

# **CAS REGISTRY<sup>SM</sup>: BLAST<sup>®</sup> similarity searching via STN Express<sup>®</sup>**

November 2008



## Table of Contents

Overview of content.....	3
Accessing CAS REGISTRY BLAST .....	4
Searching similar sequences.....	5
Viewing CAS REGISTRY BLAST results .....	9
Searching by sequence identifier .....	10
Creating alerts .....	11
Getting STN data for BLAST results.....	12
Creating a report with BLAST alignment data .....	14
For more information .....	16

## Overview of content

You can conduct similarity searching of peptides and nucleotides in the CAS REGISTRY database by using the BLAST<sup>®</sup> (Basic Local Alignment Search Tool) algorithm. CAS REGISTRY BLAST is available only for commercial accounts.

### Unique content

- Search sequences for peptides and nucleotides from GenBank<sup>®</sup> as well as from patents and journals indexed by CAS
- Find patents, journals, and other technical literature on sequences in many of the sci-tech databases available on STN, such as BIOSIS<sup>®</sup>, CAplus<sup>SM</sup>, USPATFULL and AGRICOLA

### Coverage

- Naturally occurring and synthetic proteins and nucleic acid molecules
- Regulatory elements and other defined genetic elements
- Genetically engineered proteins or genes
- Protein sequences deduced from gene translation
- Oligonucleotide probes and PCR primers
- Chemically modified peptides, proteins, and nucleotides
- Sequences with uncommon amino acids or uncommon nucleotides
- Fusion proteins, genes, and nucleic acids
- Multi component protein complexes
- Cyclic peptides, peptide metal complexes, and peptide nucleic acid (PNA) sequences

### Unparalleled update frequency

- REGISTRY is updated daily

### Flexibility

- Import or cut and paste large sequences (up to 50,000 characters) from an electronic file
- Use the common sequence file formats (ASCII, GCG, and FASTA)

### Security

- 128-bit encryption

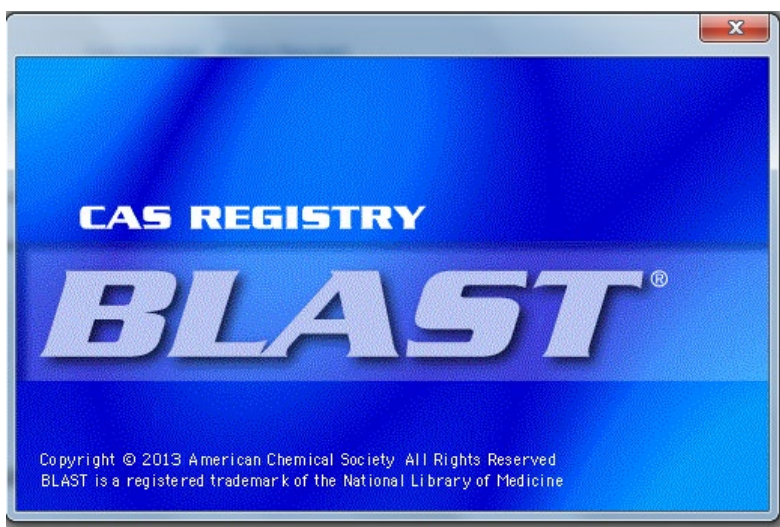
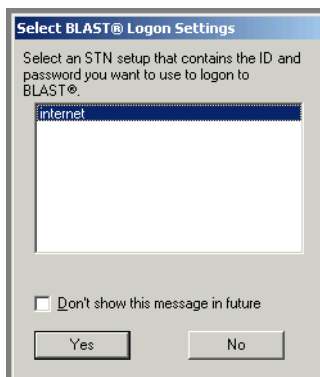
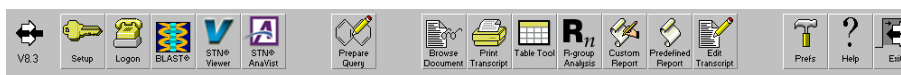
- 1 Click the **BLAST** button on the STN Express toolbar.



- 2 Select the desired setup.

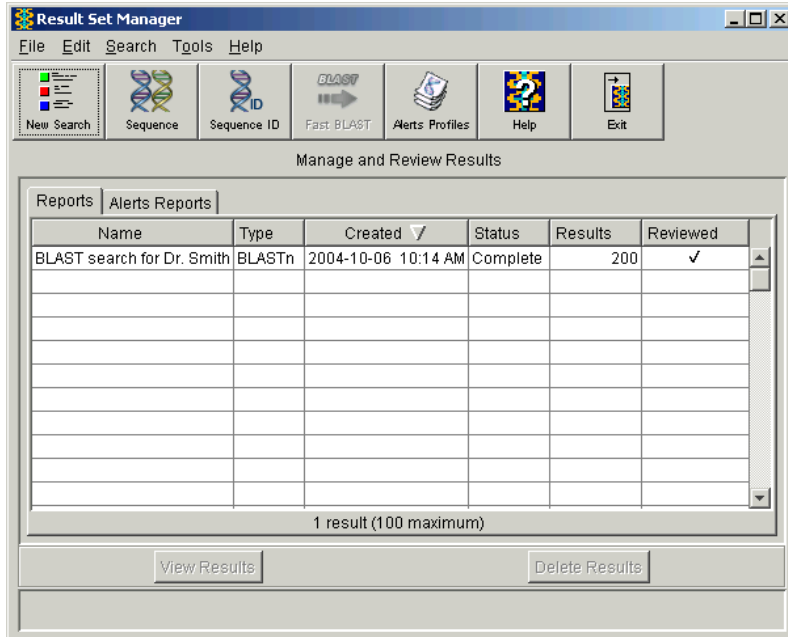
The CAS REGISTRY BLAST splash screen is displayed to indicate that you are connected.

## Accessing CAS REGISTRY BLAST

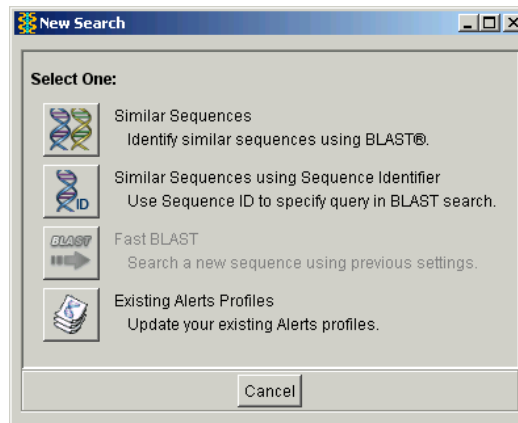


## Searching similar sequences

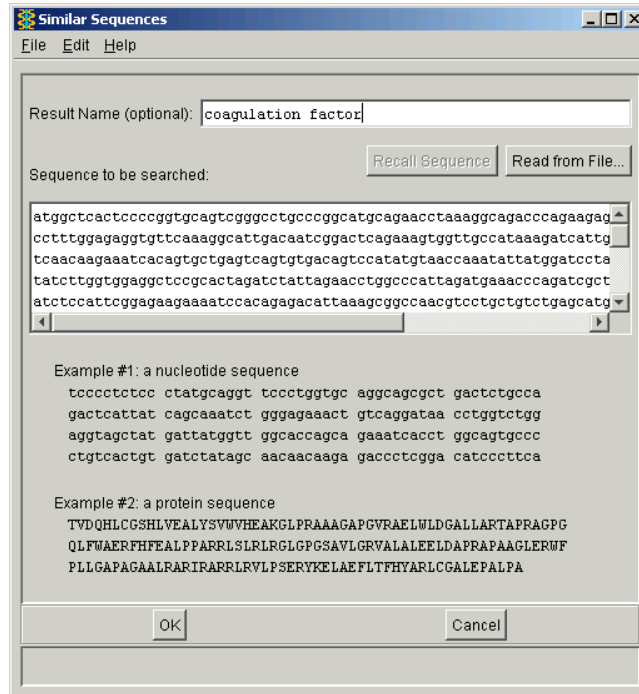
1 In the Results Set Manager window, you can select a report from a previous search or click **New Search** to conduct a new BLAST search.



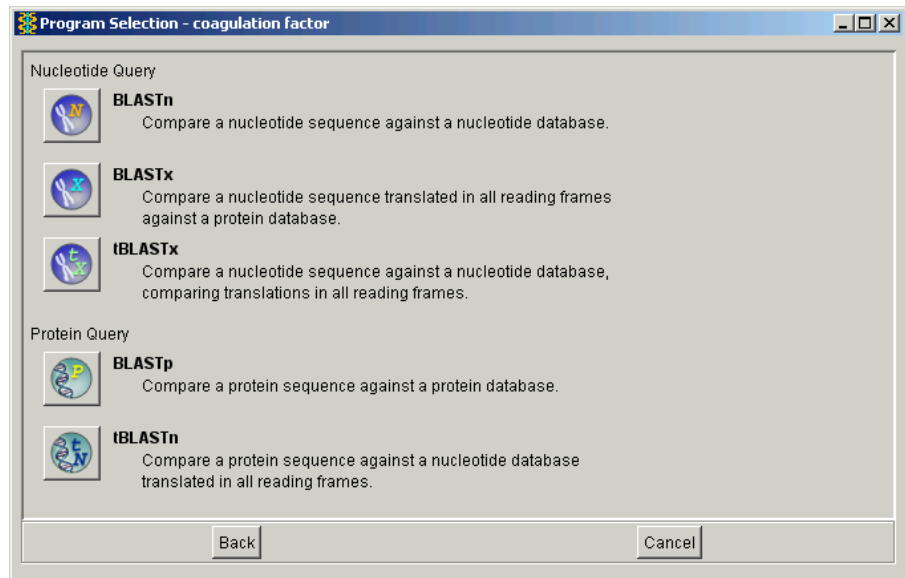
2 Click **Similar Sequences**.



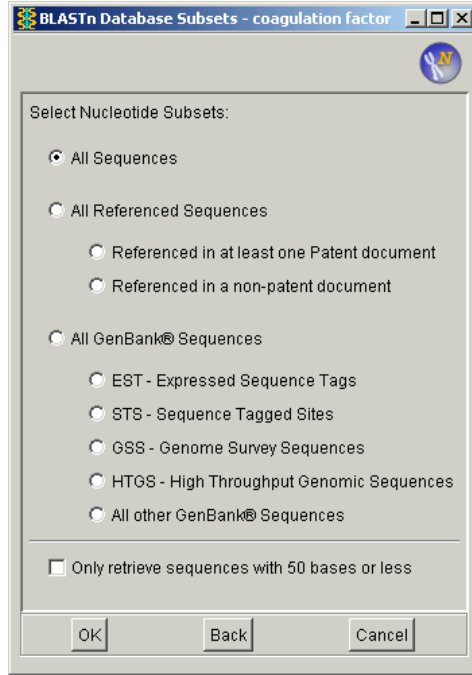
3 Enter a name to track and save the search results and the searched sequence. You can enter the sequence by typing it, copying/pasting it from another application, or "opening", "inserting", or "reading" a sequence from an existing file. The limit for entering sequences is 50,000 characters. Click **OK**.



4 Select a BLAST program.



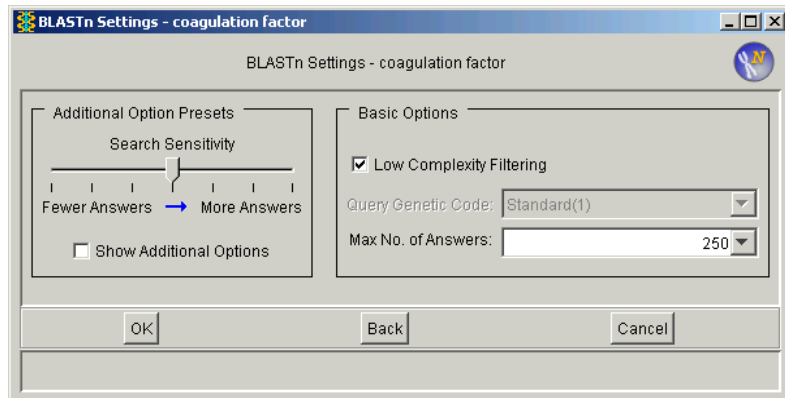
5 Select the entire database or a database subset.



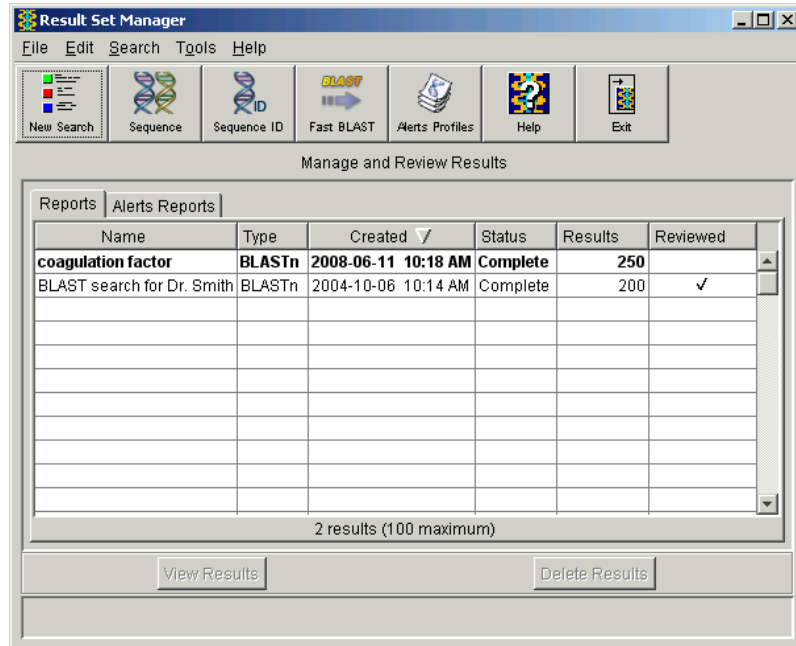
6 Specify the BLAST settings. Enter the maximum number of answers to be retrieved or accept the default. Click **OK**.

For more information on BLAST settings, refer to the NCBI BLAST web page:

[www.ncbi.nlm.nih.gov/BLAST/](http://www.ncbi.nlm.nih.gov/BLAST/)



The Result Set Manager window automatically shows you the status of your search. If you close this window before the search is complete, the search continues to run in the background. You can view the results after **Complete** is displayed in the Status column.

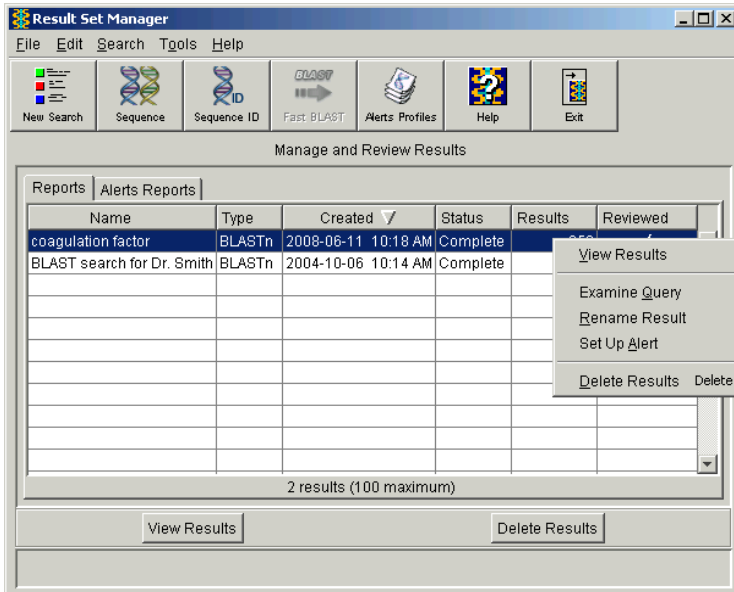


Results are kept in the table until you have 100 sets of results. When you reach 100 sets of results, you have to delete one set before you can conduct another search.



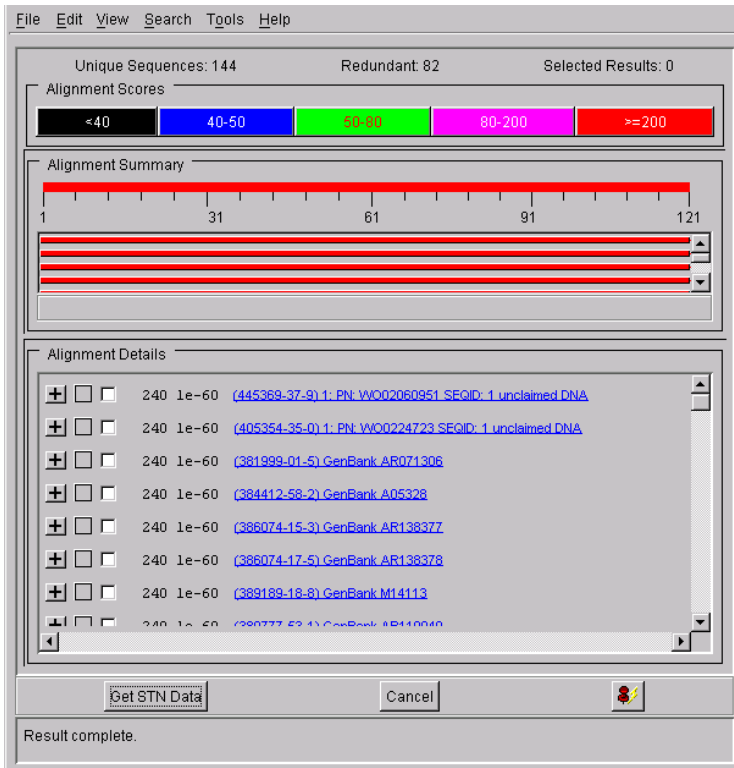
## Viewing CAS REGISTRY BLAST results

- 1 In the Result Set Manager window, right-click on the set of results and select **View Results**.



The BLAST results report shows:

- Number of sequences in the results
- Alignment scores
- Alignment summary with each line representing an individual answer, and the color of each line corresponding to the score range
- Alignment details

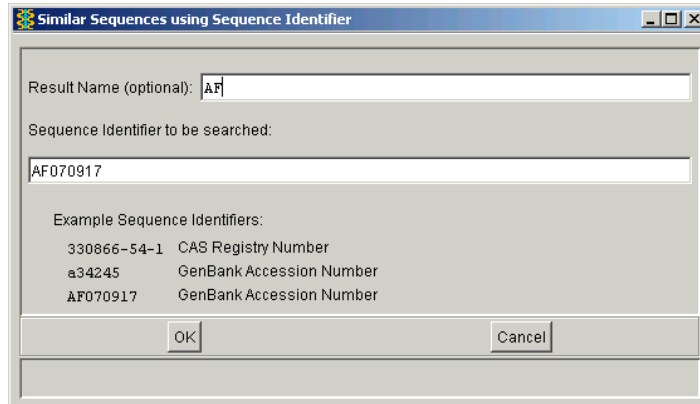
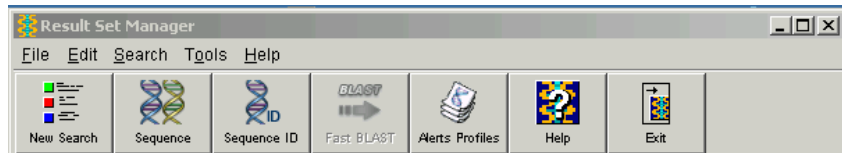


## Searching by sequence identifier

- 1 Click the **Sequences ID** button on the toolbar.



- 2 Enter the sequence identifier to be searched and a name for the results.



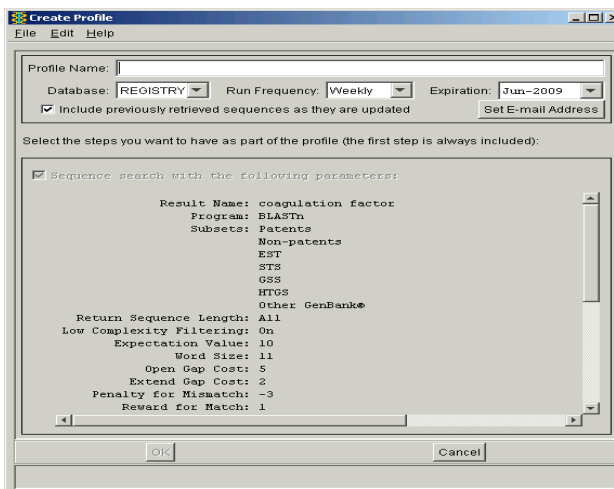
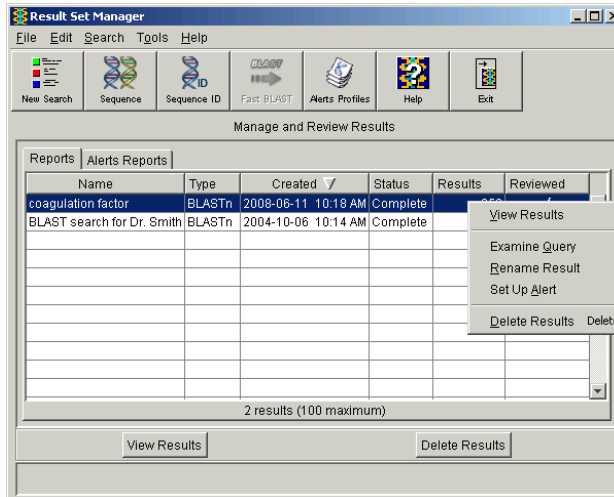
Follow the same steps as for searching similar sequences:

1. Select a BLAST program in the Program Selection window.
2. Select the database or a database subset.
3. Specify the BLAST settings.
4. Run the search.

## Creating alerts

After you conduct a CAS REGISTRY BLAST search, follow these steps to create a CAS REGISTRY BLAST alert:

- 1 In the Result Set Manager window, right-click on the set of results of interest and select **Set Up Alert**.



Define the options for the alert in the Create Profile window. Click **OK**. Click **Set E-mail Address** to specify an e-mail address for sending the alert results and for notifying about expiration dates.

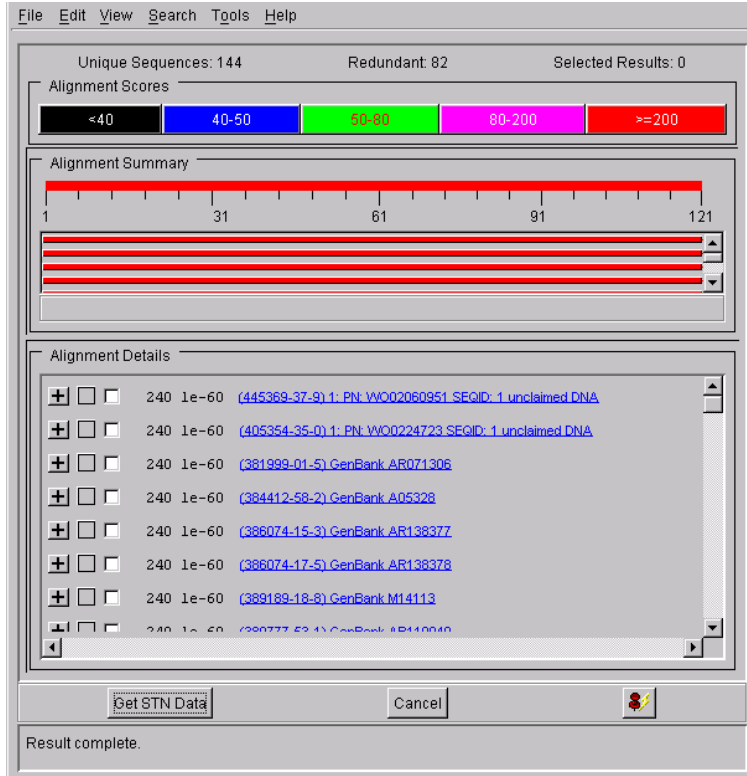
CAS REGISTRY BLAST alerts are set to expire 1 year from the date when they are created. If you provide an e-mail address, you will be notified via e-mail that your alert will soon expire.

## Getting STN data for BLAST results

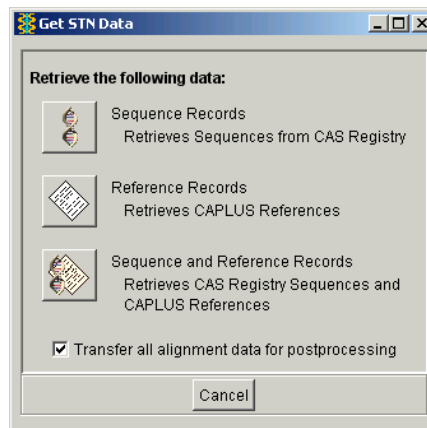
1 Select the sequences for which you want to get STN data.

- Choose **Edit > Select All** to select all sequences.
- In the **Alignment Scores** section, click the bar for the desired range to select the entire range.
- Click the selection boxes for individual sequences or their CAS Registry Numbers® to select individual sequences.

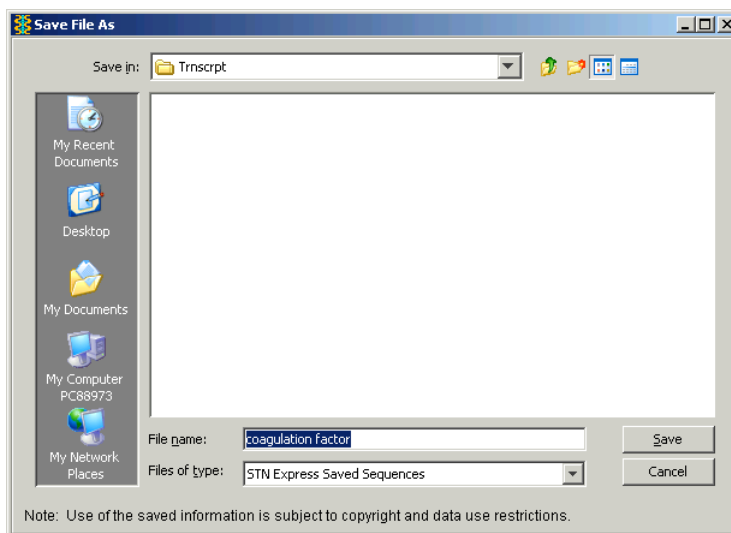
2 Click **Get STN Data**.



3 Click the desired option. The **Transfer all alignment data for postprocessing** box is checked by default so that your BLAST alignment data is saved and can be combined with your STN transcript, if you choose to create a report.



4 Enter a file name in which to save the BLAST alignment sequence data. The data are saved in an STN Express Saved Sequences (.xss) file.



The standard STN logon procedure is initiated. Once an online session is initiated, STN Express automatically enters appropriate databases and conducts searches to obtain the results that you requested in the Get STN Data window.

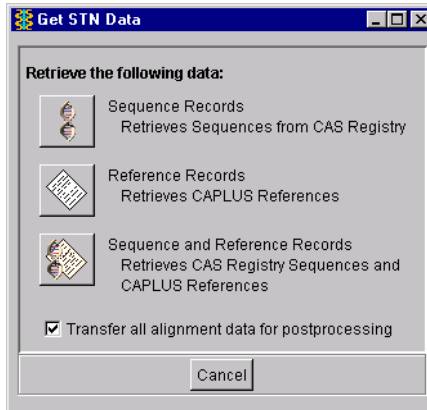
**Note:** You can create a report from the BLAST alignment data and records from a transcript. For the data from an STN record to be included in the transcript, it must first be displayed in STN. Be sure to display any data that you intend to include in a report.

- Use a display format in REGISTRY that includes CAS Registry Numbers<sup>®</sup> if you want the BLAST alignment data included within the REGISTRY records in your transcript.
- Conduct a search on the REGISTRY L-number from BLAST results in CPlus and display records in the desired format in order to include bibliographic data from the transcript in your report. For example, if you want to include bibliographic information and abstracts for CPlus records, display answers in the BIB ABS format.

## Creating a report with BLAST alignment data

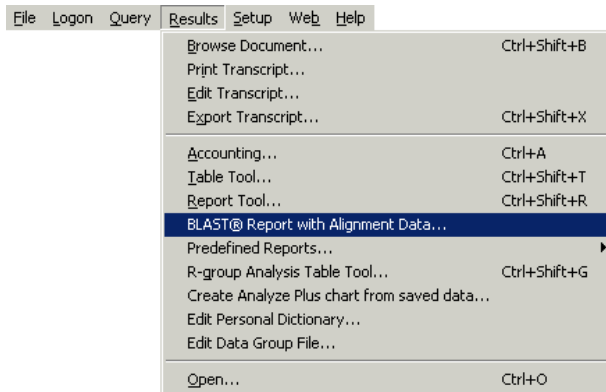
### 1 Select the data to retrieve.

To create a report that merges transcript results from your STN online session with BLAST alignment data, make sure that the box **Transfer all alignment data for postprocessing** is checked in the **Get STN Data** window of the BLAST search.

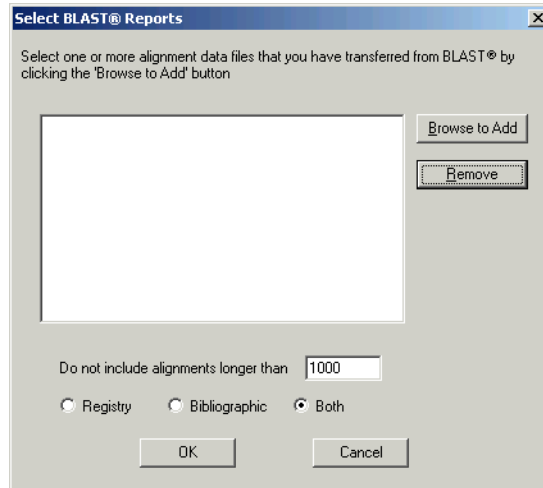


### 2 Log off STN.

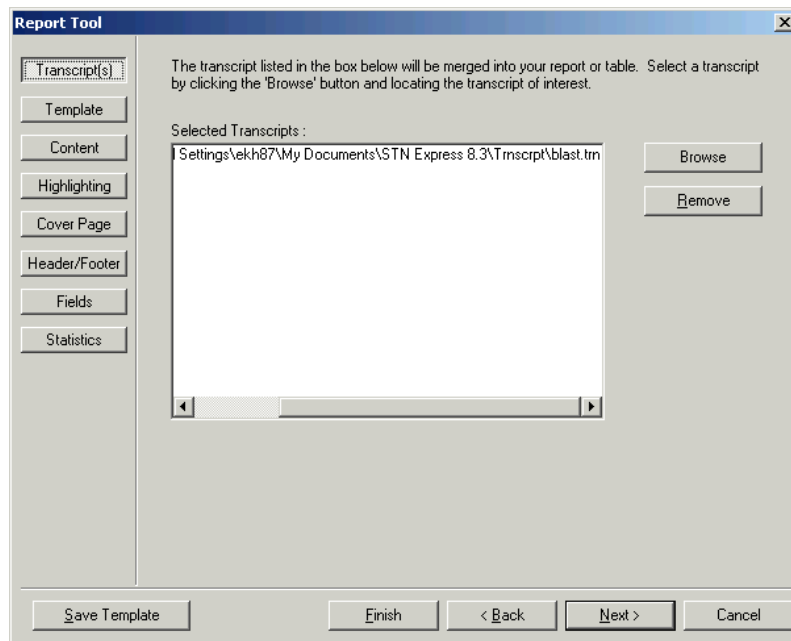
### 3 Select **Results > BLAST Report with Alignment Data**.



4 Select one or more alignment data files in the Select BLAST Reports window. Use **Browse to Add** to locate and add files. Note that the default is set to exclude alignments longer than 1000 residues. You may also choose to include only the REGISTRY data, only the bibliographic data, or both. Click **OK**.



5 The first screen of the Report Tool window shows the Transcript portion of the window with the last transcript that was created. Use the displayed transcript or click **Browse** to locate the transcript you want to use.



Select from the buttons on the left side of the Report Tool window.

<b>Use this button...</b>	<b>To...</b>
Transcript(s)	Select one or more transcripts for the report
Template	Select a template for the report
Content	Define the content of the report
Highlighting	Define highlighting in the report
Cover Page	Define the content and format for the optional cover page
Header/Footer	Define the content and format for the optional header or footer
Fields	Select and format the fields from answers
Statistics	Define charts for the report based on data from fields in the transcript(s)

### **For more information**

Refer to STN Express Training at [www.cas.org](http://www.cas.org).