

DNAMAN

Sequence Analysis Software

*Translation and
Protein Sequence
Analysis*

BIOINFORMATICS PLATFORM

Translation and Protein Sequence Analysis

- ▶ Translation from DNA Sequence
- ▶ Reverse Translation
- ▶ Protein Properties
- ▶ Protein Secondary Structure

Translation from DNA Sequence

1. Double click a sequence file from File control to load sequence into channel
2. Click Translation tool from Protein Analysis tab

The screenshot shows the DNAMAN software interface. The 'Protein & Database' tab is active, and the 'Translation' tool is highlighted with a red circle and the number '2'. The 'File' control shows a list of sequence files, with 'AY599226.seq' selected and circled in red with the number '1'. The main window displays the sequence details for 'AY599226.seq', including composition, keywords, and features. The 'ORIGIN' section shows the DNA sequence with its corresponding amino acid translation.

DNAMAN - AY599226.seq

Home Sequence Protein & Database View

Translation 2
Reverse Translation
Codon Optimization
Codon Usage RF1
Hydrophobicity Plot 1 Seq
Hydrophobicity Plot MultiSeq

Reading Frame Overview
Manager
Sequence Database
Oligo Database

File AY599226.seq

DNAMAN Files

Name

Name	Size
AF311601.seq	9
AF311601_Tran...	4
AJ403983.seq	8
AY599226.seq	5
erwinia_genom...	8,101
EXAMPLE1.SEQ	2
EXAMPLE2.SEQ	2
EXAMPLE3.SEQ	2
EXAMPLE22017...	2
EYAMDI E22017	3

AY599226.seq

SEQ AY599226 2934
Composition 746 A; 729 C; 717 G; 742 T; 0 OTHER
Percentage 25.4% A; 24.8% C; 24.4% G; 25.3% T; 0.0% OTHER
MW (kDa) 905.64 ssDNA 1808.79 dsDNA
KEYWORDS CIRCULAR
COLOURS
sequence = 1
features = 0

FEATURES

source 1..2934
/source="Cloning_vector"
misc_feature 582..671
/attributes="MCS; multiple_c;"
gene complement(1874..2734)
/name="bla"
cds complement(1874..2734)
/name="beta-lactamase"

ORIGIN

```
1 TCGCGCGTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA
61 CAGCTTGCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG TCAGGCGCGC TCAGCGGGTG
121 TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA GCAGATTGTA CTGAGAGTGC
181 ACCATATGCG GTGTGAAATA CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGGCGCC
241 ATTCGCCATT CAGGCTGCGC AACTGTTGGG AAGGGCGGAT GGTGCGGGCC TCTTCGCTAT
301 TACGCCAGCT AGAGGACCAG CCGCGTAAAC TGGCAAAATC GGTACGGTT GAGTAATAAA
361 TGGATGCCCT GCGTAAGCGG GTGTGGGCGG ACAAATAAGT CTTAAACTGA ACAAATAGA
```

Ready Channel 1:DNA AY599226 2934bp CAP NUM

Translation from DNA Sequence

1. Choose One Letter AA Symbol
2. Choose Frame one from RF
3. Choose No from Show DNA
4. Select Use Only when Annotations available
5. Check cds to use only cds in the DNA sequence for translation
6. Results shows up after clicking OK

Translation Parameters

AA Symbol
 One letter **1**
 Three letter

Reading Frame
 Frame 1 **2**
 Frame 2
 Frame 3
 All frames

Show DNA
 No **3**
 One strand
 Double strand

Space between 3-letter code

Use only **4** Exclude

cds **5**
 gene
 misc_feature
 source

20 AA per line
AA sequence to channel
NO

OK Cancel

```
AY599226_Translation
SEQ      AY599226_AA
Translation of AY599226(1-861)
Universal code
Total amino acid number: 286, MW=31503
ORIGIN
1      MSIQHFRVAL IPFFAFAFLP VFAHPETLVK VKDAEDQLGA RVGYIELDLN SGI
61     EERFPMMSTF KVLICGAVLS RIDAQEQLG RRIHYSQNDL VEYSPVTEKH LTI
121    CSAAITMSDN TAANLLTTI GGPKELTAFI HNMGDHVTRL DRWEPELNEA IPI
181    FVAMATTLRK LLTGELLTLA SRQQLIDWME ADKVAGPLLR SALPAGWFIA DK
241    RGIIAALGPD GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW*
```

6

Translation from DNA

- Overview

With a sequence loaded into channel

1. Click Reading Frame Overview

Overview tool from Protein Analysis tab

2. A graphic window displays potential translation sequences in all six reading frames

The screenshot shows the DNAMAN software interface. The 'Protein & Database' tab is active, and the 'Reading Frame Overview' tool is selected. A tooltip for the tool explains its function: 'Display the distribution of Met and Stop codons on current sequence in all RFs'. The main window displays a graphic window titled 'AY599226: Universal code' showing the distribution of Met and Stop codons across six reading frames (plus 1, plus 2, plus 3, minus 1, minus 2, minus 3). The sequence length is 176 bp. The plus frames show blue arrows indicating potential start sites, and the minus frames show red arrows indicating potential start sites. A red '1' is placed over the 'Reading Frame Overview' button in the toolbar, and a red '2' is placed over the graphic window.

DNAMAN - [AY599226_Translation]

Home Sequence Protein & Database View

Translation
Reverse Translation
Codon Optimization
Codon Usage RF1
Hydrophobicity Plot 1 Seq
Hydrophobicity Plot MultiSeq

Manager
Sequence Database
Oligo Database

Reading Frame Overview
Display the distribution of Met and Stop codons on current sequence in all RFs

AY599226: Universal code

1 293 586 880 1173 1467 176

Strand RF
plus 1
plus 2
plus 3
minus 1
minus 2
minus 3

Name
AF311601.seq
AF311601_Tran
AJ403983.seq
AY599226.seq
erwinia_genom
EXAMPLE1.SEQ
EXAMPLE2.SEQ
EXAMPLE3.SEQ
EXAMPLE22017
EXAMPLE22017

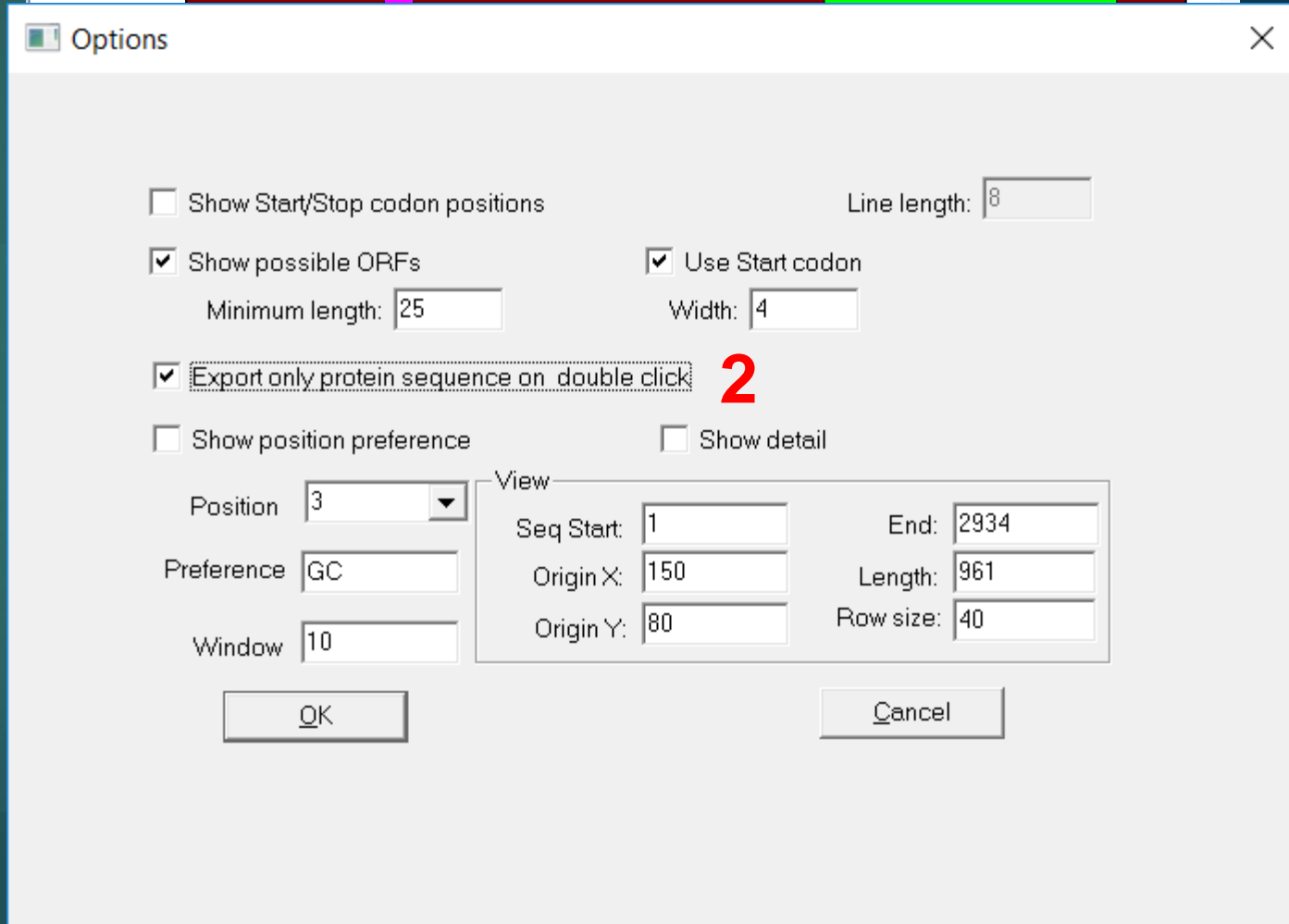
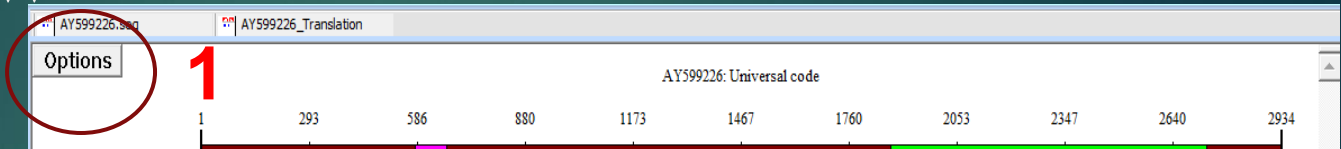
Lo... Ch... Da... File

Channel 1:DNA AY599226 2934bp CAP NUM

Translation from DNA

- Overview

1. Option button in the Overview graphic window
2. In the Option box, check Export only protein sequence on double click



Translation from DNA

- Overview

1. Double click a reading frame translation element
2. A Text window pops up to show actual translation

Options

AY599226: Universal code

Strand RF

plus 1

plus 2

plus 3

minus 1

minus 2

minus 3

beta-lactamase

1 293 586 880 1173 1467 1760 2053 2347 2640 2934

SEQ R.C.AY599226{RF-2: 1874-2732}_AA
Translation of R.C.AY599226{RF-2: 1874-2732} (1-858)
Universal code
Total amino acid number: 285, MW=31317
Original_FEATURES
gene complement (2..288)
/name="bla"
cds complement (2..288)
/name="beta-lactamase"

ORIGIN

1 MSIQHFRVAL IPFFAAFCPLP VFAHPETLVK VKDAEDQLGA RVGYIELDLN SGKILESFRP
61 EERFPMMSTF KVLCCGAVLS RIDAGQEQLG RRIHYSQNDL VEYSPVTEKH LTDGMTVREL
121 CSAAITMSDN TAANLLTTI GGPKELTAFI HNMGDHVTRL DRWEPELNEA IPNDERDTTM
181 PVAMATTLRK LLTGELLTIA SRQQLIDWME ADKVAGPLLR SALPAGWFIA DKSGAGERGS
241 RGIIAALGPD GKPSRIVVIY TTGSQATMDE RNRQIAEIGA SLIKHW

1

2

Reverse Translation

- Setup

With a protein sequence loaded into channel

1. Click Reverse Translation tool
2. Select protein sequence from channel
3. Protein sequence can be edited.
4. Select Use alternatively all codons

If codon preference file available, DNAMAN use it to optimize reverse translation

The screenshot shows the DNAMAN software interface. The 'Reverse Translation' tool is selected in the 'Protein & Database' tab, indicated by a red circle and the number 1. The 'Reverse Translation' dialog box is open, showing the 'Sequence Channel' dropdown menu set to 'CH. 2:R.C.AY599226_AA', indicated by a red circle and the number 2. The protein sequence is displayed in the dialog box, with a red circle and the number 3 highlighting the sequence. The 'Use alternatively all codons' radio button is selected, indicated by a red circle and the number 4. The 'OK' and 'Cancel' buttons are visible at the bottom of the dialog box.

Reverse Translation

Sequence Channel: CH. 2:R.C.AY599226_AA 286

```
00001 MSIQHFRVALIPFFAAFCPLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPE
00062 ERFPMMSTFKVLLCGAVLSRIDAGQEQLGRRIRHYSQNDLVEYSPVTEKHLTDGMTVRELCS
00123 AAITMSDNTAANLLLTITIGGPKELTAFLHNMGDHVTRLDRWEPELNEAIPNDERDTTMPVA
00184 MATTLRKLLTIGELLTLASRQQLIDWMEADKVGAPLLRSALPAGWFIADKSGAGERGSRGII
00245 AALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGASLIKHW
```

Show all possible codons
 Use alternatively all codons
 Use the most frequently used codon(s)
 Use alternatively the two most frequently used codons
 Use all codons according to preference
 Codon preference file

OK Cancel

Ready Channel 2:Protein R.C.AY599226_AA 286aa CAP N

Reverse Translation

- Results

The screenshot displays the DNAMAN software interface. The main window is titled "DNAMAN - Reverse_Translation". The "View" tab is active, showing a toolbar with various analysis tools. A "Reverse_Translation" channel is selected, and a window titled "Reverse_Translation" is open, displaying the following text:

```
Reverse translation/Optimization according to Universal code  
ORIGIN  
1   ATGCTATTC AACATTTTCGTTGCTTTAATCCCTTTCTTTGCCGCATTCTGTTTGCCC  
61  GTCTTTGCCCACCCAGAAACTCTTGAAAAGTGAAGGATGCTGAGGACCAGCTCGGTGCC  
121 CGCGTTGGCTATATAGAACTAGATCTGAATCCGGAAAAATTTAGAGTCATTCCGACCG  
181 GAAGAGCGGTTTCCTATGATGTCGACCTCAAGGCTTTGCTTTGCGGGCAGTACTCAGT  
241 AGAATCGACGCGGGTCAAGAACAGCTAGGCAGGCGTATACATTACAGCCAAAACGATCTG  
301 GTGGAGTATTCTCCCGTTACAGAAAAACACTTAACGGACGGAATGACTGTCCGCGAGTTG  
361 TGTTCCGCTGCCATTACCATGTCAGATAATACAGCAGCGAACCTTCTCCTAACGACTATC  
421 GGGGGTCCAAAGGAACTGACCGCTTTTTTACATAATATGGGCGACCACGTAACACGATTG
```

The status bar at the bottom indicates "Ready" and "Channel 2:Protein R.C.AY599226_AA 286aa CAP N".

Protein Properties

pI and Charges

The screenshot shows a software interface for protein analysis. The main window is titled 'Protein Charge' and displays the following information:

- Protein ID: R.CAY599226_AA
- Protein pI: 5.92
- Number of disulfide bonds (-S-S- bonds): 1
- Number of fragments: 1
- Buffer pH: 7.00
- Protein charge: -4.37

The 'Charge Table' window is open, showing a table of charges at various pH values:

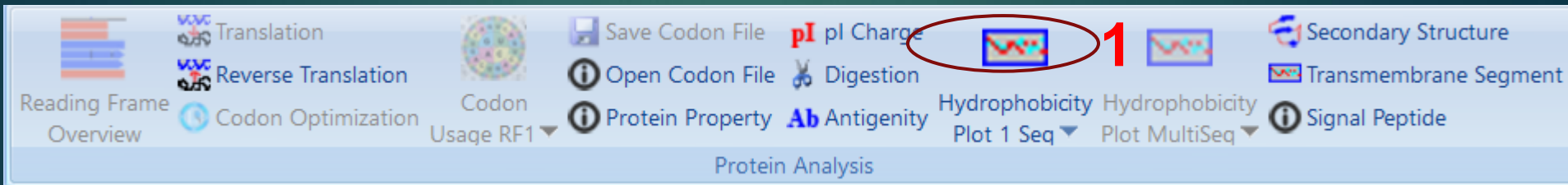
pH	Charge
2.00	37.62
2.20	37.40
2.40	37.06
2.60	36.54
2.80	35.74
3.00	34.55
3.20	32.84
3.40	30.46
3.60	27.37
3.80	23.64
4.00	19.52
4.20	15.38
4.40	11.62
4.60	8.48
4.80	6.04
5.00	4.24
5.20	2.93
5.40	1.94
5.60	1.15
5.80	0.43
6.00	-0.31
6.20	-1.11
6.40	-1.97
6.60	-2.84
6.80	-3.66
7.00	-4.37
7.20	-4.96

With a protein sequence loaded into channel

1. Click pI Charge tool
2. Select number of disulfide bonds
3. Click Charge Table button to display charges vs. pH (4)

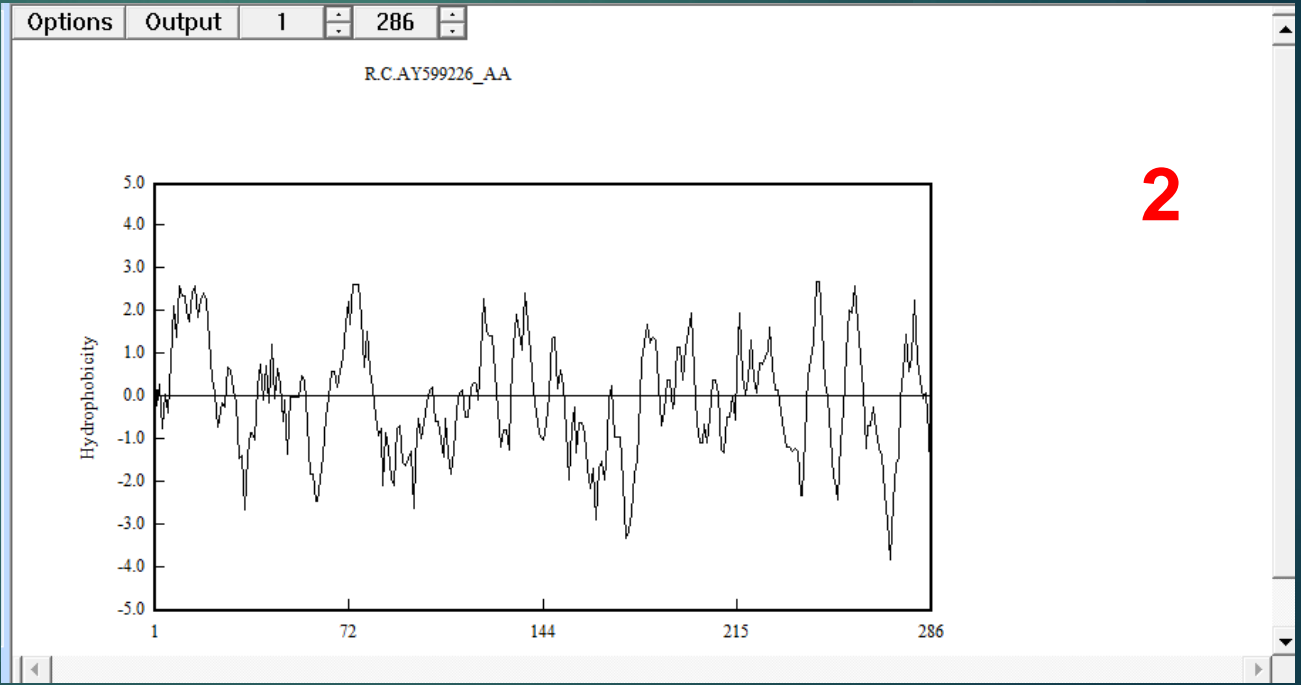
Protein Sequence Plots

Hydrophilicity/Hydrophobicity



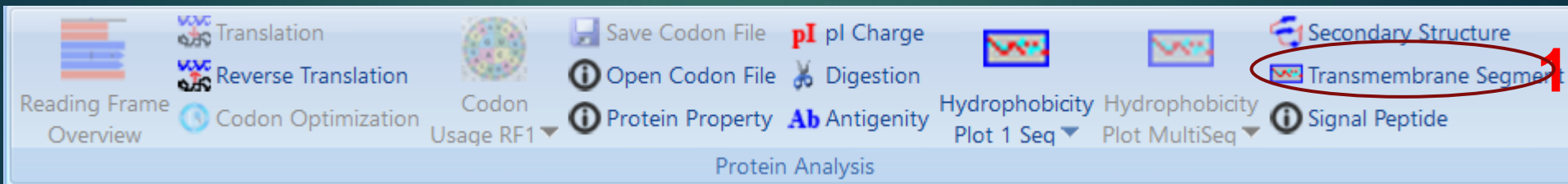
With a protein sequence loaded into channel

1. Click Hydrophobicity or Hydrophilicity tool
2. Plot is displayed in a Graphic window



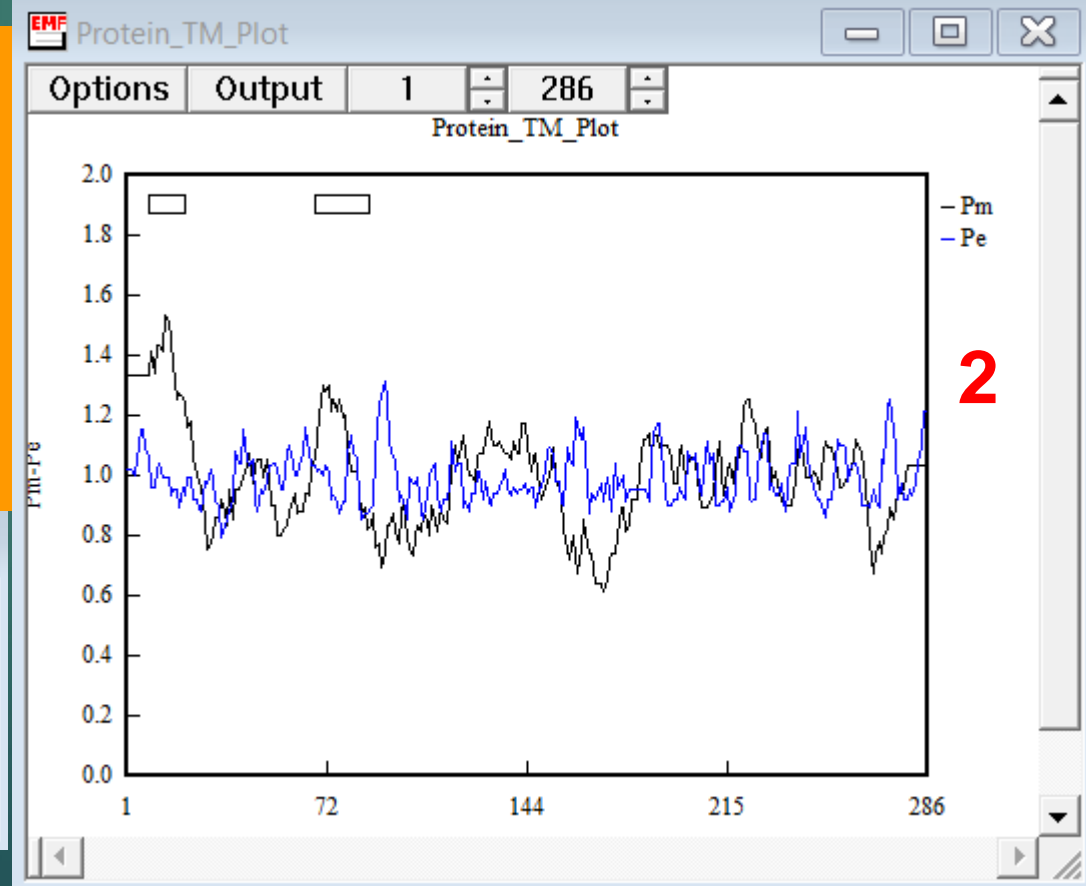
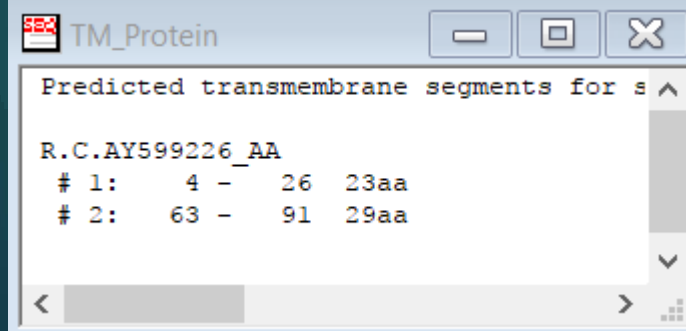
Protein Sequence Plots

Transmembrane Segments



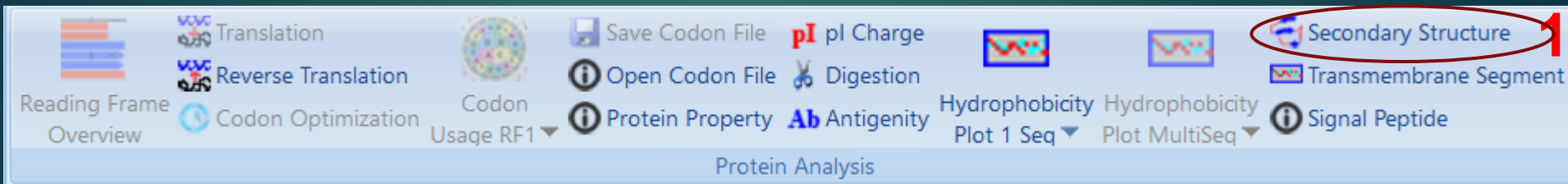
With a protein sequence loaded into channel

1. Click Transmembrane Segment tool
2. Structure plot is displayed in a Graphic window and transmembrane segments are shown in a Text window



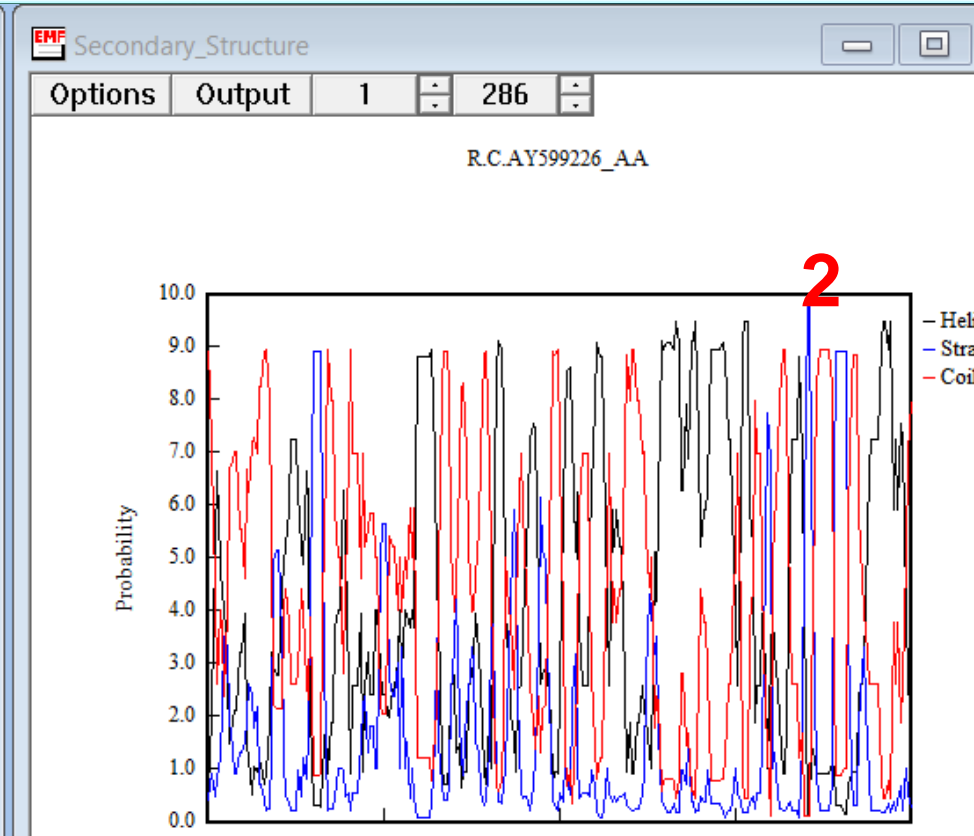
Protein Sequence Plots

Secondary Structures



Untitled

Structure	Probability	Table		
Pos aa	STR	HELIX	STRAND	COIL
	Coil	0.07	0.04	0.89
	Coil	0.11	0.07	0.82
	Coil	0.21	0.09	0.70
	Helix	0.52	0.05	0.44
	Helix	0.66	0.08	0.26
	Helix	0.48	0.12	0.40
	Helix	0.42	0.22	0.36
	Helix	0.42	0.35	0.24
	Coil	0.28	0.33	0.40
	Coil	0.14	0.19	0.67
	Coil	0.19	0.13	0.68
	Coil	0.21	0.09	0.70
	Coil	0.21	0.09	0.70
	Coil	0.32	0.13	0.55
	Coil	0.32	0.13	0.55
	Coil	0.39	0.15	0.46
	Coil	0.14	0.19	0.67
	Coil	0.12	0.26	0.62
	Coil	0.05	0.24	0.71
	Coil	0.10	0.17	0.73
	Coil	0.09	0.22	0.69
	Coil	0.11	0.07	0.82
	Coil	0.11	0.07	0.82
	Coil	0.07	0.04	0.89
	Coil	0.09	0.02	0.89



With a protein sequence loaded into channel

1. Click Secondary Structure tool
2. Structure plot is displayed in a Graphic window and structure prediction is shown in a Text window