6.24 3.35 9.00 | 3m 30m >20h 3m | 9.72 13.23 70.07 0.00 | DGDGTITTK ELGTYMR SLGQNPTEAPEFLTMMAR K |
| 23 32 39 76 | 9 9 7 8 37 40 1 1 | 06.93 04.95 71.49 46.17 | 3.92 6.24 3.35 9.00 | 3m 30m >20h 3m | 9.72 13.23 70.07 0.00 | DGDGTITTK ELGTYMR SLGQNPTEAPEFLTMMAR K |
| 23 32 39 76 | 9 9 7 8 37 40 1 1 | 06.93 04.95 71.49 46.17 | 3.92 6.24 3.35 9.00 | 3m 30m >20h 3m | 9.72 13.23 70.07 0.00 | DGDGTITTK ELGTYMR SLGQNPTEAPEFLTMMAR K |
| 23 32 39 76 | 9 9 7 8 37 40 1 1 | 06.93 04.95 71.49 46.17 | 3.92 6.24 3.35 9.00 | 3m 30m >20h 3m | 9.72 13.23 70.07 0.00 | DGDGTITTK ELGTYMR SLGQNPTEAPEFLTMMAR K |
| 23 32 39 76 | 9 9 7 8 37 40 1 1 | 06.93 04.95 71.49 46.17 | 3.92 6.24 3.35 9.00 | 3m 30m >20h 3m | 9.72 13.23 70.07 0.00 | DGDGTITTK ELGTYMR SLGQNPTEAPEFLTMMAR K |

P62152 .pro

Fragment

pl 1/2 life HPLC rt

Chymotrypsin - 30 cuts

6.4 Composition

The protein composition can be obtained from the Analysis menu.

TASK

- Select the menu cascade: Analysis > Composition
- Analysis Sites & Features Hide Available Methods Show Legend Tabular Data ... Model Structure Titration Curve... Composition.. \varTheta 🔿 🔿 Composition of P62152 .pro

Predicted Structural Class of the Whole Protein: Alpha Deléage & Roux Modification of Nishikawa & Ooi 1987

Whole Protein

16810.62 m.w.

149

Analysis Molecular Weight

- The new window will contain the • protein composition.
- In the same way as shown in the pre-• vious section, the contents can be copied.
- The resulting pasted material can be either plain text or an image depending on the software receiving the pasting.
- See above example with TextEdit

Length 1 microgram = 59.486 pMoles Molar Extinction coefficient 1490±5% 1 A(280) = 11.28 mg/ml Isoelectric Point 3.93 Charge at pH 7 -23.87 Whole Protein Composition Analys 96 hs %b∨ anbei count 54 38 14 30 43 10 Amino Acid(s) Charged (RKHYCDE) Acidic (DE) Basic (KR) Polar (NCQSTY) requency 36.24 25.50 9.40 20.13 weight 41.23 27.77 11.67 19.26 27.68 4.23 0.00 11.64 16.12 Hydrophobic (AILFWY) A Ala 28.86 6.71 0.00 11.41 7.38 0.67 5.37 5.37 5.37 6.04 4.03 3.36 8.72 4.03 3.36 8.72 4.70 0.00 0.00 0.00 C Cys D Asp E Glu F Phe G Gly H His I Ile K Lysu M Met N Asn P Q Glu S Ser T The W Thy Y Yal W Thy Y Yal W Thy Y Yas X Xxx 0 Cut Copy Paste Clear Select All Horizontal Units. 13 7.82 4.13 0.00 0.97 0.00 0.00 0.00 01000 . Ter 0.00 0.00

6.5 Titration Simulation

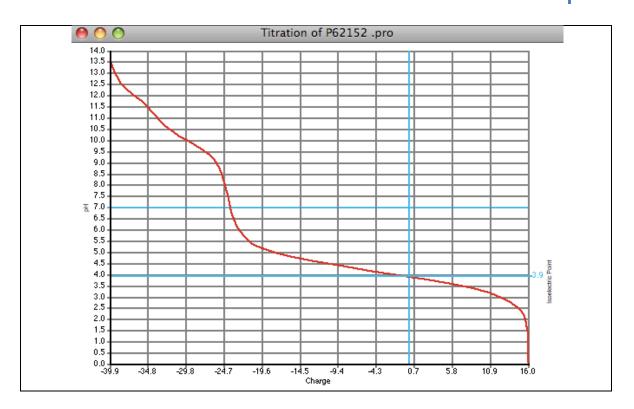


Select the menu cascade: Analysis > **Titration Curve**

Analysis Sites & Features

Hide Available Methods Show Legend

Tabular Data... Model Structure Titration Curve. Composition...



6.6 Model Structure Simulations: Helical Wheel

The prediction of secondary structure will be highlighted in the next section. However, the default view of Protean shows the predicted alpha- and beta- secondary structure. Protean also offers a graphical view for these local folds of the proteins: helical wheel, helical net, and beta net.



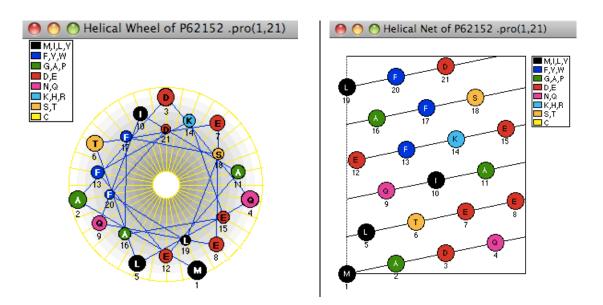
• On the main panel graphically select with the mouse the length of the first alpha helix structure on the first line: it appears as a red feature.

| | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 |
|----|---|----|----|----|----|----|----|----|
| A. | | | | | | | | |
| A | | | | | | | | |
| в | | | | | | | | |
| в | | | | | _ | | | |
| - | | | | | | | | |
| Т | | | | | | | | |

Note: On the left panel, with blue squares, you may need to switch from the "object selector" in the shape of a hand ($\)$) to the "range selector" in the shape of an arrow ($\)$).

- From the menu select: Analysis > Model Structure > Helical Wheel
- Then select: Analysis > Model Structure > Helical Net

| Analysis | Sites & Features | Options | Net Search | Windo |
|--------------------|-------------------------|----------|---------------------------|-------|
| Hide Av Show Le | ailable Methods gend | ¥[援] | | |
| Tabular | | | | |
| | tructure | | Helical Whe | el |
| Titratio | n Curve | _ | Helical Net | |
| Compos | sition | _ | Beta Net | |
| | lethod Outline | - | Linear Spac Chemical F | |



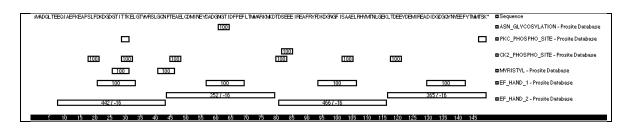
Both methods are useful for locating hydrophobic or hydrophilic patches and particularly useful for hydropathic helices.

6.7 Patterns Search: Prosite



- Within "More Methods" choose Patterns > Prosite Database
- Drag the listed item Patterns Prosite Database onto the main panel. A progress window will appear.
- The results will be displayed both within the graphical panel and under the Methods list

| Title | 0 20 30 40 |
|---|----------------|
| Ruler | |
| Sequence | |
| Proteases - Protease Map | |
| Patterns > | Prosite Databa |
| Charge Density - Charge | Ariadne file |
| | |
| Prosite Progre | \$\$\$ |
| | |
| ocating Prosite Pattern: | |
| ocating Prosite Pattern: | Stop |
| | |
| ocating Prosite Pattern: Time remaining: Less than a minute | |
| | |
| Time remaining: Less than a minute | |
| Time remaining: Less than a minute | |
| Time remaining: Less than a minute | |
| Time remaining: Less than a minute | |
| Time remaining: Less than a minute | |
| Time remaining: Less than a minute | |
| Time remaining: Less than a minute Portier over : Pietriou Curtani More Methods Patterns - Prosite Database 1 ASN_GLYCOSYLATION 1 PKC_PHOSPHO_SITE | |
| More Methods Patterns - Prosite Database ASN_GLYCOSYLATION ASN_GLYCOSYLATION CONTACT AND | |



<u>Note</u>: The EF-hand pattern is typical of calmodulin and typically exhibit a *helix-turn-helix* motif with acidic residues chelating a calcium ion.

6.8 Help - Patterns Search: user-specified descriptor (Ariadne File)

Help can be found with the last menu item on the menu bar from all Lasergene modules.

Here is an example of Help search for the Pattern Search method Ariadne.

| _ | Sear | h ariadne 🛛 🔊 |
|------|-----------|--|
| 40 4 | Help Topi | ari Adding Ariadne Patterns to Protean Pattern Matching Methods - Aria Applying the Ariadne File Method Ariadne Structure Pattern-Matchi Show All Results |

| \varTheta 🔿 🔿 Search Results: ariadne | |
|---|------|
| ▲ ► @ | 8 |
| Title | Rank |
| Help Topics | |
| 💮 ari | |
| 🞯 Adding Ariadne Patterns to Protean | |
| 🎯 Pattern Matching Methods – Ariadne File | |
| Applying the Ariadne File Method | |
| 💮 Ariadne Structure Pattern-Matching Method for Protean | |
| Modifying a Displayed Pattern | |
| 🞯 Ariadne References | |
| 💮 Changing Match Stringency | |
| 🞯 Changing Ariadne File Method Parameters | |
| 🞯 Recognized Simple Pattern Elements | |
| 🞯 Protean Analysis Methods | |
| 💩 Hierarchical Patterns | |
| 🞯 Dictionary of APSL Terms | |
| 8 | |
| 🞯 Creating a Pattern Descriptor | |
| Exporting a Protean document to GenVision | |
| Found: 17 help topics | Show |

display the list if entries in the Help search.

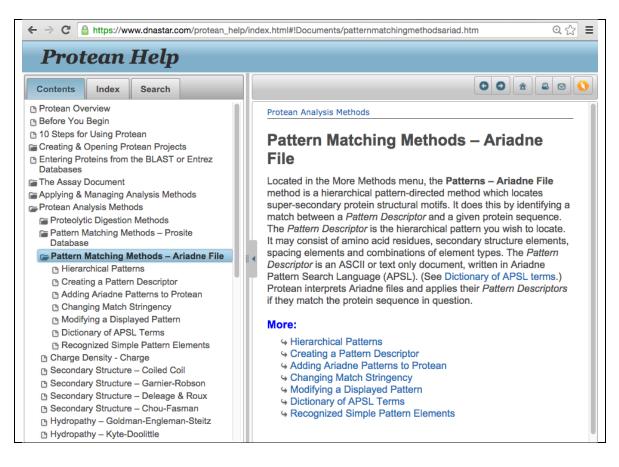
Select "Show All Results" to

The results are shown in a new window and are sorted by relevance.

The Ariadne method is different from the Prosite (http://prosite.expasy.org/) pattern search and is based on a vocabulary of keywords and variables rather then ambiguous sequence codes.

To see an example for creating a search pattern file, see the help entry titled: "**Creating a Pattern Descriptor**" towards the end of the list and "**Recognized Simple Pattern Elements**."

To understand the meaning of descriptors with an example, see "Hierachical Patterns." See https://www.dnastar.com/protean_help/index.html#!Documents/patternmatchingmethodsariad.htm or the short URL version: http://bit.ly/1FVLLV7



6.9 Quit Protean

We will reopen Protean in the next exercise.

7. Protean: Secondary Structure Predictions

INFO <u>Basics of sequence-based protein structure prediction</u>:

Very often the only information at hand is the primary sequence of a protein, often deducted from DNA sequence. Based on the protein sequence and amino acids biochemical properties, some software predict the conformation of protein secondary structure elements: *alpha helices, beta sheets, turns and random coils*.

Most of the secondary structure predictions are based on tabulations of observed amino acids conformations in solved 3D protein structures, and anticipate that similar residues will occur in analogous configurations. In the algorithms of Chou and Fasman (1978), Levitt (1978) and Garnier, Osguthorpe and Robson (1978), amino acids are classified as <u>formers</u>, <u>neutral</u> or <u>breakers</u> of the secondary structure elements.

Chou and Fasman used 19 protein structures, with 2,473 amino acids and Levitt used 11,569 residues in globular proteins to deduct predictive rules, generalized here:

- 1- find any cluster of 4 consecutive helix-forming residues within any length of 6 amino acids
- 2- propagate helix in both directions from this nucleus until:
- 3- at least 4 helix-breakers are found (tetra-peptide)
- 4- beta sheet rules are similar except that 3 out of 5 beta-formers are required to nucleate a sheet
- 5- in the case of a tie between alpha and beta, the helix usually wins
- 6- turns require 4 out of 4 residues that prefer turn configurations.

The algorithm described by Garnier *et al.* is similar to that of Chou and Fasman and is also based on observed amino acids conformations. However, this method considers a window of 17 residues (8 residues on each side of the current amino acid on the sequence) and asks whether the flanking neighbors favor helix, sheet or turn configurations. Therefore the method is not only residue specific but also sequence specific. The program will also give different predictions if the user already knows that a protein contains a particular alpha/beta ratio.

The algorithms and tables have been refined considerably but the principles remain the same today and are good probably to 55% on average. But secondary structure predictions are not very reliable, especially for proteins that are not soluble or globular. The Chou and Fasman algorithm may be better at predicting structures with large helical content.

Chou P.Y. and Fasman G.D. 1978 Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol Relat. Areas Mol. Biol. 47:45-148

Levitt, M. 1978 Conformational preferences of amino acids in globular proteins. *Biochemistry* **17**: 4277-4285 Garnier, D.G. Osguthorpe, D.J. And Robson, B. 1978 Analysis of the accuracy and implications of simple methods for predicting the secondary structure of proteins. *J. Mol. Biol.* **88**:873-894

7.1 Open the Protean software



Double-click Protean from the following location: /Applications/DNASTAR/Lasergene 12/

7.2 Open a protein sequence and display predictions



Open a protein sequence from the Entrez database with the following Protean menu cascade: **File > Enter Entrez Protein...**

On the next form **type P00568** as the accession number of a human adenylate kinase.

Click OK

Within the next window **change** the **name** to **adk.pro** and save the file in L02 on the Desktop.

Click Save

As soon as you click save, the file is saved in the L02 folder in a DNASTAR format, and the structure analysis is immediately displayed in a complex, default prediction panel:

| Position: 1 Nore Methods Hydropathy - Kyte-Doolittle And Angele Probability - Ennin Fexibility - Karplus-Schulz Fexibility - | | 21634.84, ch. + 3.2 sale pha, Regions - Gamiar- pha, Regions - Charler Ka, Regions - Charler Ka, Regions - Charler M, Amphipathic Region Setting - Karplus Setting - Karplus Setting - Karplus Setting - Karplus Setting - Karplus |
|---|--|---|
|---|--|---|

The far left column contains icons to zoom or reduce the size of the display (magnifying glass with a + or - sign within).

File Edit Analysis Sites & F

| New | | ЖN | |
|----------------------|------------------|-----|---------|
| Open | | жo | |
| Enter Entrez | Protein | ЖR | |
| | | | |
| Entrez Server: | | | |
| http://www.ncbi.nlg | ov/entrez/eutils | / 🛟 | |
| | | | |
| Enter Sequence ID or | Locus Name | | |
| P00568 | | | |
| 100200 |] | | |
| protein | ; | | |
| ? | Cancel | ОК | |
| - | Save | | - |
| | | | |
| Save As: adk pro | | | |
| WI (| | | |
| Where: Usktor |) | ÷ | |
| | | | |
| | | Ca | ancel S |

The next column is a list of methods used to calculate the current display. The column can

be made wider with the "curtain pull" icon: **Pull this to the right to make the left panel wider.**

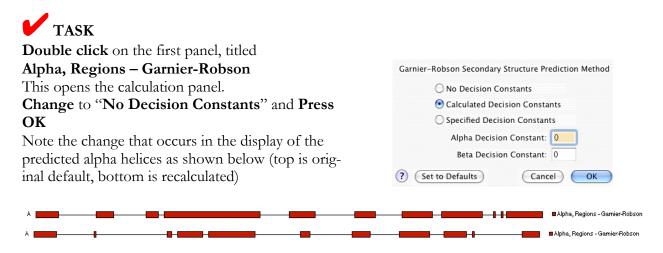
The middle panel is that of the graphical display. Each method used for calculating the display is listed on the right hand panel.

7.3 Recalculating graphics (changing defaults)

INFO

Each graphical plot is an object. Double clicking on one such object calls a dialog panel that displays the defaults used to calculate this plot; at this point parameters can changed and the plot recalculate.

For example the first plot depicts the prediction of alpha elements by the method of Garnier-Robson. The prediction is based on a "decision constant" based on the percentage of alpha and beta structures that are either known by other methods, or that could be inferred. The algorithm calculates a value for this decision constant, which can affect the results.



<u>Note</u>: Constants are based on three protein classes: with less than 20%, 20-50% and >50% alpha or beta sheet. Circular dichroïsm data is an experimental method that can help estimate the alpha-helical content of a structure. If such data is know, use the "Specified Decision Constants" option and enter values for both alpha and beta.

7.4 Recalculating options from the methods panel (left)

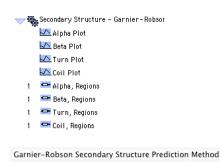


On the left panel, **click** on the **triangle** next to "Secondary Structure - Garnier-Robson." This will display subset items.

Double-click on **Alpha Plot**. This will open the same recalculation panel, this time with the calculated decision constants filled in.

Click on "Specified Decision Constants" Enter 55 and 30 for Alpha and Beta respectively Click OK

Note that changes occur for all the Garnier-Robson objects, as depicted below

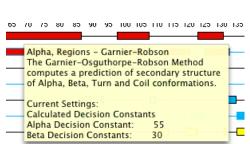


| O No Decision Constants | |
|-------------------------------|-----|
| Calculated Decision Consta | nts |
| O Specified Decision Constant | ts |
| Alpha Decision Constant: | 55 |
| Beta Decision Constant: | 30 |

Original default:



Place the mouse cursor above the Alpha, Regions – Garnier-Robson and **wait a few seconds**. A yellow window will reveal the nature of the plot and some of its parameters, such as the Decision Constants, here 55 and 30 respectively for alpha and beta.



7.5 Reorganizing the panel plots

As each plot is an independent object, it can be dragged and moved within the graphics panel.

V TASK

As an exercise **move all** the Garnier-Robson panels and place them in the same order as in the image above.

7.6 More methods

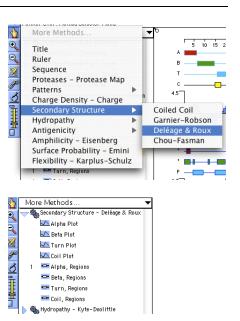
For clarity and simplicity, only a limited number of plots (based on specific methods such as Garnier-Robson) are shown at first.



Other methods or objects can be added from the More Methods... pull down menu. For example, **select** the **Secondary Structure > Deléage & Roux** method.

This will create a new entry on the left panel called "Secondary Structure – Deléage & Roux"

Click on the triangle at left to expand the visible subcontents.



Click on the **Alpha, Regions** button and drag it onto the graphics panel. A new alpha helical prediction plot is displayed. You can rearrange the panels order and place all three alpha helix structure prediction objects in a row as shown below:

| | Alpha, Regions - Gi Alpha, Regions - Ct Alpha, Regions - De |
|--|---|
| | Deléage & Roux Parameters |
| Note: as before the calculation methods can be altered by | Structural Class: 💿 Computer Calculated |
| | ○ A |
| double clicking on Alpha, Regions icon. | Ов |
| | ○ A+B |
| Here the calculation method is based on a structure class of | 🔿 A/B – A |
| protein, such as alternate alpha/beta topology. | ○ A/B - B |
| protein, such as alternate alpha/beta topology. | ○ R |
| | Set to Defaults Cancel OK |

7.7 Changing colors and patterns of plots

The other 2 helical plots have a red filled look. These cosmetic changes can be changed from the Options menus and submenus:

| Options Net Search | Window Help | Options Net Search | n Window Help | Options Net Search | Window Help | Options Net Search | Window | Help |
|--------------------|--------------|--------------------|---------------|--------------------|-------------|--------------------|--------|----------------------------|
| Line Color | Black | Line Color | • 0 | Line Color | • 0 | Line Color | . 0 | |
| Line Weight | Red | Line Weight | 0.25 Point | Line Weight | • | Line Weight | • | |
| Fill Pattern | Yellow | Fill Pattern | 0.5 Point | Fill Pattern | | Fill Pattern | • | nlm. |
| Fill Color | Green | Fill Color | ▶ ✓ 1 Point | Fill Color | | Fill Color | | Black |
| Page Color | Cyan Blue | Page Color | 2 Point | Page Caler | | Page Color | • | Red Yellow |
| Font | Magenta | | | | | Font | | and a decided extended and |
| Size | White | | | | | | | |

The default pattern (*) is empty and should be switched to completely filled (black square) or patterned designs.

To make this new panel look like the other two, follow the following menu cascade sequences:

TASKClick on the new Deléage & Roux Alpha, Regions panel objectOptions > Fill Pattern > black-squareOptions > Fill Color > Red

<u>Note</u>: Graphical interfaces are usually very redundant. All the options menus are also available by Ctrl-Clicking on the object, to open a contextual menu.

| Help | | |
|-------------------|---|---------|
| Method Parameters | | - |
| Model Structure | • | |
| Line Color | • | ✓ Black |
| Line Weight | • | Red |
| Fill Pattern | | Yellow |
| | | Green |
| Fill Color | • | Cyan |
| Page Color | | Blue |
| | | Magenta |
| | | White |

Other menus are also present in addition to the Options

7.8 Tabular Data

The graphical panel presents an overview of the computed data. However the pertinent information calculated for each residue can be displayed in a table. Only the objects present on the graphics panel are shown on the tabular panel.



To display the data in table form, select the following menu cascade: Analysis > Tabular Data...

| Analysis | Sites & Features | Options |
|----------|------------------|---------|
| Hide Av | ailable Methods | ¥[|
| Show Le | gend | 쁐] |
| Tabular | Data | |
| Model S | tructure | • |
| Titratio | n Curve | |
| Compos | sition | |

| 0 | | Tabular Data for adk.pro | | | | | | | | | | | | |
|----------------------|--------|--------------------------|---------------|----------------|--------------|--------------|---------------|-------------|---------------|-----------------------|----------------|---------------|----------------|-----------------------|
| olor | 5 | | | | | | 0.5500.57 | | | | | | | |
| ∧in ∧iax ∧iean | | Å 0.619 | Å 0.443 | Å 0.768 | B 0.0928 | B 0.294 | Т 0.113 | T 0.247 | C 0.18 | -1.93 2.68 0.47 | ; 0.459 | ; 0.479 | F 0.608 | -0.60 3.10 0.83 |
| les | Pos | Gami Alpha | Chou Alpha | Deléa Alpha | Gami Beta | Chou Beta | Garni Turn | Chou Tum | Garni Coil | Kyte Hydro | Eisen Alpha | Eisen Beta | Karpl Fle≍i | James Antig |
| let | 1 | A | A | A | | | | | | 0.58 | | + | | 0.75 |
| lu – | 2 | A | A | A | | | | | | 1.01 | + | • | | 0.7 |
| lu | 3 | A | A | A | | | | | | 1.44 | + | • | | 0.7 |
| /S | 4 | A | A | A | | | | | | 1.52 | | | | 0.7 |
| eu - | 5 | A | A | | | | | | | 1.96 | | • | F | 0.9 |
| /S | 6 | A | A | | | | | | | 1.67 | | | F | 0.9 |
| /S | 7 | A | | | | В | | | | 0.78 | | | F | 0.9 |
| r | 8 | A | | | | в | | | | 0.08 | | | F | 0.6 |
| /s | 9 | A | | | | В | | | | -0.82 | | | F | 0.4 |
| 2 | 10 | . | | | В | В | | | | -0.87 | | | | -0.3 |
| 2 | 11 | . | | | в | В | | | | -1.26 | | | | -0.6 |
| he | 12 | l . | | | в | В | | | • | -1.64 | | • | | -0.6 |
| al | 13 | · · | | | В | В | | | • | -1.54 | | | | -0.6 |
| al | 14 | l . | | | в | В | | | | -1.93 | | | | -0.6 |
| y | 15 | . | | | | • | | | C | -1.34 | | | F | -0.0 |
| У | 16 | · | | | • | • | | Т | С | -0.80 | | • | F | 0.4 |
| o | 17 | · · | • | | • | • | | Т | С | -0.06 | | • | F | 0.4 |
| У | 18 | · · | | : | • | • | Т | Т | • | 0.46 | | • | F | 1.2 |
| er | 19 | · | • | A | • | • | Т | Т | • | 1.00 | | • | F | 1.5 |
| ly | 20 | · · | • | A | • | • | Т | Т | • | 1.34 | | • | F | 1.8 |
| /S | 21 | · · | • | A | • | • | Т | Т | : | 1.02 | | • | F | 2.3 |
| ly | 22 | · · | • | Ą | • | • | • | T | c | 1.23 | | : | F | 2.2 |
| nr | 23 | : | : | Ą | • | • | • | Т | с | 1.62 | | • | F | 3.0 |
| In | 24 | A I | Ą | Ą | • | • | | • | • | 1.03 | | • | F | 2.1 |
| ys I | 25 | A I | Ą | Ą | · · | • | • | • | • | 0.52 | : | • | F | 1.3 |
| 21 | - ac - | L | L L | | I I | I | | I | I | 0.40 | • | I | | 1 1 |

245

INFO

<u>Note</u>: One or more column can be selected and then pasted within a spreadsheet such as Microsoft Excel for further analysis. When the columns are copied and pasted within Excel, the first 2 columns (A and B) of the new spreadsheet are automatically populated with the residue name (A) and sequence position number (B) for convenience. On this example the antogenic and surface index (last 2 columns of the Tabular data) are shown when pasted into Excel.

| 0 | A | B | C | D |
|----|---------|----------|---------------|--------------------|
| 1 | Minimum | | -0.6 | 0.05 |
| 2 | Maximum | | 3.1 | 4.54 |
| 3 | Mean | | 0.83 | 1.18 |
| 4 | Residue | Position | Antigenic Ind | Surface Probabilit |
| 5 | Met | 1 | 0.75 | 2.22 |
| 6 | Glu | 2 | 0.75 | 1.43 |
| 7 | Glu | 3 | 0.75 | 2.24 |
| 8 | Lys | 4 | 0.75 | 4.54 |
| 9 | Leu | 5 | 0.9 | 3.78 |
| 10 | Lys | 6 | 0.9 | 4.36 |
| 11 | Lys | 7 | 0.9 | 1.53 |
| 12 | Thr | 8 | 0.6 | 1.3 |
| 13 | Lys | 9 | 0.45 | 0.56 |
| 14 | Ile | 10 | -0.3 | 0.21 |
| 15 | Ile | 11 | -0.6 | 0.11 |
| 16 | Phe | 12 | -0.6 | 0.05 |

7.9 Reality check



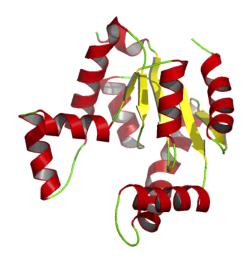
It happens that this protein has been crystallized and is therefore a known 3D structure with PDB ID code 3ADK.

If you have time, you can use the program PyMOL to "fetch" entry 3ADK and represent it as a cartoon structure:

Names Panel: 3adk > S > Cartoon

colored by secondary structure:

Names **Panel: 3adk > C > ss > choose** option with red helix /3adk 2 6 11 16 21 26 31 36 41 46 51 56 63 M SO4 <u>E SO4 EKLKKSKIIFVVGGPGSGKGTQCEKIVQKYGYTHLSTGDLLRAEVSSGSARGKMLSEIM</u>



You can then display the sequence (Top menu **Display > Sequence On**) to more easily see the amino acids involved in alpha helices.

From the tabular data of Protean (top menu **Analysis > Tabular Data...**), you can also see which individual amino acids are predicted as alpha helices by various methods. It is therefore possible to see how well some helices are predicted and even count them.

8. Protean 3D

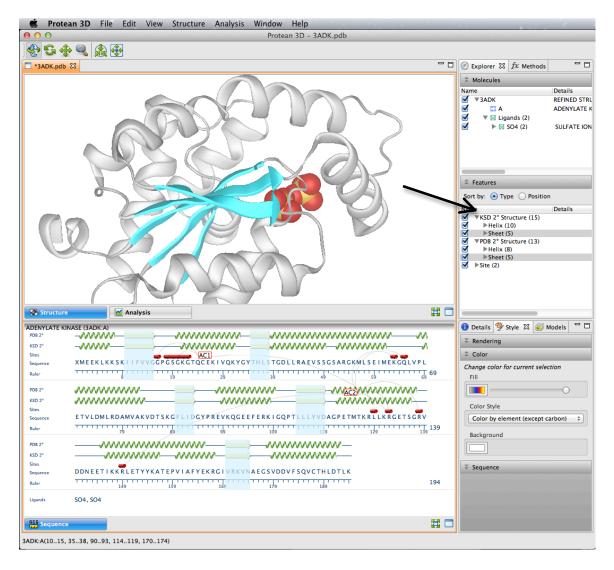
Protean3D is a new addition to Lasergene for protein analysis that includes 3D (*i.e.* PDB files) and variations of the various "methods" explored within Protean.

8.1 Try Protean3D

Optionally you can explore the 3ADK structure with Protean3D that you will find with the other Lasergene modules in /Applications/DNASTAR/Lasergene 12/

With this version it is possible to open the file directly from the Protein Data Bank web site (www.rcsb.org) with the menu cascade File > Open From PDB...

Protean 3D will open the file and automatically show the protein as ribbons and any solvent as spheres.



In the middle panel to the right the "Features" section will list the secondary structure found within the PDB file as well as in the KSD¹ secondary structure database (see arrow in image.)

On the image above it was chosen to select the Beta sheets from both offered methods. This has the effect to color the sheets in cyan-blue within the 3D viewer as well as the bottom panel that illustrates in cartoon form the secondary structure.

If you have time you can explore other aspects of Protean3D such as the "Rendering" and "Color" panels.

| 2 Color | â Rendering |
|--|--|
| Change color for current selection | Change rendering for current selection Backbone |
| | ✓ Sticks Ball and Stick Wireframe |
| Color by element (default) | Spheres |
| ✓ Color by element (except carbon) Color by chain Color by amino acid chemistry Fill with solid color | Ribbon Ribbon (cartoon) Ribbon (solid) Tube Tube (temperature) Tube (occupancy) |
| | Trace |
| | Trace (wireframe) Trace (spline) |
| | Points |
| | Invisible |

Here the background was changed from the default black to white for printing purposes.

8.2 Analysis panel

All or most of the methods available in Protean are also present in Protean3D and visible when clicking on the Analysis button (circled in image below) next to the Structure button in the middle pane

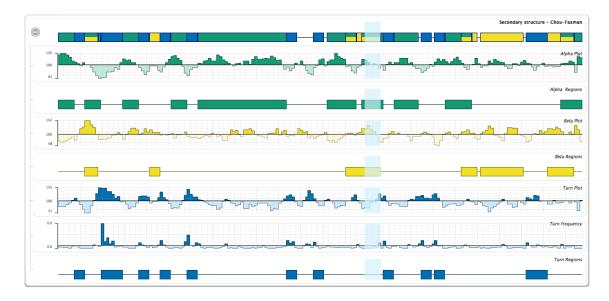
¹ **Kinase Sequence Database** is a collection of protein kinase sequences grouped into families by homology of their catalytic domains. http://sequoia.ucsf.edu/ksd/

Example: click on one beta-sheet arrow within the Sequence panel (the bottom panel) for example around amino acid residue 114-119.

Once clicked the same area will be highlighted on all methods (enhanced with black outline here.)

| 00 | otean 3D File Edit View Structure Analysis Window Help Protean 3D - 3ADK.pdb | | ዋ 🛞 🛋 🦉 🧭 🧰 00:06 |
|-------------------|--|----------------------|--|
| 🎨 Ġ 🛛 | | | |
| *3ADK.pd | pdb 83 | - D | 🖉 Explorer 🕄 🏂 Methods 👘 |
| 3ADK:A | A | | Molecules |
| JADICA | | Tarre - Chox-Tasman | Name Details |
| 0 | XUNNY XV | are - Onu-rashan | 🗹 🗆 A ADENYLA |
| | | | |
| | Inverse | icity - Jameson-Wolf | |
| 0 | | | |
| | | | |
| 0 | | athy - Kyte-Doolttie | |
| | | | |
| | Hydrog | authy - Hepp-Woods | |
| 0 | | | |
| | | | |
| 0 | 04 | rge density - Charge | 2 Features |
| | | | Sort by: 💿 Type 🕐 Position |
| | Pote | ity - Karphus-Schulz | Name Details Name Details Image: Structure (15) Structure (15) |
| 0 | | | ✓ ► PDB 2° Structure (13) ✓ ► Site (2) |
| | | | Site (2) |
| | | Author Annotations | |
| | | | |
| 0 | | ~~~~ | 😗 Details 🦻 Style 🕮 🥃 Models 📟 |
| | ¥60.2° Gravita | | * Rendering |
| | | ~~~~ | - Color |
| | ∧C1 ♦ ∧C2 | | - Sequence |
| Structu | ture (1 M Analysis | H 🗆 | |
| | E KINASE DADA | 144 L | |
| PD8 2* | | | |
| KSD 2* Sites | | | |
| Sequence | XMEEKLKKSKI I FVVGGPGSGKGTQCEKI VQKYGYTHLSTGDLLRAEVSSGSARGKMLSE I MEKCQLVPLETVLDMLRDAMVAKVDTSKGF | | |
| Ruler | | | |
| PD8 2* KSD 2* | | | |
| Sites | | | |
| Sequence Buler | KQGEEFERKI GQPTLLLYVI AGPETMTKRLLKRGETSGRVDDNEETI KKRLETYVKATEPVI AFYEKRGI VRKVNAEGSVDDVFSQVCTHL | | |
| | | | |
| Liganda | 504, 504 | | |
| | | | |
| Sequen | ence | 🛱 🗖 | |

Clicking on the ^(III) sign at the left will open each method to full display, for example with Chou and Fasman here, where the blue highlight is also shown.



To close click on the \bigcirc sign.

8.3 Learn more from DNASTAR tutorials and videos

To learn more about Protean3D DNASTAR provides tutorials online including <u>video</u> tutorials! The web Protean3D page is at: http://www.dnastar.com/t-protean-3d.aspx



From the web page:

"Protean 3D is Lasergene's application for exploring macromolecular structure, motion, and function. Rich, synchronized graphical views allow you to see the 3D molecular structure, the annotated sequence, and the analysis of applied prediction methods simultaneously, enabling easy identification and analysis of secondary structure elements. Protean 3D also provides access to the Motion Library, where you can browse and search over 300 animated and annotated macromolecular conformational changes, and to <u>NovaFold</u>, which allows you to predict three-dimensional structures for protein sequences."

Video tutorials are at: http://www.dnastar.com/t-protean-videos.aspx

| we | Training and Support training videos below to s binar for whatever workflow you need. te: all of the videos below are hosted on YouTu | ee how it's done. Don't see what you're l | | | |
|----|---|--|---|---|---------|
| | ideos available by: Product e ArrayStar With QSeq (27) GeneQuest (8) GenVision (5) MegAlign And MegAlign Pro (17) Protean And Protean 3D (14) SeqMan NGen (41) SeqMan Pro (49) SeqNinja (3) | Protean and Protean 3D DNASTAR - Proteomics Softw DNASTAR - An Introduction 1 DNASTAR - Locating PROSITI DNASTAR - Locating PROSITI DNASTAR - Locating PROSITI DNASTAR - Locating PROSITI DNASTAR - Protein 3D'S Anal DNASTAR - Protein 3D'S Anal DNASTAR - Rotating and Mov DNASTAR - Protein Structure DNASTAR - Protein Structure DNASTAR - Frotein Predictio DNASTAR - Identifying Protea DNASTAR - Ligner Protein Started V DNASTAR - Starte Started V | D Protean 3D View Dotalis Patterns on a Protein Se ng Styles to Highlight Fea Library View Details ysis View View Details pitope Prediction in Prot and Adjusting Lighting on Alignment in Protean 3D 1 Using the DNASTAR Mu se Cleavage Sites Using F ith NovaFold View Details | tures of Interest View Details ean 3D View Details n Structures in Protean 3D View D View Details thod View Details | Search: |
| | TUTORIALS | VIDEOS | REQUES | | FAQS |

9. Quit All Programs

- 1) Quit Protean and/or PyMOL
- 2) Close all windows.



L02 – Supplemental: Primer Design

1. L02 - Setup:



If it does not exist yet, create a folder named L02

| | File | Edit | View | Go | Window | ŀ |
|---|------|----------|----------|----|--------|---|
| | Ne | w Find | ж | N | | |
| t | Ne | w Fold | | N | | |
| | | C | A 17 A 1 | | 37.001 | |

BACKGROUND

Primers are small, oligonucletotides complementary to a specific location of a nucleic acid strand. Primers are used widely in PCR (Polymerase Chain Reaction) to amplify specific regions of nucleic acids, for example genomic DNA fragments or mRNA molecules. In this case primers are designed in pairs. For the generation and subsequent cloning cDNA fragments primers can be designed as a single sequence.

<u>Primer design</u> is one of the most critical parameter in designing PCR experiments as most of the outcome of the experiment depends on many of the primer's properties. Here are some critical or important parameters to consider in the design of primers:

<u>Primer length</u>: oligonucleotides between 18 and 24 bases in length are the most sequence specific (at the optimal annealing temperature). If the primer is too long the annealing is less efficient.

<u>Annealing temperature</u>: It has been determined empirically that a good annealing temperature is approximately 5°C lower than the melting temperature. (Tm).

<u>Melting Temperature</u> (Tm) : a PCR reaction uses 2 primers. Both primers should have a similar melting temperature. If the Tm are too dissimilar the high Tm primer will mis-prime at lower temperatures while the low Tm primer may not be efficient at the higher temperatures. Nearest neighbor thermodynamic calculations is the most accurate method to calculate the melting temperatures of primers. However in the 18 to 24 length range the formula Tm = 2(AT) + 4(GC) can be used as an approximation. A melting temperature of 55°C - 72°C gives the best results.

<u>Specificity</u>: depends in part on the primer length. Repetitive elements in the DNA sequence may be responsible for random annealing resulting in a smear on the gel. Short 3' complementary stretches may cause non-specific annealing at lower temperatures.

<u>Complementarity within primer sequences</u>: Primers should not contain stretches of more than 3 complementary base pairs in a row, as this would result in "snap back" configurations: partially double-stranded structures of the primer folding back onto itself. Similarly, sequence similarity should be limited between the 2 primers used in a reaction to avoid primer dimers.

<u>G/C content, homopolymeric</u>, (T, C) and (A, G) stretches: Ideally the chosen primers will have a near random sequence with between 45 to 55% in GC content. PolyC and polyG promote non-specific annealing while polyA and polyT stretches can facilitate "breathing" after annealing. Similarly polypyrimidine (T, C) and polypurine (A, G) stretches should be avoided.

<u>3'-end sequence</u>: a G or C residue should be included at the 3' end of primers since a "GC Clamp" helps correct binding and minimizes breathing.

- Summary Design Rules -

- 1. Primers should be 17-28 bases in length;
- **2.** Base composition should be 50-60% (G+C);

3. Primers should end (3') in a G or C, or CG or GC: this prevents "breathing" of ends and increases efficiency of priming;

4. Tms between 55-80 °C are preferred;

5. 3'-ends of primers should not be complementary (ie. base pair), as otherwise primer dimers will be synthesised preferentially to any other product;

6. Primer self-complementarity (ability to form 20 structures such as hairpins) should be avoided;

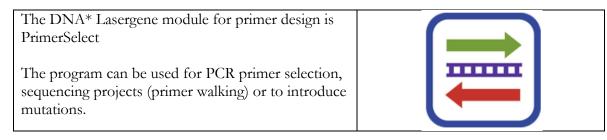
7. Runs of three or more Cs or Gs at the 3'-ends of primers may promote mispriming at G or C-rich sequences (because of stability of annealing), and should be avoided.

<u>Adapted from</u>: Innis, M.A. and D.H. Gelfand. 1990. Optimization of PCRs. In PCR protocols: A guide to methods and applications (ed. M.A. Innis, D.H. Gelfand, J.J. Sninsky, and T.J. White), pp. 3–12. Academic Press, San Diego, CA.-See

http://bioweb.uwlax.edu/GenWeb/Molecular/Seq_Anal/Primer_Design/primer_design.htm#designrules or short URL: http://bit.ly/1jN2U9v

2. DNASTAR PrimerSelect

DNA* Lasergene 12 modules are located within Macintosh HD > Applications > DNASTAR > Lasergene 12



F

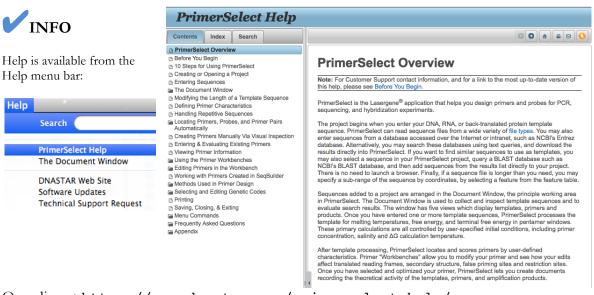
OK

2.1 Launch PrimerSelect

V TASK

• **Double-click** the PrimerSelect icon

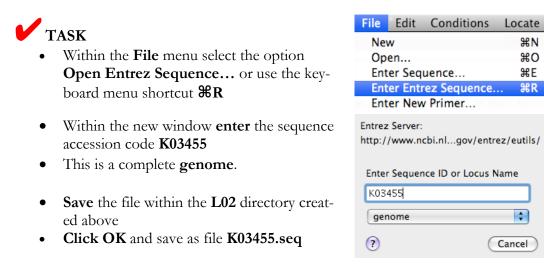
2.1.1 PrimerSelect Help



Or online at https://www.dnastar.com/primerselect_help/ 2.2 Importing a sequence

We will select primers flanking the region of the HIV-1 envelope poly-protein in order to amplify it. The same primer pairs could then be used systematically on various isolates to collect and compare sequences.

Within this sequence the coding region for the envelope poly-protein is from 6225 to 8795.



The new sequence will be shown as a simple long rectangle, shorter than 10,000 bases long within a new, untitled window:

| 0 | 00 | | Untitled | | | | | | |
|--------|-----|------|----------|------|------|--|--|--|--|
| | | | | | | | | | |
| 7 | | | | | | | | | |
| ∆G | 200 | 4000 | 6000 | 8000 | 1000 | | | | |
| ∆G | | | | | | | | | |
| | | | | | | | | | |
| € O | | | | | | | | | |

Various icons at the left of the PrimerSelect window are switches to display various properties of the template sequence, such as the Tm over a window range () or free energy plots (e.g.). The default is the schematic presentation, toggled with the in button at the top.

2.3 Limiting the Annealing Region

Allow approximately 200 bases as a priming area:

| | Conditions Locate Log Repor | t Opt | | | | |
|--|--|-----------------------|--|--|--|--|
| TASK Select Conditions > Primer Loca- | Sequence Positions and Limits $\Re =$ | | | | | |
| tions or choose the shortcut #L | Initial Conditions Primer Characteristics | жк | | | | |
| | Primer Locations | #L | | | | |
| | Mispriming | | | | | |
| • From the pull-down menu at the top of the dialog window select the " Upper and Lower Primer Rang- | Product Length Upper Primer Range and Lower Primer Range and Restrict Locations by V Upper and Lower Primer | Product Length | | | | |
| es" and enter 6000 to 6225 for the | Minimum | Maximum | | | | |
| Upper Primer Locations and 8795 to | | | | | | |
| 9000 for the Lower Primer Location. | Upper Primer Locations: 6000 | 6225 | | | | |
| | Lower Primer Locations: 8795 | 9000 | | | | |
| Click OK | Avoid Locations Containing Repetitive Sequen | ce within 10 bp of 3' | | | | |
| | ? Set to Defaults | Cancel OK | | | | |

Note that triangular marks are now displayed within the schematics presentation

| | 00 | | Untitled | _ |
|----|------|---------|----------|-----------|
| | | | N | |
| 7 | | | | |
| ∆G | 2000 | 4000 | 6000 | 10000 |
| ∆G | | KØ3455. | .seq | |
| | | | | |

2.4 Calculating primer pairs:

TASK

• Select Menu Bar: Locate > PCR Primer Pairs The results appear in a separate window.

| <u> </u> | mary Pairs, 7 | | | | | | | | M | | | | |
|----------|---------------|------|------|------|------|---------------|------|------|-------------|-------|------|-------|--------|
| | 1000 | 2000 | 3000 | 4000 | 5000 | 6000 | 7000 | 8000 | 9000 | Score | ∆Tm | Ta-Tm | Leng |
| | | | | | | | | | | 29.1 | 12.1 | 32.7 | 2.813K |
| | | | | | | ······ D····· | | | ·····• | 27.3 | 5.1 | 28.7 | 2.951K |
| | | | | | | ·····• | | | | 26.5 | 4.0 | 31.7 | 2.714K |
| | | | | | | ······ | | | | 13.5 | 4.2 | 33.5 | 2.871K |
| | | | | | | | | | | 10.2 | 11.8 | 32.5 | 2.865K |
| | | | | | | | | | ·····ā····· | 9.8 | 4.0 | 33.2 | 2.747K |
| | | | | | | | | | | 6.4 | 10.7 | 32.7 | 2.999K |
| Alte | rnate Pai | rs: | | | | | | | | | | | |
| | | | | | | ····· D····· | | | ·····• | 26.9 | 10.6 | 32.7 | 2.814K |
| | | | | | | ·····• | | | ······ | 19.1 | 13.3 | 33.9 | 2.813K |
| | | | | | | | | | | 16.6 | 11.8 | 33.9 | 2.814K |
| | | | | | | ····· | | | ·····ā····· | 14.0 | 9.6 | 32.7 | 2.812K |
| | | | | | | ······ā····· | | | ā | 10.5 | 10.8 | 33.9 | 2.812K |

• On the <u>original "Untitled" window</u> Click on the sequence view icon on the left side

• This will present the sequence, in split windows

|) 🔘 |) 🔿 Unt | titled 2 | | | | |
|----------------------|--|--|--|--|--|--|
| | | | | | | |
| ™ 5' ∆G 3' | стоютссятаютаята поставляется по стоются по стоются по стоются по стоются по стоются по стоются по станования по | GRTGGGTGGCRARGTGGTCRARARAGTAGTGTGTATTGGATGGCCTRCTGTARGGGRARGARTGA CTACCCACCGTTCACCAGTTTTTCATCACCACTAACCTACCGGATGACATTCCCTTTCTTACT TCACACTAACCTACCCGATG | | | | |
| | 6120 6140 6150 Натантанская тактотототототототополата таконала с 160 | 88/99 88/20 88/49 САТБССТОССАНСТССТАНАЛАСТАСТСТАНТССАТСССТАСТСТАНСССАНАСАНСАНТСА | | | | |

2.5 Evaluating primer pairs: menu "Report".

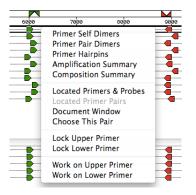
VINFO

A report on primer pairs can be called on either from the menu bar "Report" or using the keyboard shortcut **#D**.

| Report | Options | Net Searc | h Wir |
|--------|-------------|-----------|-------|
| Primer | Self Dime | 'S | ЖD |
| Primer | Pair Dime | rs | _ |
| Primer | Hairpins | | - |
| Amplif | ication Sur | nmary | |
| Compo | osition Sun | nmary | - |
| Locate | d Primers d | & Probes | ж[- |
| Locate | d Primer Pa | airs | ¥] |
| Docum | nent Windo | w | ¥. |

| Locate | Log | Report | Option | | | |
|-------------------------|---------|--------|--------|--|--|--|
| Primers & Probes #F | | | | | | |
| Only Catalogued Primers | | | | | | |
| PCR Pr | rimer P | Pairs | ЖY | | | |

| | PCR Prim | er Pairs | S | жү |
|----------------|----------|-----------|---------|---------|
| d Primer Pairs | | | | |
| <u> </u> | | | | đ |
| 7000 8000 9000 | Sec | ore ∆T | m Ta-Tm | Length |
| | 29 | 9.1 12. | 1 32.7 | 2.813KE |
| | 27 | ?.3 5. | 1 28.7 | 2.951Kb |
| | | i.5 4. | 0 31.7 | 2.714КЬ |
| | 13 | 3.5 4. | 2 33.5 | 2.871Kb |
| | | 9.2 11. | 8 32.5 | 2.865Kb |
| | ç | 9.8 4. | 0 33.2 | 2.747КЬ |
| | e | 5.4 10. | 7 32.7 | 2.999Kb |
| | | | | |



Another method is to **right-click** on the primer to be studied.

We will use the first pair at the top of the list as the print-out examples.

Using either method to report (Report menu or right-click) explore the 5 menu options:

Primer Self Dimers

| \varTheta 🔿 🔿 Self Dimer Formation |
|--|
| 11 dimens found. 11 dimens found. 16052.0271,2 bp. 40 = -3.1 kc/m (worst= -35.0) 5' ROTRCHTGTRACCARCCTAT 3' 1 3' TATCCARCECARTGTRCATGA 5' |
| +[68526072], 2 bp, 40 = -2.0 kc/m (worst= -35.0) 5' AGTACHTOTAGGERACTAT 3' 1 11 11 3' TATCCARCGCARTOTACATGA 5' |
| +[68526072], 2 bp, 46 = -1.6 kc/m (worst= -35.0) 5' AOTACATOTARCCCARCTAT 3' 1 1 1 1 3' TATCCARCCCARTOTACATOR 5' |
| +[60526072], 2 bp, ΔG = -1.5 kc/m (worst= -35.0) 5' AGTACATGTAACGCAACCTAT 3' |
| 3' TATCCARCGCARTGTACATGA 5' |
| -[88648841], 2 bp, 40 = -1.3 kc/m (worst= -43.7) 5' GCTCTTCTTTTTTCCTTACH 3' 1 1 3' ACHTTCCCTTTCTTACTCCCC 5' |
| +(60526072), 2 bp, 40 = -1.3 kc/m (worst= -35.0) 5' AGTACATOTACCORACTAT 3' 1 1 1 1 1 1 1 1 3' TATCCARCCCARTOTACATOR 5' |
| +[60526072], 2 bp, 46 = -1.3 kc/m (worst= -35.0) 5' AGTACATGARGEGRACETAT 3' 3' TATCCARGEGRATETACATER 5' |
| +[60526072], 2 bp, ∆0 = -1.0 kc/m (worst= -35.0) 5' AGTACATGTARGGGAACCTAT 3' 11 3' TATCCARGCGAATGTACATGA 5' |
| +[60526072], 2 bp, 40 = -1.0 kc/m (worst= -35.0) 5' AGTACATGTARCGCARCCTAT 3' 1 3' TATCCARCGCARTGTACATGA 5' |
| -[88648841], 2 bp, AG = -1.0 kc/m (worst= -43.7) 5' GCTCGTCTCATTCCCTTRCA 3' 3' ACATTCCCTTTCTTATCTGCTCG 5' |
| +[68526072], 2 bp, 46 = -1.0 kc/m (worst= -35.0) 5' AGTACRIGTARCCCARCCTAT 3' 1 3' TATCCARCECARTGTACATON 5' |

Report > Primer Pair Dimers

| \varTheta 🔿 🔿 🛛 🖡 | air Di | mer | Formatio | n | | | |
|---|--------|-----|-----------|------|---------|----------------|---|
| 8 dimers found. | | | | | | | _ |
| +[60526072] vs[886488 5' AGTACATGTAACGCAACCTAT 3' ACATTCCCTTTCTTAC | 31 | ., | ∆G = −6.2 | kc/m | (bad!) | (worst= -35.0) | > |
| +[60526072] vs[886488 5' AGTACATGTAACGCAACCTAT 3' ACATTCCCTTTCTTACTCTGCT | 3' | bр, | ∆G = −2.3 | kc∕m | (worst= | -35.0) | |
| +[60526072] vs[886488 5' AGTACATGTAACGCAACCTAT 3' ACATTCCCTTT | 3111 | ., | | kc∕m | (worst= | -35.0) | |
| +[60526072] vs[886488 5' AGTACATGTARCGCAACCTAT 3' ACATTCCCTTTCTTACT | 3111 | ., | ∆G = -1.9 | kc∕m | (worst= | -35.0) | |
| +[60526072] vs[886488 5' AGTACATGTARCGCAACCTAT 3' ACATTCCCTTTCTTACTCTG | 3111 | ., | ∆G = −1.9 | kc/m | (worst= | -35.0) | |
| +[60526072] vs[886488 5' AGTACATGTAACGCAACC 3' ACATTCCCTTTCTTACTCGCT | TAT 3' | bр, | ∆G = -1.6 | kc∕m | (worst= | -35.0) | l |
| +[60526072] vs[886488 5' AGTACATGTAACGCAACCTI 3' ACATTCCCTTTCTTACTCTGCT | ат З' | bр, | ∆G = −1.3 | kc∕m | (worst= | -35.0) | l |
| +[60526072] vs[886488 5' AGTACATGTARCGCARCCTAT 3' AGTACATGTARCGCARCCTAT | 31 | | | | (worst= | -35.0) | l |
| 3 HOHITO | | THU | 000100 5 | | | | |

• Menu Bar: Report Primer Hairpins

| | Hairpin Forma | ation | |
|--|---------------------|---------------|--------------------|
| Primer: +[60526072 Primer: -[88648841 +[60526072], 2 bp 5' AGTACATOTA A 3' TATCCAACGC | 1], 1 hairpins, | but the prime | r is OK r is OK |
| +[60526072], 2 bp 5' AGTACA 3' TATCCAACGCAATO | а-, Т | 2.8 kc/m | |
| +[60526072], 2 bp 5' AGTACATGTARC~ 3' TATCCAACG~ | (Loop=6), ∆G =] | 3.0 kc/m | |
| +[60526072], 2 bp 5' AGTACATGTAACGO 3' TATCCAF | ີ່່ | 3.1 kc/m | |
| +[60526072], 2 bp 5' AGTACATGTAA 11 1 3' TATCCAACGC | (Loop=8), ∆G = | 3.1 kc/m | |
| +[60526072], 2 bp 5' AGTACATGTAACG- 3' TATCCAAC- | ר ר | = 3.4 kc/m | |
| +[60526072], 2 bp 5' AGTAC 3' TATCCAACGCAAT(| ²⁸ | 3.5 kc/m | |
| +[60526072], 2 bp 5' AGTACATGT~ 3' TATCCAACGCAA~ | | = 3.6 kc/m | |
| +[60526072], 2 bp 5' AGTACATGT A 3' TATCCAACGCA- | (Loop=3), ∆G = | 3.6 kc/m | |
| -[88648841], 2 bp 5' GCTCGTCTCATTC- 3' ACATTCCCTT- | (Loop=15), ΔG | = 4.5 kc/m | |

• Menu Bar: Report > Amplification Summary

| Amplification Summary | | | | | | | | | |
|--|---|---|--|--|--|--|--|--|--|
| | | | | | | | | | |
| Upper Primer: 21-mer 5' AGTACATGTARCGCARCCTAT 3' Lower Primer: 24-mer 5' GCTCGTCTCATTCTTTCCCTTACA 3' | | | | | | | | | |
| DNA 250 | pM, Sal | t 50 mM | | Upper Prim | er Lowe | r Primer | | | |
| | Tm Overall S Location | Stability | | 43.6 °C -35.0 kc/ 6052607 | m -43 | 5.7 °C 1.7 kc/m i48841 | | | |
| Product Tm - Primer Tm 32.7 °C Primers Tm Difference 12.1 °C Optimal Annealing Temperature 51.6 °C | | | | | | | | | |
| Product Product | : Length : Tm (%GC : GC Conte : Tm at 6; | ent | | 2813 bp 76.3 °C 40.7% 97.9 °C | | | | | |
| | Product | Melting | Temper | ature (%GC | Method) | | | | |
| | Salt | | | Form | amide | | | | |
| mМ | ×SSC | ×SSPE | ØX | 10% | 20% | 50% | | | |
| 1 10 50 165 330 500 1000 | 0.005 0.051 0.256 0.846 1.692 2.564 5.128 | 0.005 0.052 0.312 1.031 2.052 3.125 6.250 | 48.1 64.7 76.3 84.9 89.9 92.9 97.9 | 69.8 78.4 83.4 86.4 | 35.1 51.7 63.3 71.9 76.9 79.9 84.9 | 15.6 32.2 43.8 52.4 57.4 60.4 65.4 | | | |
| 195 | 1.000 | 1.219 | 0.0 | %formamic | de = Tm 8 | 36.1 °C | | | |

Menu Bar: Report > Composition Summary

| | J | |
|------------|---|--|
| 0 | | |
| \bigcirc | | |

| \varTheta 🔿 Compos | sition Summa | ary | | | | |
|--|-------------------------------------|---|---|--|--|--|
| Upper Primer: 21-mer 5' AGTACATGTAACGCAACCTAT 3' Lower Primer: 24-mer 5' GCTCGTCCATTCTTTCCCTTACA 3' | | | | | | |
| Primers | Upper Prime | r Lowe | r Primer | | | |
| Single Strand Mr Extinction Coefficient 1/E | 6.5 k 4.72 nM/A26 30.6 μg/A26 | | | | | |
| Product | Composition | Quantity | Per Cent | | | |
| Upper Strand Mr 874.6 k Lower Strand Mr 863.4 k Both Strands Mr 1737.9 k Length 2813 bp Tm (XGC Method) 76.3 °C Tm at 6xSSC 97.9 °C GC Content 40.7% | A C G T I A+T G+C | 991 467 677 678 0 1669 1144 | 35.2 16.6 24.1 24.1 0.0 59.3 40.7 | | | |

2.6 Locating Primers



Primers location along the sequence can be reviewed with the either the menu cascade Report > Located Primer Pairs or with a right-click > Located Primers & Probes on top of a primer within the "Located Primer Pairs" graphical window as illustrated above.

The primer pairs selected within the "Located Primer Pairs" graphical window will appear highlighted within the resulting report window.

| • |) () (| | Lo | cated | Primer | s | |
|-----|----------|-------|-----------|-------|--------|----------|------|
| √ | Start | End | Length | Tm | ΔG | ∆Profile | Name |
| Upp | ber Prim | mers: | 15 Locate | ed | | | |
| < | 6000 | 6019 | 20-mer | 43.9 | -32.7 | 4.1 | |
| √ | 6031 | 6050 | 20-mer | 39.0 | -32.0 | 6.4 | |
| √_ | 6031 | 6052 | 22-mer | 41.4 | -34.3 | 6.6 | |
| ٧, | 6047 | 6069 | 23-mer | 46.5 | -37.8 | 9.6 | |
| 1 | 6047 | 6070 | 24-mer | 47.9 | -39.4 | 16.4 | |
| √ | 6052 | 6071 | 20-mer | 42.4 | -33.6 | 17.8 | |
| √ | 6052 | 6072 | 21-mer | 43.6 | -35.0 | 23.9 | |
| ٧, | 6125 | 6144 | 20-mer | 40.6 | -32.1 | 7.5 | |
| √ | 6125 | 6145 | 21-mer | 41.4 | -33.1 | 8.4 | |
| √ | 6125 | 6146 | 22-mer | 43.2 | -34.6 | 9.5 | |
| ٧, | 6127 | 6150 | 24-mer | 44.7 | -37.3 | 10.0 | |
| √ | 6151 | 6171 | 21-mer | 38.7 | -33.5 | 2.6 | |
| √ | 6151 | 6172 | 22-mer | 41.0 | -35.1 | 2.8 | |
| ٧, | 6163 | 6186 | 24-mer | 47.7 | -39.8 | 1.1 | |
| √ | 6198 | 6221 | 24-mer | 47.5 | -38.5 | 3.9 | |
| | ver Prim | mers: | 19 Locate | ed | | | |
| < | 8805 | 8826 | 22-mer | 46.9 | -36.3 | 3.3 | |
| √ | 8806 | 8826 | 21-mer | 43.7 | -34.4 | 2.2 | |
| √ | 8807 | 8826 | 20-mer | 41.3 | -32.4 | 1.3 | |
| √ | 8821 | 8840 | 20-mer | 48.7 | -36.1 | 50.1 | |
| √_ | 8825 | 8846 | 22-mer | 51.2 | -39.3 | 35.9 | |
| √_ | 8826 | 8846 | 21-mer | 48.2 | -37.3 | 34.2 | |
| √ | 8840 | 8863 | 24-mer | 53.3 | -42.2 | 14.5 | |
| √ | 8841 | 8864 | 24-mer | 55.7 | -43.7 | 55.6 | |
| ٧, | 8842 | 8865 | 24-mer | 54.2 | -43.3 | 37.8 | |
| ٧, | 8855 | 8871 | 17-mer | 47.2 | -32.6 | 18.5 | |
| √ | 8875 | 8896 | 22-mer | 59.1 | -44.0 | 73.4 | |
| ٧, | 8899 | 8922 | 24-mer | 61.1 | -46.5 | 11.1 | |
| √ | 8900 | 8923 | 24-mer | 58.2 | -44.9 | 10.2 | |

2.7 Upper and lower primers

It is possible to specify whether you want to work specifically with one of the primers within a primer pair with the Edit menu bar options: Work on Upper Primer and Work on Lower Primer. These 2 options can also be engaged from the right-click menu shown above.

| | Edit Conditions Loca | te Log |
|--|----------------------|------------|
| TASK | Can't Undo | ЖZ |
| Engage working with the upper primer of the first line | Cut | ЖX |
| with e.g. the menu cascade | Сору | жc |
| with e.g. the menu caseade | Paste | жv |
| | Clear | |
| Edit > Work on Upper Primer | Select All | ЖA |
| | Work on Upper Primer | % 1 |
| On the keyboard shortout 921 | Work on Lower Primer | Ж2 |

Or the keyboard shortcut **#1**

This will pop up yet another graphical window labeled "Upper Primer Workbench" showing the primer bound to the sequence.

NOTE: To work on UPPER primer only:

• Double-click on specific primer from within the report window Upper Primer WorkBench <u>00 0</u> Length = 21, Tm = 43.0 Y TEATE TATCAAAGCAGTAAG ▶ÀGTACATGTAACGCAACCTAT TTCTTCGCCTCTGTCGCTGCTTCTCGAGTAGTCTTGTCAGTCTGAGTAGTTCGAAGAGATAGTTTCGTCAT з:ГС GGTTATCATCGTTATCATCGTAATCATCGTT CATGTACATTGCGT GerPheSerIleLysAlaValSerberThrCysAsnAlaThrTyrThrAsnSer GerThrCysAsnAlaThrTyr GluGluAlaGluThrAlaThrLysSerSerGerGluGlnSerAsp Primer Frame 2 sGinterValValHisValThrGinProIisProIieValAlaIieValAlaLeuValValAlaI gLysLysArgArgGlnArgArgArgAldHisGlnAsnSerGlnThrH ValValHisValThrGlnProIls Primer Frame 3 lyArgSerGlyAspSerAspGluGluLeuIleArgThrValArgLeuIleLysLeuLeuT Primer terTyrMetterArqAsnLeuTyr

=0 \$ SerLysterTyrMetterArgAsnLeuTyrGInterterGInterterHisterterGIn Priming 4000 1000 2000 3000 5000 6000 7000 8000 9000 Sites No dimers > 2 bp Hairpin 2 bp, 2.4 ko/m 5' AGTACATGTA-C Cancel OK Name: Note: No Frames Frame 1 Frame 2 Frame 3 **Top Strand Frames** The buttons at left are used to show or hide the transla-Frame 4 tion frames all of them shown by default. Frame 5 Frame 6 Bottom Strand Frames All Frame





3. Designing Primers in SeqBuilder

Acknowledgment: this section is adapted directly from the guide "Getting Started with DNASTAR® Lasergene® - For Macintosh® and Windows® - Version 8.0"

SeqBuilder enables you to design primers for regions of interest on your sequence. If desired, once a primer pair is selected and modified, SeqBuilder allows you to cut and clone the PCR insert (with corresponding primer features) into a vector.

The data for the tutorials in this section can be found in the following location:

Hard Drive > Applications > DNASTAR > Lasergene 12 > Demo Data

3.1 Creating and Modifying Primer Pairs for a Region of Interest

Objective: To create a primer pair for a region of interest on your sequence.

V TASK

- Launch SeqBuilder.
- Select File>Open to navigate to the Demo Data folder and choose Tn5wPCR.sbd. The sequence opens in the Linear Map view.
- Choose the neomycin/kanamycin resistance feature labeled aminoglycoside-3'-O-phosphotransferase as the region of interest by single clicking on that feature. The sequence range 1559-2353 will be selected.

| 00 | @ Tn5wPCR.sbd |
|--------------------------|---|
| 🔻 🚞 Views | R Selection: Top strand 15592353, length = 795 5834 bp |
| Sequence | Acc51 Banli Kasl BsrFI BspMPfol BmB15d101 Bpml BlogtAlli Sfol BsrFI Sspl |
| 📓 🛛 Feature List | ABvandpil BrneT1101 Nari Sapi BsriFilogi Bsrki7ati Banili BuzantT110Bbel NrneAlli BrneT1140incil Bisevi/I Aval Bisevi-Alli Bolet BsrFi AvalSali HBnsadiX'i Xmrni B1Ka731 BstAPi Bocqi Bsevi/Itsol |
| Comment | GSBadol BsrFI BsrFilmeAlli <mark>BspMil Bse</mark> YI Sammal NarNmeAlKasi BsrFI Bytanni Hindili Psti Banil BseYi HinNari Bogi B5BatAPI Kasi <mark>Gsal BravaNitSbi</mark> lBmeT1N100H BogiNN168anil BseYi BsrFi BsrFi KasiGsal |
| 📥 Linear Map | Ssp\$rfol Bogl' BpmAlwN <mark>I Narl Bahll</mark> FtoBsnBbgbe⊫ BamBbbel BsnFBmeTAlMobbyl Masal Bsal AvaNarl Sapl |
| 🖒 Circular Map | Kpri Ka8abpi Ahdi Hin8abp <mark>ii Stol. Tsol</mark> BprMenincil38167i BespMAlii ABaa3Yi BespMali Tsol GeberneAliPtol Ptol Nn6adali Tsol AniilBse <mark>YiBbel NmeAliiSa</mark> pi BanikA8gogiKa8adaXi Bs16878gi Gs88meT.8656abi AlwBisa Ahdi Sa¢abbel AlwNi |
| Primer Design | BbBeev(Psti BsrFi Skral PstiBsp51)hi E⊧prBicg1'KasaNari SfolBc5915g1 AvAhmBe4mBerFi BohmeeAlli Bogi SfoTsol Banli |
| Primer List | 500 1000 1 <mark>500 2000</mark> 2500 3000 3500 4000 4500 5000 5500 |
| Minimap | |
| Site Summary | transposon Tn5 |
| Enzymes Displayed | |
| ▶ 🚞 🗹 Features Displayed | aminoglysferase |
| ORFs | |
| ▶ 🧰 Ruler | protein #3 |
| | protein #4 |
| | |
| | (►) (►) (►) (►) (►) (►) (►) (►) (►) (►) |
| Views mRN | IA 🗾 91558 |
| Feature mRM | IA ➡ 14683658 |
| Feature mich | |

• Select menu **Priming>Create Primer Pairs**. The Create Primer Pairs dialog appears.

| Priming | Format | View | Net Se | arch |
|----------|------------|--------|--------|------|
| Create | Primer Pai | rs | | |
| Create | Primer Fro | m Sele | ction | • |
| Pair Sel | ected Prin | ners | | |
| | | | | |

| • | Click the triangles next to Condi- |
|---|--------------------------------------|
| | tions and Primer Characteristics |
| | to expand those sections so that the |
| | dialog appears as follows: |

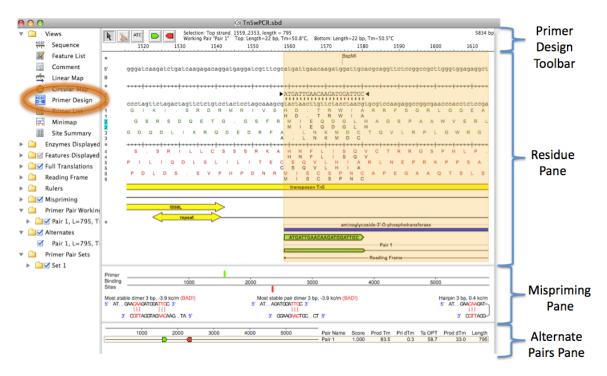
| $\Theta \odot \odot$ | Create Primer Pairs |
|----------------------|--|
| Locations | |
| | • Primers end exactly at selection |
| | Choose optimal primer location |
| | Stay within 0 bp of selection |
| | Display top 1 primer pairs |
| | Amplify 5' - 1559 to 2353 - 3' |
| rimer Cha | Salt concentration 50.0 mM |
| | ngth: minimum 17 maximum 24 |
| Target Tr | m 60.0 C 3' pentamer stability 8.0 -kc/M |
| Unique at | t 3' for 12 bp Ignoring duplexing 2 bp from 3' |
| Avoid | repeats in catalog: Choose |
| ? Defaul | ts Cancel OK |

INFO: The **Locations** section of the Create Primer Pairs dialog allows you to specify parameters that define where SeqBuilder will search for primers on your sequence. The **Conditions** section allows you to specify the initial salt concentration, which affects the calculation of predicted melting temperature. The **Primer Characteristics** section enables you to limit your search for primers based on primer length, target melting temperature, and primer interactions at the 3' end. You also have the option in this dialog to avoid known repetitive sequences that are likely to cause mispriming.

• Adjust the Target Tm to 55.0°C. Leave the rest of the values as they are, and then click OK. SeqBuilder will choose the best primer pair with the characteristics speci-

fied that lies exactly within the selected region and display it in the Primer Design View. By default, the view focuses on the top strand primer first.

• If necessary, **resize the Primer Design View** pane by clicking and dragging the dividing bar between panes so that you can see all 3 sections of the view: the Residue Pane, the Mispriming Pane, and the Alternate Pairs Pane:



- Notice in the Mispriming Pane, the **Most stable 261-mer** and **Most stable pair dimer** are <u>both</u> labeled "BAD!". This indicates that the final pentamer value of the dimer exceeds the 3' pentamer stability threshold defined in the Primer Characteristics section of the Create Primer Pairs dialog. In general, you should avoid using primers with dimers or hairpins labeled "BAD!" unless you have experimental evidence that they will function as desired, or if you have no other option.
- Switch to the Primer List view by selecting it from the curtain shown on the left side of the SeqBuilder screen. The single primer pair located by your search is displayed.



• Click the triangle next to Set 1 to expand the primer pair and view each individual primer within the pair:

| 000 | | | | | 🕑 Tn5wPCR.sbd | | | | |
|---------------|-----|----------|-----------|-----|---|--------|-----------------|-----------------|---------|
| Views | | Set Name | Pair Name | T/B | Primer Sequence | Length | Tm | dG | Current |
| Sequence | | ▼ Set 1 | Pair 1 | | <atgattgaacaagatggat< td=""><td>795</td><td>Tm=83.5, TaOPT=</td><td>=58.7, dTm=33.0</td><td>~</td></atgattgaacaagatggat<> | 795 | Tm=83.5, TaOPT= | =58.7, dTm=33.0 | ~ |
| | | | | Т | ATGATTGAACAAGATGGATTC | 22 | 50.8 | -38.6 | |
| 🌠 🛛 Feature L | ist | | | В | TCAGAAGAACTCGTCAAGAA | 22 | 50.5 | -38.2 | ন |
| Comment | | | | | | | | | |
| 떀 Linear Ma | p | | | | | | | | |

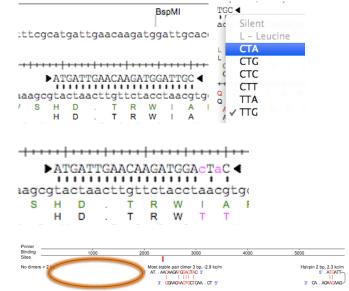
- Notice that the melting temperatures for the individual primers are both close to 50° C, which is lower than our Target Tm of 55°C.
- Switch back to the Primer Design view to view the top strand primer:

| R | 📐 ATC 🕞 | | ction: Top strand king Pair "Pair 1" | | ength = 52 22 bp, Tm=50.8°C | , Bottom: Le | ngth=22 bp, Tm=5 |
|----|---|---|---|---|--------------------------------|--------------|---|
| | 1520 | 1530 | 1540 | 1550 | 1560 | 1570 | 1580 |
| 0 | | | | | | | BspMI |
| 5' | ggatcaaga | tctgatcaa | gagacaggato | gaggatcgtt | tege <mark>atgattg</mark> | aacaagatg | gattgcacgcag |
| в | | | | | | | |
| 0 | +++++++++++++++++++++++++++++++++++++++ | +++++++++++++++++++++++++++++++++++++++ | +++++++++++++++++++++++++++++++++++++++ | +++++++++++++++++++++++++++++++++++++++ | | ••••• | +++++++++++++++++++++++++++++++++++++++ |
| т | | | | | ► ATGATTG | AACAAGATG | |
| 3' | cctagttct | agactagtt | ctctgtcctad | etectageaa | | | ctaacgtgcgtd |

• Select the codon change mode by clicking on the button in the Primer Design Toolbar. Notice that your cursor changes and that each codon is highlighted as you hover over the sequence.



- Introduce a codon change by clicking on the **TTG** triplet in your primer, located at 1577-1579 of your sequence. From the list of codons given, select **CTA**.
- Notice that changes to the primer sequence and translation are shown in magenta.
- Notice also that the Mispriming Pane has now updated to show that there is no longer a stable dimer conformation.



•

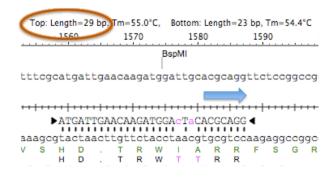
- Display the bottom strand primer by • Selectio ATC clicking the **button** from the Pri-Working mer Design Toolbar. 1550 1560 2330 2340 2350 The Primer Design o View scrolls to show ctatcgccttcttgacgagttcttctgageggg ▶GGAAGAACTGCTCAAGAAGACT◀ 5 the bottom strand priв o mer in the center of the т view: 3' gatagcggaagaactgctcaagaagactcgccc P S.RVLL S.RVLL 1 s 1
- Increase the length of the bottom strand primer from 22 bp to 23 bp by clicking • and dragging the triangle shown on the left (3') end of the primer sequence. The length of the primer is displayed and updated automatically in the Primer Design Toolbar shown at the top of the view:

| ATC | | Position: 1578 Working Pair "Pair 1" | Top: Leng | th=22 bp, Tm=44.1° | C, Bottom: Lei | ngth=23 bp | m=54.4°C |
|------|------|---|-----------|--------------------|----------------|------------|----------|
| 2300 | 2310 | 2320 | 2330 | 2340 | 2350 | 2360 | 2370 |

Notice that once the primer length is increased to 23 bp, the melting temperature for • the bottom strand primer increased to 54.4°C, which is much closer to our Target Tm. Also notice that the **Most stable pair dimer** has been updated in the Mispriming Pane and no longer has the indication "BAD!"

Switch back to the top strand •

primer by clicking on the button. Increase the length of the top strand primer from 22 bp to 29 bp by clicking and dragging the triangle shown on the **right (3') end** of the primer sequence. Notice that the melting temperature of the top strand primer has now increased to 55.0°C.



- Now that the primers have been modified to our satisfaction, change back to the • Primer List view by selecting it from the curtain on the left.
- Expand the pair again by clicking on the triangle next to Set 1. Notice that the edits to our primer pair are reflected here as well. Nucleotide changes to the primer sequence are shown in lowercase.

| | | | | ⑦ Tn5wPCR.sbd | | | | |
|---|----------|-----------|-----|--|--------|----------------|-----------------|---------|
| | Set Name | Pair Name | T/B | Primer Sequence | Length | Tm | dG | Current |
| • | Set 1 | Pair 1 | | <atgattgaacaagatggactacacgcagg><gcct< td=""><td>795</td><td>Tm=83.5, TaOPT</td><td>=59.9, dTm=29.1</td><td></td></gcct<></atgattgaacaagatggactacacgcagg> | 795 | Tm=83.5, TaOPT | =59.9, dTm=29.1 | |
| | _ | | Т | ATGATTGAACAAGATGGAcTaCACGCAGG | 29 | 55.0 | -46.8 | |
| | | | В | TCAGAAGAACTCGTCAAGAAGGC | 23 | 54.4 | -41.4 | |

• Name the modified pair by clicking on Pair 1 in the Pair Name column and typing Modified pair.

| | Set Name | Pair Name | T/B | Primer Sequence | Length | Tm | dG | Current |
|---|----------|---------------|-----|--|--------|-----------------|-----------------|---------|
| • | Set 1 | Modified pair | 1 | <atgattgaacaagatggactacacgcagg><gcct< th=""><th>795</th><th>Tm=83.5, TaOPT=</th><th>=59.9, dTm=29.1</th><th>~</th></gcct<></atgattgaacaagatggactacacgcagg> | 795 | Tm=83.5, TaOPT= | =59.9, dTm=29.1 | ~ |
| | | | т | ATGATTGAACAAGATGGAcTaCACGCAGG | 29 | 55.0 | -46.8 | ~ |
| | | | В | TCAGAAGAACTCGTCAAGAAGGC | 23 | 54.4 | -41.4 | ~ |
| | | | | <i>&\beta</i> | | | | |

4. Primer3, free web-based application

Acknowledgment: Steve Rozen and Helen J. Skaletsky (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365-386. http://jura.wi.mit.edu/rozen/papers/rozen-and-skaletsky-2000-primer3.pdf Archive link: http://bit.ly/1j5Nj48

Note: Primer3 is open source: http://primer3.sourceforge.net/

Objective: create a primer based on a free web resource

4.1 Launch web browser

TASK: Open <u>One</u> the following URL:

```
http://frodo.wi.mit.edu/
```

Above URL will switch to: http://bioinfo.ut.ee/primer3-0.4.0/ http://fokker.wi.mit.edu/primer3/input.htm

```
Above URL will switch to: http://bioinfo.ut.ee/primer3-0.4.0/primer3/input.htm http://biotools.umassmed.edu/bioapps/primer3 www.cgi
```

4.2 Enter data

READ As input sequence you can use for the HIV-1 reference genome you retrieved in a previous exercises. Simply open EditSeq to retrieve the version saved earlier with the accession number: K03455, and use the Copy/Paste method.

As a target sequence we will this time focus on the viral polymerase (*pol*) which spans nucleotides 2358 to 5096.



- Open K03455.seq with EditSeq
- Select and Copy sequence in the clipboard
- Switch to the web browser
- Under Sequence ID enter K03455 or another ID meaningful to you
- Under **Targets** enter **2358,200 5096,200** (Enter numbers on the same line, separated by blank space. The represent the surrounding regions for the target primers around the coding region of *pol*)

<u>Note</u>: This selection is different than the one we used in DNASTAR PrimerSelect. This selection will create two 200 length targets, one at the beginning and one at the end of the coding region of *pol*.

INFO

We will leave the other optional areas within the page such as

General Primer Picking Conditions Other Per-Sequence Inputs Objective Function Penalty Weights for Primers etc.

with their respective default values which are sometimes blanks.

The top portion of your browser should look similar to the following:

| r | | | | | | | | |
|--|---|--|--|------------------------|--------------------------|--|--|--|
| Primer3 (v. 0.4.0) Pick primers from a DNA sequence. | | Primer3plus interface More primer/oligo tools Old (0.3.0) interface | | disclaimer cautions | Primer3 Home FAO/Wiki | | | |
| [(, | | | 0.3.0) interface | cautions | FAQ/WIKI | | | |
| Paste source sequence below (5'->3', string undesirable sequence (vector, ALUs, LIN recare protected page of the sequence (vector, ALUs, LIN recare protected page of the sequence of the sequence age attracted page of the sequence of the sequence treage to sequence of the sequence of th | Es, etc.) or use a <u>Mispriming Li</u> agagagaagtgttagagtggggggtttgacagcog agaactgctgacatcgagcttgctacaagggact gggagtgggggggccctcagatcctgcatataagc tgggagcticttggctaactagggaacccact | brary (repeat library) | | STA format | ok. Please N-out | | | |
| | | | | | | | | |
| Pick left primer, | Pick hybridization probe (i | nternal | Pick right primer, or use right primer below | | | | | |
| or use left primer below: | oligo), or use oligo below: | | (5' to 3' on opposite strand): | | | | | |
| | | | | | | | | |
| Excluded E.g. | 50,2 requires primers to surroun ATCT[CCCC]TCAT means 401,7 68,3 forbids selection of p ence with < and >: e.gATCT | that primers must fla primers in the 7 bases | ank the central CCCC. starting at 401 and the 3 base | es at 68. Or n | , | | | |
| Product Size Ranges 150-250 100-300 301 | -400 401-500 501-600 601-700 70 | 01-850 851-1000 | | | | | | |
| Number To Return 5 | Max 3' Stability 9.0 | | | | | | | |
| Max Repeat Mispriming 12.00 Pair | Max Repeat Mispriming 24.00 | | | | | | | |
| Max Template Mispriming 12.00 Pair Ma | ax Template Mispriming 24.00 | | | | | | | |
| Pick Primers Reset Form | | | | | | | | |
| General Primer Picking Condition | ons | | | | | | | |
| | D | NASTAR Las | ergene: PrimerS | elect - | page - 265 | | | |



• Click on the "Pick Primers" button

A new window will appear summarizing results.

• Explore the proposed data.

If you get result for only one of the region, try each Target region separately: 2358,200 alone and then 5096,200 alone.

>>>>> and <<<<< mark the primer sequence on the result page and ****** mark the product

Example output:

Primer3 Output

PRIMER PICKING RESULTS FOR K03455

| No mispriming library specified Using 1-based sequence positions OLIGO <u>start len tm</u> <u>gc%</u> any <u>3' seq</u> LEFT PRIMER 5058 20 59.97 50.00 3.00 0.00 ggtgatgattgtgtggcaag RIGHT PRIMER 5354 21 58.95 52.38 4.00 1.00 tggtctgctagttcagggtct SEQUENCE SIZE: 9719 INCLUDED REGION SIZE: 9719 |
|--|
| PRODUCT SIZE: 297, PAIR ANY COMPL: 3.00, PAIR 3' COMPL: 2.00 TARGETS (start, len)*: 2358,200 5096,200 |
| 1 tggaagggctaattcactcccaacgaagacaagatatccttgatctgtggatctaccaca |
| 2341 atacagtattagaagaaatgagtttgccaggaagatggaaaccaaaaatgatagggggaa ********************** |
| 2401 ttggaggttttatcaaagtaagacagtatgatcagatactcatagaaatctgtggacata |
| 2461 aagctataggtacagtattagtaggacctacacctgtcaacataattggaagaaatctgt ********************************** |
| 2521 tgactcagattggttgcactttaaattttcccattagccctattgagactgtaccagtaa ********************************** |
| 2581 aattaaagccaggaatggatggcccaaaagttaaacaatggccattgacagaagaaaaaa |
| 5041 atggaaaacagatggcaggtgatgattgtgtggcaagtagacaggatgaggattagaaca >>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>> |
| 5101 tggaaaagtttagtaaaacaccatatgtatgtttcagggaaagctaggggatggttttat |
| 5161 agacatcactatgaaagccctcatccaagaataagttcagaagtacacatcccactaggg |

```
5221 \ gatgctagattggtaataacaacatattggggtctgcatacaggagaaagagactggcat
           ******
 5281 ttgggtcagggagtctccatagaatggaggaaaaaggagatatagcacacaagtagaccct
                                                                                       <<<<<
 5341 gaactagcagaccaactaattcatctgtattactttgactgtttttcagactctgctata
         <<<<<<<
 5401\ agaaaggccttattaggacacatagttagccctaggtgtgaatatcaagcaggacataac
KEYS (in order of precedence):
***** target
>>>>>> left primer
<<<<< right primer
ADDITIONAL OLIGOS
                             start len
                                                  tm gc% any 3' seq
 1 LEFT PRIMER 5058
5355

        5058
        20
        59.97
        50.00
        3.00
        0.00
        ggtgatgattgtgtggcaag

        5355
        22
        60.30
        50.00
        4.00
        1.00
        ttggtctgctagttcagggtct

    PRODUCT SIZE: 298, PAIR ANY COMPL: 3.00, PAIR 3' COMPL: 2.00

        59.97
        50.00
        3.00
        0.00
        ggtgatgattgtgtggcaag

        59.00
        50.00
        4.00
        2.00
        tggtctgctagttcagggtcta

                              5058 20
5354 22
 2 LEFT PRIMER
    RIGHT PRIMER
    PRODUCT SIZE: 297, PAIR ANY COMPL: 3.00, PAIR 3' COMPL: 1.00
 3 LEFT PRIMER
                              5058 20
                                                59.97 50.00 3.00 0.00 ggtgatgattgtgtggcaag
                              5354 23 59.82 52.17 4.00 2.00 tggtctgctagttcagggtctac
    RIGHT PRIMER
    PRODUCT SIZE: 297, PAIR ANY COMPL: 3.00, PAIR 3' COMPL: 1.00

        60.26
        52.38
        3.00
        0.00 tggcaagtagacaggatgagg

        58.95
        52.38
        4.00
        1.00 tggtctgctagttcagggtct

 4 LEFT PRIMER
                               5071
                                         21
                              5354
    RIGHT PRIMER
                                       21
    PRODUCT SIZE: 284, PAIR ANY COMPL: 5.00, PAIR 3' COMPL: 1.00
Statistics
            contooinnotmhighhighsidmanytarexclbadGCtootooany3'
                                                                                                             high
                                                                                                   poly
                                                                                                            end

        ered
        Ng
        etcl
        reg
        GC%
        clamp
        low
        high
        compl
        compl

        42479
        0
        226
        0
        995
        0
        18788
        10805
        0
        0

        62569
        0
        223
        0
        1723
        0
        29856
        13976
        0
        12

                                                                                                     X stab
                                                                                                                        ok
Left 42479
                                                                                                     297
                                                                                                              616 10752
Right 62569
                                                                                              12 257 770 15752
Pair Stats:
considered 180834628, no target 1335657, unacceptable product size 179498950, high end
compl 3, ok 18
primer3 release 1.1.0
```

(primer3_results.cgi 0.4.0 modified for WI)

5. DNASTAR Resources and Support

Support for DNASTAR Lasergene is at http://www.dnastar.com/t-support-training.aspx : This is a repository of tutorials on all DNASTAR software and not only Lasergene.

DNASTAR Lasergene: PrimerSelect - page - 267



The "Getting Started with DNASTAR® Lasergene® - For Macintosh® and Windows® - Version 8.0" publication is available online, the direct link to the PDF file (*GettingStartedGuide8.0.pdf*) is on an FTP server. Quick link: http://bit.ly/lsuGSLx

6. Ending session

- **Quit** all programs
- **Close** all Macintosh windows.
- Move to the trash the files you created today



