

SRS USER GUIDE

Version 7.1

Copyright (c) 2003 LION bioscience AG (LION). All rights reserved.

LION Bioscience SRS 7.1 Documentation

This manual, as well as the software described in it, is furnished under license and may only be used or copied in accordance with the terms of such license. The information in this manual is furnished for information only, is subject to change without notice, and should not be construed as a commitment by LION bioscience AG. LION bioscience AG assumes no responsibility or liability for any errors or inaccuracies that may appear in this book.

Except as permitted by such license, no part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, recording, or otherwise, without the prior permission of LION bioscience AG.

Comments about the documentation are welcome at: documentation@uk.lionbioscience.com

For customer support issues please e-mail support@uk.lionbioscience.com

TABLE OF CONTENTS

CHAPTER 1: SRS QUICK TOUR	1
1.1 Introduction	2
1.2 Starting an SRS Project	2
1.3 SRS Queries	4
Making a Query using SRS Quick Search	4
Making a Query Using the Standard Query Form	7
1.4 Sorting Results	10
1.5 Links	11
Linking to Related Information	11
1.6 Using Views	13
Creating a View	14
Applying a View	17
Deleting a View	20
1.7 Using Analysis Tools	20
Running a Batch Queued Tool	24
Running an Interactive Tool	27
Viewing Tool Results	30
1.8 Saving your Results	31
1.9 Review	35
CHAPTER 2: SRS PROJECTS	37
2.1 Introducing Projects	38
Temporary Projects	38
Permanent Projects	38
Starting a Project	39

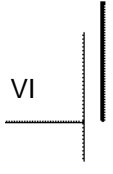
2.2	Temporary Projects	41
	Starting a Temporary Project	41
	Managing Temporary Projects	42
2.3	Permanent Projects	43
	Starting a Permanent Project	44
	Managing Permanent Projects	46
2.4	Using the Project Manager	47
	Creating a New Project	48
	Switching to Another Project	48
	Copying Project Information	49
	Saving a Project	50
	Opening a Saved Project	52
	Deleting a Project	52
	Renaming a Project	53
CHAPTER 3:	QUERYING WITH SRS	55
3.1	Search Terms	56
	Single Word Searches	56
	Multiple-Word Searches	57
	Numbers and Dates	57
	Regular Expressions	59
	Wildcards	59
3.2	Quick Searches	60
	SRS Quick Search Page	60
	About Quick Search	60
	Using Quick Search	61

3.3 Query Forms	62
About Query Forms.	62
Using the Standard Query Form	62
Using the Extended Query Form.	65
3.4 Subentries	68
About Subentries.	68
Use of Subentries	68
Querying Using Subentries	69
3.5 Expression Queries	79
Searches using a Query Expression	79
Using Expression Queries	80
3.6 Sorting Results.	82
What is Sorting?	82
Sorting a Set of Results	82
3.7 Browse Index.	86
About Browsing Indices	86
Browsing Indices.	86
Getting to the Field Information Page.	88
CHAPTER 4: LINKS TO ADDITIONAL DATA.	91
4.1 What is a Link?	92
4.2 Hypertext Links	92
4.3 Index Links	95
Searching for Data using Index Links.	95
Link Options	98
4.4 Getting to the Link Page	99

	Finding Links from the Manage your Query Results Page	99
	Finding Links from the Query Result Page	101
	Finding Links from the Entry Page	103
4.5	Expression Linking	104
	Expression Linking Procedure	104
	Linking Operators	105
	Expression Linking Examples	107
CHAPTER 5:	VIEWS	109
5.1	What is a View?	110
5.2	Using Views	110
	How to Apply Views	111
5.3	Creating Views	123
	Creating Views from the Query Forms	124
	Creating Views using the View Manager Pages	129
CHAPTER 6:	ANALYSIS TOOLS	137
CHAPTER 7:	COMMAND LINE SRS	149
7.1	getz	150
	getz Usage	151
CHAPTER 8:	SRS QUERY LANGUAGE	155
8.1	Introduction	156
8.2	Searching in Indices	157
	Introduction	157
	General Syntax	157
	Search Strings	158

Searching Using Numerical Ranges	161
Searching for Dates	163
Searching Multiple Databanks	164
8.3 Combining Searches	165
Introduction	165
General Syntax	165
Operands	167
Operators	168
Use of Operators to Combine Search Items	171
8.4 Entries and Subentries	173
Links with Sets Containing Subentries	174
8.5 Storing Intermediate Results in Sets	176
INDEX	145

VI



1

SRS QUICK TOUR

SRS is the world's premier data integration, analysis and display tool for bioinformatic, genomic and related data.

This chapter contains a guided tour of SRS which will help you become familiar with the SRS web interface. During the guided tour, you will learn:



- How to start an SRS project.
- How to perform a query using SRS.
- How to link your query results to other databanks.
- How to change the way in which your results are displayed by creating your own view.
- How to run an application.
- How to save your working project.

1.1 Introduction

This chapter is intended to introduce new users to SRS. It will take you through some basic procedures that you can expect to do with SRS. It is intended as an introduction and will get you started, but it is not intended as a complete description of the software.

If you want more detail of how to do a certain task refer to the relevant chapter of this guide. The notes in the margins tell you which chapter is relevant to each set of steps.

If you want to know more about a particular SRS page, then you should refer to the relevant page in the SRS Online Help.

You can get to the Online Help by clicking  on your SRS page at any time. This will take you to the information relevant for the page you were using when you pressed .

1.2 Starting an SRS Project

See “SRS Projects” - Chapter 2 for more information.

1. When you access SRS the **Start** page is usually shown first. From here you can start a permanent project, do a quick search (if you have the relevant databanks installed on your installation) or access the online help files.

Note: Your SRS Administrator should be able to give you the web address of your local server, or of an external site to which you have access.

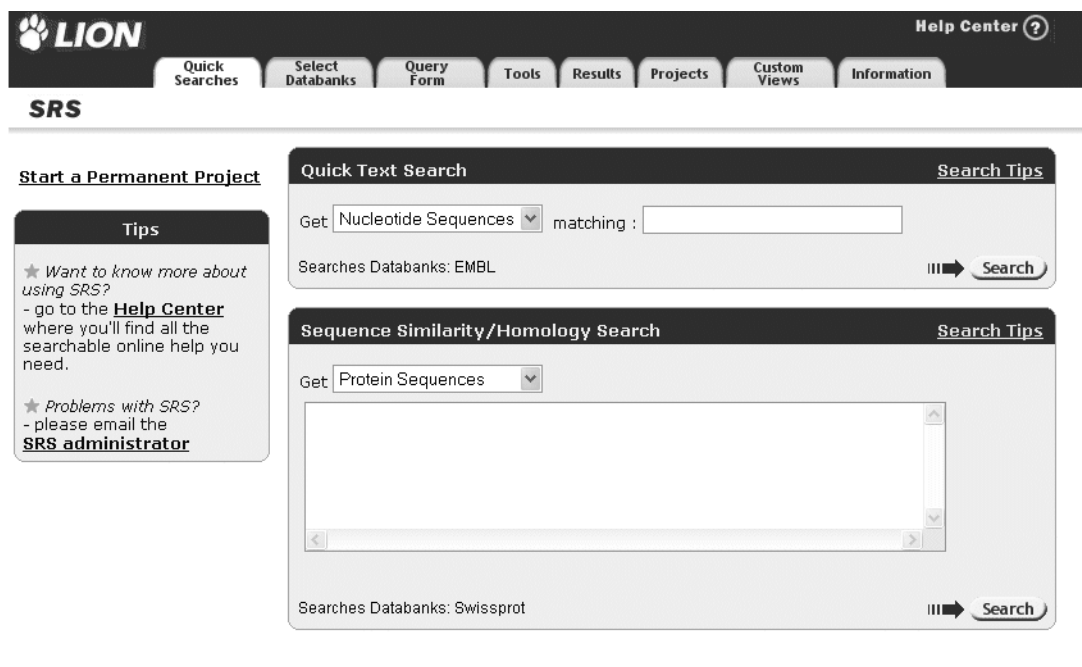


Figure 1.1 The **Start** page.

Note: This page is almost identical to the **SRS Quick Search** page, that will be introduced shortly. The main difference is that where the **Start** page has a link to allow you to start a permanent project, the **SRS Quick Search** page displays the current project ID.

2. For this tour a temporary project will be used. Such a project will be started for you as soon as you start to make a query or select a databank; you do not need to start a temporary project explicitly.

1.3 SRS Queries

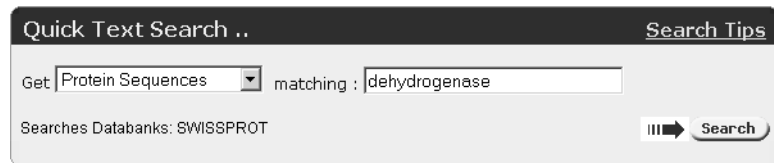
1.3.1 Making a Query using SRS Quick Search

Quick Searches allow users to make a number of searches without needing to learn how to use SRS in depth. The searches query some of the common databanks without having to go and select them explicitly and without the need to understand the **SRS Query Forms**.

Quick Searches can be performed from either the **Start** page (when you first open SRS) or the **SRS Quick Search** page (when you are already in a project).

Quick Text Search

1. You should already be on the **Start** page, if you are not, for any reason, click on the **Quick Searches** tab. This will take you to the **SRS Quick Search** page and you can perform the same queries from there.
2. Select **Protein Sequences** from the **Get** drop-down list. This tells SRS to search the SWISS-PROT databank, and this is indicated below the drop-down list.
3. Type **dehydrogenase** in the **matching:** text box to tell SRS what you want to find.



Quick Text Search .. [Search Tips](#)

Get matching :

Searches Databanks: SWISSPROT

Figure 1.2 Quick Text Search box showing a search for **Protein Sequences** (SWISS-PROT entries) that match **dehydrogenase**.

4. Click the button. This runs the query and then displays the **Query Result** page, showing your results.

SRS

Reset Query "[SWISSPROT-alltext:dehydrogenase*]" found 3987 entries next

Apply Options to:

selected results only

unselected results only

Result Options

Launch analysis tool:

BlastP

Show tools relevant to these results:

Link to related information:

Save results:

Display Options

View results using:

SeqSimpleView

Sort results by:

unsorted

ascending

descending

Show 30 results per page

Printer friendly view

SWISSPROT	Accession	Description	SeqLength
<input type="checkbox"/> SWISSPROT:12AH_CLOS4	P21215	12-alpha-hydroxysteroid dehydrogenase (EC 1.1.1.176) (Fragment).	29
<input type="checkbox"/> SWISSPROT:25KD_SARPE	P23170	Development-specific 25 kDa protein.	258
<input type="checkbox"/> SWISSPROT:2BHD_STREX	P19992	20-beta-hydroxysteroid dehydrogenase (EC 1.1.1.53).	255
<input type="checkbox"/> SWISSPROT:3BH1_MESAU	Q60555	3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type I (3Beta-HSD I) [Includes: 3-beta-hydroxy-delta(5)-steroid dehydrogenase (EC 1.1.1.145) (3-beta-hydroxy-5-ene steroid dehydrogenase) (Progesterone reductase); Steroid delta-isomerase (EC 5.3.3.1) (Delta- 5-3-ketosteroid isomerase)].	372
<input type="checkbox"/> SWISSPROT:3BH1_MOUSE	P24815	3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type I (3Beta-HSD I) [Includes: 3-beta-hydroxy-delta(5)-steroid dehydrogenase (EC 1.1.1.145) (3-beta-hydroxy-5-ene steroid dehydrogenase) (Progesterone reductase); Steroid delta-isomerase (EC 5.3.3.1) (Delta- 5-3-ketosteroid isomerase)].	372
<input type="checkbox"/> SWISSPROT:3BH1_RAT	P22071	3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type I (3Beta-HSD I) [Includes: 3-beta-hydroxy-delta(5)-steroid dehydrogenase (EC 1.1.1.145) (3-beta-hydroxy-5-ene steroid dehydrogenase) (Progesterone reductase); Steroid delta-isomerase (EC 5.3.3.1) (Delta- 5-3-ketosteroid isomerase)].	372
<input type="checkbox"/> SWISSPROT:3BH2_MESAU	Q64421	3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type II (3Beta-HSD II) [Includes: 3-beta-hydroxy-delta(5)-steroid dehydrogenase (EC 1.1.1.145) (3-beta-hydroxy-5-ene steroid dehydrogenase) (Progesterone reductase); Steroid delta-isomerase (EC 5.3.3.1) (Delta-5-3-ketosteroid isomerase)].	372
<input type="checkbox"/> SWISSPROT:3BH2_MOUSE	P26149	3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type II (3Beta-HSD II) [Includes: 3-beta-hydroxy-delta(5)-steroid dehydrogenase (EC 1.1.1.145) (3-beta-hydroxy-5-ene steroid dehydrogenase) (Progesterone reductase); Steroid delta-isomerase (EC 5.3.3.1) (Delta-5-3-ketosteroid isomerase)] (Fragment).	265
<input type="checkbox"/> SWISSPROT:3BH2_RAT	P22072	3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type II (3Beta-HSD II) [Includes: 3-beta-hydroxy-delta(5)-steroid dehydrogenase (EC 1.1.1.145) (3-beta-hydroxy-5-ene steroid dehydrogenase) (Progesterone reductase); Steroid delta-isomerase (EC 5.3.3.1) (Delta-5-3-ketosteroid isomerase)].	372

Figure 1.3 The results of your query, shown on a **Query Result** page.

Note: If you want to look at a complete entry, click on its hyperlink.

You can sort your results to help you find the ones you are most interested in. See “Sorting Results” (p. 10) for more information on how to do this.

The next section will take you through making a query using the **Standard Query Form**.

1.3.2 Making a Query Using the Standard Query Form

See “Querying with SRS” - Chapter 3 for more information.

This example will take you through making a query, which uses the SRS **Standard Query Form** to search for the word **kinase** in the Description field of the SWISS-PROT databank.

Before you start your query you must choose the databank(s) in which you wish to search. To do this you have to use the **Select Databanks to Search** page.

1. Click the **Select Databanks** tab.
2. On the **Select Databanks to Search** page, select the SWISS-PROT databank by ticking the check box to the left of the **SWISSPROT** hyperlink.

LION Help Center ?

Quick Searches | Select Databanks | Query Form | Tools | Results | Projects | Custom Views | Information

SRS

Reset Quick Search

Search Options

1. Select the **databanks** you want to search
2. Enter your **search terms** in the **Quick Search** box, or choose a **query form** from below

[Standard Query Form](#)

[Extended Query Form](#)

You can **browse** through all the **entries** in any **databanks**. First, **select** the **databanks** you want to browse, then click:

[Browse Entries](#)

Available Databanks

Expand all Collapse all Show databanks tooltips:

EMBOSS Results

Sequence databanks - complete

all EMBL SWISSPROT

Sequence databanks - subsections

all EMBL (Release) EMBL (Updates) SWISSPROT (Release)

SWISSPROT (Updates)

Tool Results

all BLASTP BLASTN BLASTX TBLASTN

TBLASTX CLUSTALW NCLUSTALW PROSITESEARCH

RESTRICTIONMAP HMMBUILD

Tips

► bookmark this [link](#) to return to your project

SRS Release 7.1 Copyright © 1997-2003 LION bioscience AG. All Rights Reserved. [Terms of Use](#) [Feedback](#)

Figure 1.4 The **Select Databanks to Search** page with the SWISS-PROT databank selected.

Note: This page can be reached, at any time from within SRS, by clicking on the **Select Databanks** tab.

3. Click the [Standard Query Form](#) button to display the **Standard Query Form**. From here you can search databanks in many different ways.

LION Help Center

Quick Searches Select Databanks Query Form Tools Results Projects Custom Views Information

SRS

Reset search SWISSPROT

Search Options

Combine search terms with: & (AND)

Use wildcards

Get results of type: Entry

Result Display Options

View results using: SeqSimpleView

or

Create a view

Show 30 results per page

Fields you can search Your search terms

In a single field, you can separate multiple values by &, |, !

AllText

AllText

AllText

AllText

Create a view

Select the fields you want displayed in your view and choose the format

Choose 1 or more fields: ID, Accession Number, Primary Accession Number, Description, Gene Name, Keywords, Entry Creation Date

Display As: Table List

Sequence Format: swiss

Search

Tips

To do more advanced

Figure 1.5 Standard Query Form for a query on the SWISS-PROT databank.

4. In the first field box, select the **Description** field using the drop-down list.
5. Type **kinase** in the text box beside it.

Fields you can search Your search terms

In a single field, you can separate multiple values by &, |, !

Description kinase

AllText


AllText

AllText

Search

Figure 1.6 Search the Description field for kinase.

Note: If you want to know more about the chosen field, click on the **i** icon beside it. This will display the **Field Information** page (see also section “Information Pages - Field Information” in the *Online Help*). Use the **Back** button on your browser to return to the **Standard Query Form**.

6. Click one of the  buttons. This runs the query and then displays the **Query Result** page, showing your results.

1.4 Sorting Results

See section “Sorting Results” (p. 82) for more information.

Results are listed, by default, in the order in which they are stored in the databank (**unsorted**). To make it easier for you to find the results you want, you can use sorting to reorganize them. Sorting is particularly useful if you have a large number of results.

1. To sort a set of results, choose the required sort order from the **Sort results by:** drop-down list in the **Display Options** box, e.g. **Organism Name**, **Sequence Length**.
2. Use the option buttons to choose whether to sort in **ascending** or **descending** order.

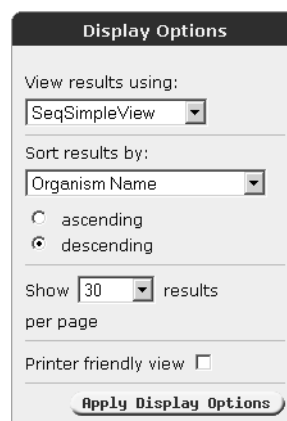


Figure 1.7 Sort by **Organism Name** in an **descending** order.

3. Press the **Apply Display Options** button to sort the results.

1.5 Links

See “Links to Additional Data” - Chapter 4 for more information.

One of the many features of SRS is its ability to search for links between your current results and related information in other databanks. In this example you will search for links from one of your SWISS-PROT results to related entries in EMBL.

1.5.1 Linking to Related Information

1. On the **Query Results** page, tick the check box next to the entry for which you want to find related items.
2. Click the **Link** button in the **Result Options** box to display the **LINK** page.

LION Help Center ?

Quick Searches Select Databanks Query Form Tools Results Projects Custom Views Information

SRS

Reset Find entries related to current query: "[SWISSPROT-ID:1433_OENHO]" in other databanks

Link Options

Select the databanks you want to search for related information

Find related entries

Refine Query - show only results with related entries

Show only results without related entries

Search

Display Options

Show 30 results per page

Databanks Available to Link to

Expand all Collapse all

To Parent Entry

EMBOSS Results

Sequence databanks - complete

all EMBL SWISSPROT

Sequence databanks - subsections

Tool Results

SRS Release 7.1 Copyright © 1997-2003 LION bioscience AG. All Rights Reserved. Terms of Use Feedback

Figure 1.8 LINK page.

3. Tick the check box to the left of the databank in which you wish to find links, e.g. **EMBL**.
4. Click the **Search** button to search for the related results.

The result will be a list of all the EMBL entries that are related to the SWISS-PROT entry (or entries) with which you started. These will be displayed on the **Query Result** page.

LION Help Center ?

Quick Searches Select Databanks Query Form Tools Results Projects Custom Views Information

SRS

Reset Query "[SWISSPROT-ID:ISPE_ZYMMO] > EMBL)" found 6 entries

Apply Options to:

selected results only

unselected results only

EMBL:ECAPH

EMBL:ECT903

EMBL:ISTN90

EMBL:TT903A2

EMBL:TT903A3

EMBL:PRTSG

Figure 1.9 Query Result page showing the EMBL entries that had links to the SWISS-PROT entry: SWISSPROT:ISPE_ZYMMO (Accession number: Q9X3W5), from the above query for **kinase**.

If there are no items linked to your selection, then go back to the original **Query Result** page, choose a different selection, and try again.

If you want to search all your results for links to EMBL, from a **Query Result** page, click the **unselected results only** option in the **Apply Options to:** area, ensure that the check box beside each of the entries is unticked, and repeat the search.

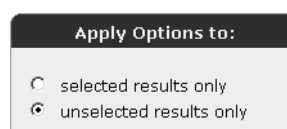


Figure 1.10 Apply Options to: box.

Similarly, by ticking several entries in your list, you can search all those you have selected, or all those that are not selected using the option buttons.

Note: Searching a large number of results for links might take some time when large databanks are involved.

1.6 Using Views

See "Views" - Chapter 5 for more information.

SRS allows you to customize the way in which you display data. This is usually done using the **View Manager** pages, although simple view creation is also possible from both of the

Query Forms (see section 5.3.1 “Creating Views from the Query Forms”, p. 124).

1.6.1 Creating a View

This example explains how to create a table view that can be used for SWISS-PROT, and for which links to EMBL should be shown.

1. Click the **Custom Views** tab to display the **View Manager** page 1.
On this page you can select databanks and choose a name for your customized view. See Figure 1.11.
2. Give your view a name by typing it in the **View name** box. This example uses the name **myTestView**.
3. Use the **Display results as** options to indicate whether you want to create a table or list view.
4. In the **Show fields from:** options, indicate whether you want to be able to select fields from all of those available for your chosen databanks, or only common fields.
5. In the list under **Databanks to define a view for**, click **SWISSPROT** to select it.
6. In the list under **Databanks to be linked to** click **EMBL**.

SRS

Reset

Create View Options

View name:

Display results as

table

list

Databanks to define a view for	Databanks to be linked to
Show fields from: <input type="radio"/> All fields in databanks <input checked="" type="radio"/> Common fields only	
EMBL EMBL_reference EMBL_features EMBL_counter SWISSPROT SWISSPROT_reference SWISSPROT (comment) SWISSPROT (feature) SWISSPROT (counter) ~~~~~ EMBL (Release) EMBLRELEASE_features EMBLRELEASE_reference EMBLRELEASE_counter EMBL (Updates)	EMBL EMBL_reference EMBL_features EMBL_counter SWISSPROT SWISSPROT_reference SWISSPROT (comment) SWISSPROT (feature) SWISSPROT (counter) ~~~~~ EMBL (Release) EMBLRELEASE_features EMBLRELEASE_reference EMBLRELEASE_counter EMBL (Updates)

Create New View

Delete View: *Names only* Delete View

Figure 1.11 Create your own views (Steps 2 - 7).

7. Click the **Create New View** button to display the **View Manager** page 2 from which you can select the fields for your views (see Figure 1.12).
8. Choose the fields you want to view by ticking the check boxes beside them. For the purpose of this example, select the **Description** field for both databanks.

Note: The field names are also hyperlinks to the **Field Information** page (see also section “Information Pages - Field Information” in the *Online Help*), which contains information about each field.

Create View Options

Select the **datafields** you want **displayed in your view** using the checkboxes.

View name: myTestView

Save view:

Datafields for your primary databanks

SWISSPROT

<input type="checkbox"/> ID	<input type="checkbox"/> <u>AccNumber</u>	<input type="checkbox"/> <u>PrimAccNumber</u>	<input checked="" type="checkbox"/> <u>Description</u>
<input type="checkbox"/> <u>GeneName</u>	<input type="checkbox"/> <u>Keywords</u>	<input type="checkbox"/> <u>DateCreated</u>	<input type="checkbox"/> <u>LastSequenceUpdate</u>
<input type="checkbox"/> <u>LastAnnotationUpdate</u>	<input type="checkbox"/> <u>Organism</u>	<input type="checkbox"/> <u>Taxon</u>	<input type="checkbox"/> <u>NCBI_TaxId</u>
<input type="checkbox"/> <u>Organelle</u>	<input type="checkbox"/> <u>ProteinID</u>	<input type="checkbox"/> <u>Checksum</u>	<input type="checkbox"/> <u>DbName</u>
<input type="checkbox"/> <u>DBxref</u>	<input type="checkbox"/> <u>SeqLength</u>	<input type="checkbox"/> <u>Sequence</u>	<input type="text" value="fasta"/>

Fields of subentry **Reference**

<input type="checkbox"/> <u>Authors</u>	<input type="checkbox"/> <u>Title</u>	<input type="checkbox"/> <u>Journal</u>	<input type="checkbox"/> <u>VolumeNo</u>
<input type="checkbox"/> <u>FirstPage</u>	<input type="checkbox"/> <u>Year</u>	<input type="checkbox"/> <u>MedlineID</u>	<input type="checkbox"/> <u>PubMedID</u>
<input type="checkbox"/> <u>RefPosition</u>	<input type="checkbox"/> <u>RefCommentCode</u>	<input type="checkbox"/> <u>RefComment</u>	

Fields of subentry **Comment**

<input type="checkbox"/> <u>CommentType</u>	<input type="checkbox"/> <u>Comment</u>
---	---

Fields of subentry **Feature**

<input type="checkbox"/> <u>FtKey</u>	<input type="checkbox"/> <u>FtLength</u>	<input type="checkbox"/> <u>FtDescription</u>
---------------------------------------	--	---

Fields of subentry **Counter**

<input type="checkbox"/> <u>CountedItem</u>	<input type="checkbox"/> <u>CountedN</u>
---	--

Datafields from databanks to be linked to

EMBL

<input type="checkbox"/> ID	<input type="checkbox"/> <u>Division</u>	<input type="checkbox"/> <u>AccNumber</u>	<input type="checkbox"/> <u>PrimAccNumber</u>
<input type="checkbox"/> <u>SeqVersion</u>	<input type="checkbox"/> <u>Molecule</u>	<input checked="" type="checkbox"/> <u>Description</u>	<input type="checkbox"/> <u>Keywords</u>
<input type="checkbox"/> <u>Organism</u>	<input type="checkbox"/> <u>Taxon</u>	<input type="checkbox"/> <u>Organelle</u>	<input type="checkbox"/> <u>Comment</u>
<input type="checkbox"/> <u>DateCreated</u>	<input type="checkbox"/> <u>LastUpdated</u>	<input type="checkbox"/> <u>SeqLength</u>	<input type="checkbox"/> <u>Link</u>
<input type="checkbox"/> <u>Sequence</u>	<input type="text" value="fasta"/>		

Fields of subentry **References**

<input type="checkbox"/> ID	<input type="checkbox"/> <u>Authors</u>	<input type="checkbox"/> <u>Title</u>	<input type="checkbox"/> <u>Journal</u>
<input type="checkbox"/> <u>VolumeNo</u>	<input type="checkbox"/> <u>FirstPage</u>	<input type="checkbox"/> <u>Year</u>	<input type="checkbox"/> <u>MedlineID</u>
<input type="checkbox"/> <u>PubMedID</u>	<input type="checkbox"/> <u>RefPosition</u>		

Fields of subentry **Features**

<input type="checkbox"/> ID	<input type="checkbox"/> <u>AccNumber</u>	<input type="checkbox"/> <u>FtKey</u>	<input type="checkbox"/> <u>FtQualifier</u>
<input type="checkbox"/> <u>ProteinID</u>	<input type="checkbox"/> <u>Gene</u>	<input type="checkbox"/> <u>FtDescription</u>	<input type="checkbox"/> <u>FtLength</u>
<input type="checkbox"/> <u>Sequence</u>	<input type="text" value="gCg"/>		

Fields of subentry **Counters**


<input type="checkbox"/> ID	<input type="checkbox"/> <u>CountedItem</u>	<input type="checkbox"/> <u>CountedN</u>
-----------------------------	---	--

Use explicit link:

Use view to display entries:

Display only number of linked entries only

Figure 1.12 View Manager page 2.

9. Click the  in the **Create View Options** box to save your view.
Your view will be saved and you will be returned to the **View Manager** page 1, where you can create more views.

Your saved views will be added to the drop-down list in the **Results Display Options:** box, and can be applied to your results in the same way that any other view is applied (see section 1.6.2 “Applying a View”, p. 17).

1.6.2 Applying a View

This section demonstrates how to apply a view to your results. The example uses the view created in section 1.6.1 “Creating a View”, p. 14, but can equally refer to any other view in the drop-down list.

1. Click the **Results** tab to get to the **Manage your Query Results** page.
2. Choose a set of results that you wish to view, e.g. by ticking the check box beside them (on the **Manage your Query Results** page). See Figure 1.13. Tick the check box beside the query labelled **Q1**. This was the first query you made (hence **Q1**).
3. Select **myTestView** from the drop-down list in the **Results Display Options** box.

The screenshot shows the SRS interface with a search bar and a results table. The search bar contains a query expression: `{[SWISSPROT-ID:ISPE_Z]`. The results table shows four query sets: Q4 (link, 6 results), Q3 (select, 1 result), Q2 (query, 3741 results), and Q1 (workflow, 3987 results). The Q1 result set is selected. The 'Results Display Options' panel is open, showing a dropdown menu with options: '* Complete entries *', 'default view', '* Names only *', '* Complete entries *', 'SeqSimpleView', 'FastaSeqs', 'SwissView', 'proteinChart', and 'myTestView'. A mouse cursor is pointing at 'myTestView'.

Results Options

Options are **applied** to **selected** queries only

Save results: [Save](#)

Delete results: [Delete](#)

Combine queries with: [& \(AND\)](#) [Combine](#)

Find related info: [Link](#)

Results Display Options

View results using:

- * Complete entries *
- default view
- * Names only *
- * Complete entries *
- SeqSimpleView
- FastaSeqs
- SwissView
- proteinChart
- myTestView

Search using a query expression: [Search](#)

Result History

Name	Type	Total No.	From	No.	Query Expression	Comment
<input type="checkbox"/> Q4	link	6	EMBL	6	{[SWISSPROT-ID:ISPE_Z]	
<input type="checkbox"/> Q3	select	1	SWISSPROT	1	[SWISSPROT-ID:ISPE_Z]	
<input type="checkbox"/> Q2	query	3741	SWISSPROT	3741	[swissprot-Descripti]	
<input checked="" type="checkbox"/> Q1	workflow	3987	SWISSPROT	3987	[SWISSPROT-alltext:d]	

Figure 1.13 View a set of results using your own view.

- Click the [Rerun Query](#) button to display the query result set, **Q1**, using **myTestView**. The entries in **Q1** will be displayed using the view you selected.

LION Help Center ?

Quick Searches | Select Databanks | Query Form | Tools | Results | Projects | Custom Views | Information

SRS

Reset Query "[SWISSPROT-alltext:dehydrogenase*]" found 3987 entries next

Apply Options to:

selected results only

unselected results only

Result Options

Launch analysis tool:

BlastP Launch

Show tools relevant to these results: Tools

Link to related information: Link

Save results: Save

Display Options

View results using:

myTestView

Sort results by:

unsorted

ascending

descending

Show 30 results per page

Printer friendly view

Apply Display Options

SWISSPROT	Description	EMBL	Description
<input type="checkbox"/> SWISSPROT:12AH_CLOS4	12-alpha-hydroxysteroid dehydrogenase (EC 1.1.1.176) (Fragment).		
<input type="checkbox"/> SWISSPROT:25KD_SARPE	Development-specific 25 kDa protein.	EMBL:SPDEVG	
<input type="checkbox"/> SWISSPROT:2BHD_STREX	20-beta-hydroxysteroid dehydrogenase (EC 1.1.1.53).		
<input type="checkbox"/> SWISSPROT:3BH1_MESAU	3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type I (3Beta-HSD I) [Includes: 3-beta-hydroxy-delta(5)-steroid dehydrogenase (EC 1.1.1.145) (3-beta-hydroxy-5-ene steroid dehydrogenase) (Progesterone reductase); Steroid delta-isomerase (EC 5.3.3.1) (Delta-5-3-ketosteroid isomerase)].	EMBL:MAORF	
<input type="checkbox"/> SWISSPROT:3BH1_MOUSE	3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type I (3Beta-HSD I) [Includes: 3-beta-hydroxy-delta(5)-steroid dehydrogenase (EC 1.1.1.145) (3-beta-hydroxy-5-ene steroid dehydrogenase) (Progesterone reductase); Steroid delta-isomerase (EC 5.3.3.1) (Delta-5-3-ketosteroid isomerase)].	EMBL:MMHSD3B	
<input type="checkbox"/> SWISSPROT:3BH1_RAT	3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type I (3Beta-HSD I) [Includes: 3-beta-hydroxy-delta(5)-steroid dehydrogenase (EC 1.1.1.145) (3-beta-hydroxy-5-ene steroid dehydrogenase) (Progesterone reductase); Steroid delta-isomerase (EC 5.3.3.1) (Delta-5-3-ketosteroid isomerase)].	EMBL:RN3BHSDA	
<input type="checkbox"/> SWISSPROT:3BH2_MESAU	3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type II (3Beta-HSD II) [Includes: 3-beta-hydroxy-delta(5)-steroid dehydrogenase (EC 1.1.1.145) (3-beta-hydroxy-5-ene steroid dehydrogenase) (Progesterone reductase); Steroid delta-isomerase (EC 5.3.3.1) (Delta-5-3-ketosteroid isomerase)].	EMBL:MAHSD3B	
<input type="checkbox"/> SWISSPROT:3BH2_MOUSE	3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type II (3Beta-HSD II) [Includes: 3-beta-hydroxy-delta(5)-steroid dehydrogenase (EC 1.1.1.145) (3-beta-hydroxy-5-ene steroid dehydrogenase) (Progesterone reductase); Steroid delta-isomerase (EC 5.3.3.1) (Delta-5-3-ketosteroid isomerase)] (Fragment).	EMBL:MM3BHIDRX	
<input type="checkbox"/> SWISSPROT:3BH2_RAT	3 beta-hydroxysteroid dehydrogenase/delta 5-->4-isomerase type II (3Beta-HSD II) [Includes: 3-beta-hydroxy-delta(5)-steroid dehydrogenase (EC 1.1.1.145) (3-beta-hydroxy-5-ene steroid dehydrogenase) (Progesterone reductase); Steroid delta-isomerase (EC 5.3.3.1) (Delta-5-3-ketosteroid isomerase)].	EMBL:RN3BHSDB EMBL:S63167	

Figure 1.14 Query Result page showing typical SWISS-PROT entries and any linked EMBL entries, shown using myTestView.

If you cannot see any entries with links, try looking for the entry for which you found the original link. If you had to try

linking to a databank other than EMBL, you may want to create a view that will show SWISS-PROT and your chosen databank, rather than SWISS-PROT and EMBL. (Do not worry, at this stage, if you cannot find your result.)

1.6.3 Deleting a View

If you no longer need a view, you can delete it.

1. From the **View Manager** page 1 (use the **Custom Views** tab to get there) select the view you want to delete, e.g. **myTestView**.



Figure 1.15 Selecting a view to delete.

2. Click the **Delete View** button.

1.7 Using Analysis Tools

See “Analysis Tools” - Chapter 6 for more information.

SRS is able to analyze the results of your search using many bioinformatics analysis tools or applications. This enables you to seek out further information that may be relevant to your initial search. This section will take you through running a tool on a set of results.

1. Go to a **Query Result** page.
(If you can't remember how to get to the **Query Result** page, use the **Results** tab to go to the **Manage your Query Results** page, tick the check box beside a set of

- results, set the view to something appropriate (e.g. **default view**), and click the **Rerun Query** button to take you to the **Query Result** page for that set of results.)
2. Select an entry from the list of the results, by ticking the check box beside it. It will be the information in this entry that is then used to run the tool.
 3. Click the **Tools** button. This will take you to the **Tool Select** page which shows all the applications that can be used on the current entry. You may need to open up the lists of tools by clicking on the **+** next to a group.

Query "[SWISSPROT-ID:12AH_CLOS4]" found 1 entries

Quick Launch

Launch analysis tool:

Packages Information

[BLAST](#)
[EMBOSS](#)

Available Analysis Tools - listed by type

Expand all Collapse all

- Alignment Tools
- Display Tools
- Edit Tools
- Information Tools
- Nucleic Tools
- Protein Tools
- Similarity Search Tools

BlastP BLASTP Protein-Protein Sequence Similarity Search - [Launch More Info...](#)

TBlastN Sequence Similarity Search - [Launch More Info...](#)

Figure 1.16 Tool Select page.

Note: You can also access many of the most commonly used tools using the drop-down list in the **Result Options** box. Selecting the tool from the drop-down list and pressing the button takes you directly to the **Launch** page, see step 4.

4. If you have selected a query for which **BlastN** is available, then click on the **Launch** hyperlink for **BlastN**.

Similarity Search Tools

- BlastP** BLASTP Protein-Protein Sequence Similarity Search - [Launch More Info...](#)
- BlastX** Sequence Similarity Search - [Launch More Info...](#)
- BlastN** Sequence Similarity Search - [Launch More Info...](#)
- TBlastX** Sequence Similarity Search - [Launch More Info...](#)
- TBlastN** Sequence Similarity Search - [Launch More Info...](#)
- HmmBuild** Search a sequence for matches to profiles - [Launch More Info...](#)

Figure 1.17 Selecting **BlastN**.

5. This will take you to the **Launch** page.
 If **BlastN** is not available, then try again with another

query, or select one of the other tools that is available, noting that the images below refer to BLASTN, and will be slightly different for your case.

SRS

Reset

BlastN [More Info...](#)

Job name: temp Database to search: EMBL (Updates) Launch

or enter file name Browse...

Strand of query sequence to use: Both

Job Options

Note: this tool is run by LSF batch queueing system.
Select a queue to submit jobs to: blast-m sisley(batch)

Parameter set options

Save current parameter set as:

Output Options

Number of hits and alignments to show: 50
Number of best hits from a region to keep: 100

Search Parameters

Filter query sequence
Penalty for a nucleotide mismatch: -3
Reward for a nucleotide match: 1
The E value: 10.000000
word size: Default
Perform gapped alignment
Cost to open a gap: Default
Cost to extend a gap: Default
Use lowercase filtering of query sequence

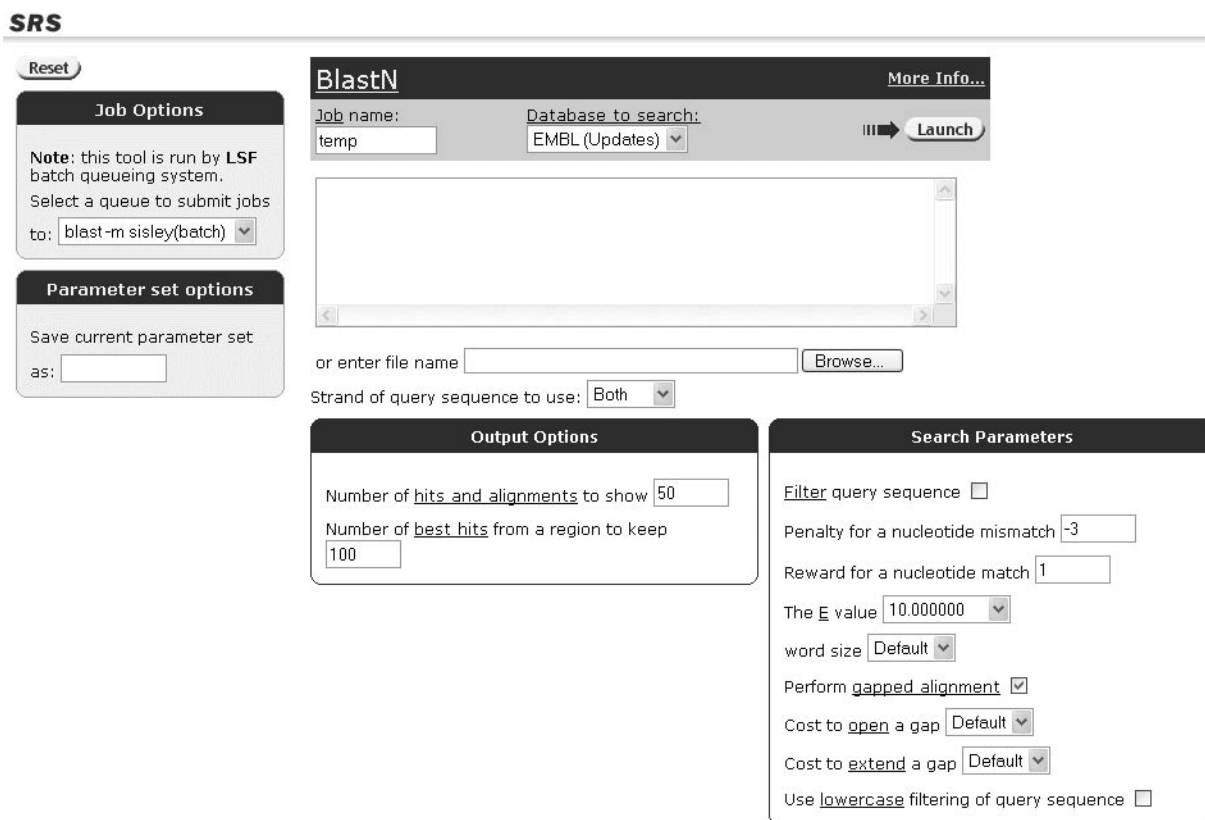


Figure 1.18 Default settings for the **Launch** page.


Note: If you want to know more about any of the tools, click on the **More Info** link beside any of the tools listed on the **Tool Select** page.


- When you run a tool, you can set values for the run-time options on the **Launch** page. For this example, use the **Default parameters**. More information about tool options is available in chapter 6, “Analysis Tools”, (p. 137).


In addition, for tools that will be sent to a batch queue, there is a drop-down list that allows you to choose which batch queue, where there is more than one batch queue available to you. This is not shown in Figure 1.18 because there is only one batch queue available on the installation used.

Note: Depending on how your installation is set up, the tool may be run as a batch job, using a batch queueing system, or it may be run interactively. Figure 1.18 shows that BLASTN will be run as a batch job for this SRS installation, but this may not be the case on your installation. Section “Running a Batch Queued Tool” (p. 24) demonstrates how to run a batch application, while section “Running an Interactive Tool” (p. 27) describes how to run an interactive application.

1.7.1 Running a Batch Queued Tool

1. Click the  button to start the tool running (see Figure 1.19).



LION  Help Center ?

Quick Searches | Select Databanks | Query Form | Tools | Results | Projects | Custom Views | Information


SRS


Tool was submitted to Queue:blast -m sisley(batch).
Tool command:

```
blastall -p blastn -d '/srsdata/flatfiles/blast/emblnew' -i ./temp_blastn_1_in
```

Use *Batch job status page* to view the results

Figure 1.19 Tool Invocation page.

Note: The  icon that appears at the top left of your SRS window, indicates that you have a batch job running. (See section 6.4.3 “Batch Tools”, p. 142 for more information on this icon and others used when running batch applications.)

- Click on the , at the top left corner, to see the **Job Status** page, which will show the status of any batch queue jobs you have run.



LION  Help Center ?

Quick Searches | Select Databanks | Query Form | Tools | Results | Projects | Custom Views | Information

SRS

Job Options

Options are **applied** to **selected** jobs only

Delete jobs:

View job results using:
Complete entries

Run job again with different options:

List of Batch Jobs

Job Name	Status	Start Date	Results from	Result Set	Queue Name
<input type="checkbox"/> temp_blastn_1	<input checked="" type="checkbox"/>	18-Mar-2003 12:40	BLASTN	Q2: 50	priority -m klee(batch)

Figure 1.20 Job Status page.

3. The ✓ icon shows that the application has finished. Click on the hyperlink [temp_blastn](#) to access the results.

The screenshot shows the SRS BLASTN interface. At the top, there's a navigation bar with 'LION' logo, 'Quick Searches', 'Select Databanks', 'Query Form', 'Tools', 'Results', 'Projects', 'Custom Views', and 'Information'. Below this, the 'SRS' logo is visible. The main content area shows a search query: "[BLASTN-JobName:temp_job1]" found 50 entries. There are 'Reset' and 'next' buttons. On the left, there are three panels: 'Apply Options to:' (with radio buttons for 'selected results only' and 'unselected results only'), 'Result Options' (with 'Launch analysis tool:' set to 'HmmBuild', 'Tools', 'Link', and 'Save' buttons), and 'Display Options' (with 'View results using:' set to '*Complete entries*', 'Show 30 results per page', 'Printer friendly view' checkbox, and 'Apply Display Options' button). The main results area shows two entries. The first entry is for 'BLASTN:temp_job1 emblnew 1 AY069346' with a score of 61.9 bits (31), Expect = 4e-08, and 58/67 (86%) identities. The second entry is for 'BLASTN:temp_job1 emblnew 2 AK004606' with a score of 54.0 bits (27), Expect = 1e-05, and 42/47 (89%) identities. Both entries show sequence alignments with gaps represented by vertical bars.

Reset Query "[BLASTN-JobName:temp_job1]" found 50 entries next

BLASTN:temp_job1 emblnew 1 AY069346

>emb|AY069346|AY069346 Drosophila melanogaster LD05574 full length
cdNA.
Length = 2285

Score = 61.9 bits (31), Expect = 4e-08
Identities = 58/67 (86%)
Strand = Plus / Plus

Query: 155 ggcactgccggttatgcctccatcaacacgcacatctcggcgtcgcagcagtcctcgcctgac 214
|||||
Sbjct: 999 ggcactgccggttatgcctccatcaacacgcacatctcggcgtcgcagcagtcctcgcctgac 1058

Query: 215 gacatgg 221
|||||
Sbjct: 1059 gacatgg 1065

Score = 54.0 bits (27), Expect = 1e-05
Identities = 42/47 (89%)
Strand = Plus / Plus

Query: 3 ttcattccaccgcgacatcaagcctgacaacttccttatgggattgg 49
|||||
Sbjct: 846 ttcattccaccgcgacatcaagcctgacaacttccttatgggattgg 892

BLASTN:temp_job1 emblnew 2 AK004606

>emb|AK004606|AK004606 Mus musculus adult male lung cDNA, RIKEN

Figure 1.21 Entry page showing the application results.

1.7.2 Running an Interactive Tool

1. If your tool is not run via a batch queueing system (i.e. it is run interactively) then the **Launch** page should look something like that shown in Figure 1.22.

LION Help Center ?

Quick Searches Select Databanks Query Form Tools Results Projects Custom Views Information

SRS

Reset

Result Display Options

View results using:
Blast_View

Show 30 results per page

Show results automatically:

Parameter set options

Save current parameter set as:

BlastX More Info...

Job name: temp Database to search: SWISSPROT (Updates) Launch

EMBL:MMNKX

begin 1

end 1972

1 11 21 31 41 51
 cgaggggcccaggtacaggggcaggcaggcctgagctctgtagaggaggtactcgggggt
 61 71 81 91 101 111
 gtgtcagagcagcaggtcaccagagcttctgagcaaaaagaggagacctcaagactat
 121 131 141 151 161 171
 ctccccattggccaagagagaaagcttccagggcacaggcaagaggaggagggcaccgag

Strand of query sequence to use: Both Codon Translation table
 Standard Genetic Code

Output Options

Number of hits and alignments to show 50

Number of best hits from a region to keep 100

Search Parameters

Filter query sequence

Scoring matrix BLOSUM62

The E value 10.000000

word size Default

Perform gapped alignment

Cost to open a gap Default

Cost to extend a gap Default

Use lowercase filtering of query sequence

Figure 1.22 Launch page for an interactive tool.

2. Use the default parameters (see step 6, above).
3. Choose a view from the drop-down menu, unless you want to use the default view, (Figure 1.23).

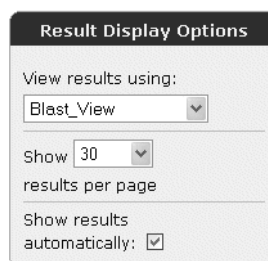
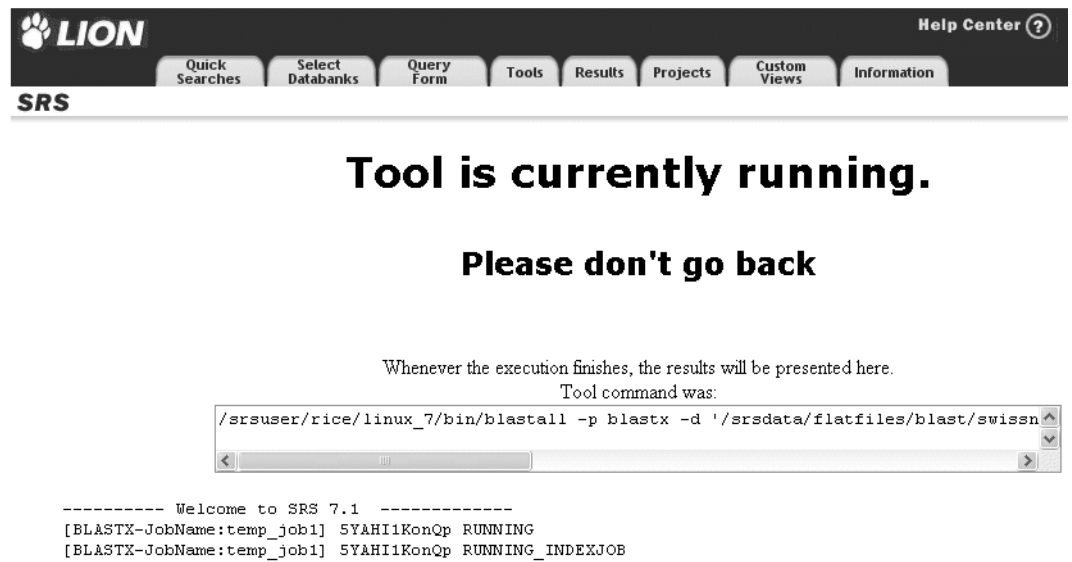


Figure 1.23 Choosing a view with which to display the results of a BLASTX analysis, run interactively.

4. Click the **Launch** button to start the tool running (see Figure 1.24).



The screenshot shows the SRS (Sequence Retrieval System) web interface. At the top, there is a navigation bar with the LION logo on the left and a 'Help Center' link on the right. Below the navigation bar, there are several menu items: 'Quick Searches', 'Select Databanks', 'Query Form', 'Tools', 'Results', 'Projects', 'Custom Views', and 'Information'. The main content area displays a large, bold message: 'Tool is currently running. Please don't go back'. Below this message, there is a text box that says 'Whenever the execution finishes, the results will be presented here.' and 'Tool command was:'. A terminal window is shown with the command: `/srsuser/rice/linux_7/bin/blastall -p blastx -d '/srsdata/flatfiles/blast/swissn`. Below the terminal window, the output of the command is displayed: `----- Welcome to SRS 7.1 -----
[BLASTX-JobName:temp_job1] SYAH11KonQp RUNNING
[BLASTX-JobName:temp_job1] SYAH11KonQp RUNNING_INDEXJOB`

Figure 1.24 Tool Invocation page for an interactive tool.

5. When the tool finishes running, the results will be displayed automatically. See Figure 1.25.

Reset Query "[BLASTX-JobName:temp_job1]" found 50 entries next

Apply Options to:

selected results only
 unselected results only

Result Options

Launch analysis tool:
 BlastP

Show tools relevant to these results:

Link to related information:

Save results:

Display Options

View results using:
 Blast_View

Show 30 results per page
 Printer friendly view

BLASTX	Query	Search Database	Hit
<input type="checkbox"/> BLASTX:temp_job1_swissnew_1_HMX_STRPU	EMBL:MMNKX	swissnew	>sp Q26656 HMX_STRPU
<input type="checkbox"/> BLASTX:temp_job1_swissnew_2_HK32_MOUSE	EMBL:MMNKX	swissnew	>sp P97503 HK32_MOUSE
<input type="checkbox"/> BLASTX:temp_job1_swissnew_3_HK32_HUMAN	EMBL:MMNKX	swissnew	>sp P78367 HK32_HUMAN
<input type="checkbox"/> BLASTX:temp_job1_swissnew_4_HM1D_DROAN	EMBL:MMNKX	swissnew	>sp P22544 HM1D_DROAN
<input type="checkbox"/> BLASTX:temp_job1_swissnew_5_HMX2_COTJA	EMBL:MMNKX	swissnew	>sp P23410 HMX2_COTJA
<input type="checkbox"/> BLASTX:temp_job1_swissnew_6_NK2E_XENLA	EMBL:MMNKX	swissnew	>sp P42583 NK2E_XENLA
<input type="checkbox"/> BLASTX:temp_job1_swissnew_7_HMX2_CHICK	EMBL:MMNKX	swissnew	>sp P28362 HMX2_CHICK
<input type="checkbox"/> BLASTX:temp_job1_swissnew_8_SAX1_MOUSE	EMBL:MMNKX	swissnew	>sp P42580 SAX1_MOUSE
<input type="checkbox"/> BLASTX:temp_job1_swissnew_9_HMBP_DROME	EMBL:MMNKX	swissnew	>sp P22809 HMBP_DROME

Figure 1.25 Query Result page showing the tool results.

1.7.3 Viewing Tool Results

1. SRS provides many ways to view tool results.
2. Using the drop-down list below the button, select **BlastAlignment**. If this option is not available, choose another view type, noting that the figures presented below

may refer to **BlastAlignment** and were generated for the BlastN results created above.

- Click the **View** button. This displays the results of the tool using the **BlastAlignment** view. See Figure 1.26.

BLASTX	Score	Alignment	Description	SeqLength
<input type="checkbox"/> BLASTX:temp_job1_swissnew_1_HMX_STRPU	162	Frame = +2 Query: 1262 ESPEKKPACRKKKTRTVFSRSQVFQLESTFDMKRYLSSSERAGLAASLHLETQVKIWFQ 1441 +SP+KK KKKTRTVFSRSQVFQLESTF++KRYLSSSERAGLAA+LHLETQVKIWFQ Sbjct: 247 QSPQKK---KKKTRTVFSRSQVFQLESTFEVKRYLSSSERAGLAANLHLETQVKIWFQ 302 Query: 1442 NRRNKWKRQLAAELEAANLSHAA-----AQRIVRVPILYHENSAAE 1564 NRRNKWKRQ+AAELE+ANL+HAA AQR+VRVPILYHEN Sbjct: 303 NRRNKWKQMAAELESANLAHAAQIRAQANLAQVSAVHVHAYAQRMVVPVILYHENPTT 362	Homeobox protein (SpHmx) (H6-like) (Fragment).	405
<input type="checkbox"/> BLASTX:temp_job1_swissnew_2_HK32_MOUSE	101	Frame = +2 Query: 1037 DSSPASGTD-----DSPEPLKADPD-HKELDSKSPDEIILEESDSEEGKKEGEAV- 1189 D ASGT R D P L D +E +S E+ S + E ++V Sbjct: 109 DVPGASGTGRARVTLGLDQPGCELHPAKDLEEEAPVRSDSSENSASVSGDHSRPGEDD5VT 168 Query: 1190 PGAAGTTVGATTATPGSEWAKAGAESPEKKPAC---RKKKTRTVFSRSQVFQLESTFDMK 1360 P G A A G + +AG E++PA RKK++R FS +QVF+LE F + + Sbjct: 169 PGGARVPGLRGAAGSGASGGQAGVVEEEPAAPKPRKKRSRAAFSHAQVFELERRFNHQ 228	Homeobox protein NKX-3.2 (Bagpipe homeobox protein homolog 1).	333
<input type="checkbox"/> BLASTX:temp_job1_swissnew_3_HK32_HUMAN	99	Frame = +2 Query: 1106 KELDSKSPDEIILEESDSEEGKKEGEAV-PGAAGTTVGATTATPGSEWAKAGAESPEKKP 1282 +E +S E+ S + E + V P A + + A G AG E++P Sbjct: 140 EEAAGRSDSENSASVSGDRSPRTEDDGVGPRGAHVSAALCSGAGGGGGSPAGVAEEEEEP 199 Query: 1283 AC---RKKKTRTVFSRSQVFQLESTFDMKRYLSSSERAGLAASLHLETQVKIWFQNRN 1453 A RKK++R FS +QVF+LE F + RYLS ERA LAASL LLETQVKIWFQNRN Sbjct: 200 AAPKPRKKRSRAAFSHAQVFELERRFNHQRVLSGPERADLAASLKLLETQVKIWFQNRN 259	Homeobox protein NKX-3.2 (Bagpipe homeobox protein homolog 1).	333
<input type="checkbox"/> BLASTX:temp_job1_swissnew_4_HM1D_DROAN	92	Frame = +2 Query: 1097 PDHKELDSKSPDEIILEESDSEEGKKEGEAVPGAAGTTVGATTATPGSE-----WKAGAE 1264 P ELD S E+ DS+EG G G+ + + G +D K+G+ Sbjct: 262 PAGAELDDSSDYHEENEDCSDSEGGGAGGGGGGSHMDDHSHVCSNGKDDGNSIKSGST 321 Query: 1265 SPEKKPACRKKKTRTVFSRSQVFQLESTFDMKRYLSSSERAGLAASLHLETQVKIWFQ 1444 S + +++K RT F+ Q+ LE +F+ ++YLS ER LA L L++ QVK W+QN Sbjct: 322 SDMSGLSKQKQKARTAFTHQLQTLKSFERQKYLVSQERQELAHKLDLSDCQVKTWYQN 381	Homeobox protein OM (1D).	606

Figure 1.26 Part of a **Query Result** page showing one of the pre-defined SRS views.

1.8 Saving your Results

See section “Temporary Projects” (p. 41) for more information.

- SRS allows you to save queries so you can use them later. Click the **Results** tab to take you to the **Manage your Query Results** page. Here you can choose which queries you want to save.

2. Select a query that you want to save by ticking the check box beside it.

The screenshot shows the SRS web interface. At the top, there is a navigation bar with the LION logo and a Help Center link. Below the navigation bar, there are several tabs: Quick Searches, Select Databanks, Query Form, Tools, Results, Projects, Custom Views, and Information. The main content area is divided into several sections:

- Search Bar:** A search bar with the text "Search using a query expression:" and a "Search" button.
- Results Options Panel:** A panel on the left with the title "Results Options". It contains:
 - A note: "Options are **applied** to **selected** queries only"
 - "Save results:" with a "Save" button (highlighted by a mouse cursor) and a tooltip that says "Select options to save data".
 - "Delete results:" with a "Delete" button.
 - "Combine queries with:" with a dropdown menu set to "& (AND)" and a "Combine" button.
 - "Find related info:" with a "Link" button.
- Results Display Options Panel:** A panel below the first one with the title "Results Display Options". It contains:
 - "View results using:" with a dropdown menu set to "* Complete entries *".
 - "Show 30 results per page" with a dropdown menu.
 - A "Rerun Query" button.
- Result History Table:** A table with the following columns: Name, Type, Total No., From, No., Query Expression, and Comment.

Name	Type	Total No.	From	No.	Query Expression	Comment
		50	BLASTX	50	[BLASTX-JobName:temp]	
<input type="checkbox"/> Q2	workflow	3	EMBL	3	[EMBL-alltext:rankx*]	
<input checked="" type="checkbox"/> Q1	workflow	586050	EMBL	586050	[EMBL-alltext:kinase]	

Figure 1.27 Select a query to save.

3. Click the **Save** button to display the **Save Options** page.
4. Use the **Save Options** page to specify what is saved, the way the output is saved and where it goes (to a text file, or to the screen).

LION Help Center ?

Quick Searches Select Databanks Query Form Tools Results Projects Custom Views Information

SRS

Reset

Saving: Query "[EMBL-alltext:kinase*]" 586050 entries

Output To:

Browser Window (HTML) File (text)

Save As:

ASCII text/table

Save with view: *Complete entries*

Column Separator: \t Record Separator: \n

Generic XML format

Using the loader: seqsimple

Specific XML format

Using the loader: seqsimple

Using XML PrintMetaphors with id: no valid metaphors found

Save

Figure 1.28 Set the save options on the **Save Options** page.

5. Click the **Save** button. Your browser's **File Download** dialog box (Figure 1.29) will be displayed.

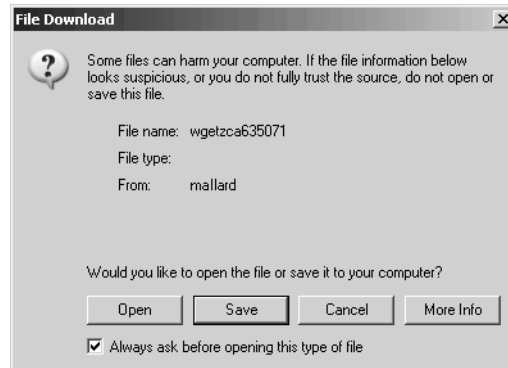


Figure 1.29 The **File Download** dialog box.

Note: Dialog boxes shown throughout this User Guide were generated in the WindowsNT version of Internet Explorer 5. The dialog box you see may look different depending on your system and web browser.

6. Click the **Save** button. This will display the **Save As** dialog box.

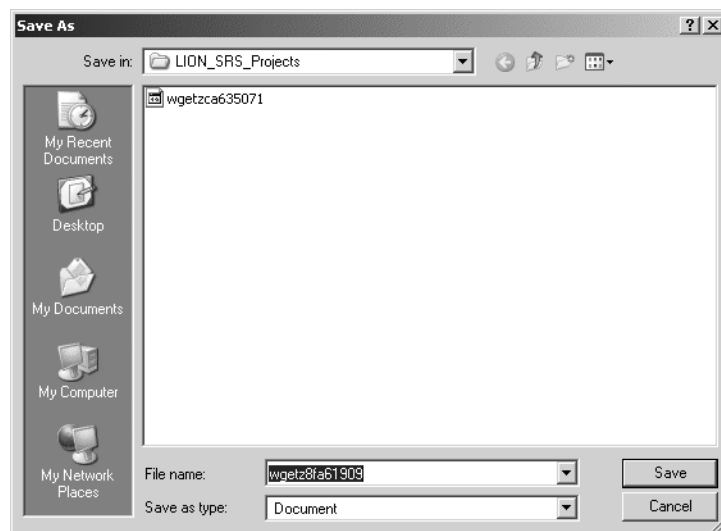
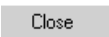


Figure 1.30 The **Save As** dialog box.

7. Select a location for the file and type a suitable name.

Note: You should not use filenames which contain spaces or other special characters, because these can cause problems on some systems. Use an underscore or start new words with capital letters instead.

8. Click  to save the file.

9. Click the  button on the **Download Complete** dialog, if necessary.

1.9 Review

You have now completed a brief overview of the key features of SRS and should be ready to start using it on your own. You should now know how to start an SRS project, be able to perform a query using SRS, link your query results to other databanks, change the way in which your results are displayed by creating your own view, run an application, and save your working project. If you want more information on any particular subject, refer to the relevant chapter.

SRS PROJECTS

All the work you do using SRS will be within projects. These are simply a way of keeping related work together, and any queries, views, etc. that are created in a project will be stored in the project history. This chapter introduces SRS projects. By the end of this chapter you will have learned more about:

- The benefits of using temporary and permanent projects.
- How to start or return to a project.
- Features that are available in permanent projects.
- How to make temporary project data available in other projects.
- How to move temporary project work to a permanent

2.1 Introducing Projects

There are two types of project that can be used in SRS: temporary and permanent. These are described below.

2.1.1 Temporary Projects

When you use a temporary project, your queries and views are stored in a temporary location. They may remain available for a time after you have finished working, but you should not rely on this. If you bookmark the page in your web browser, you should be able to return to the project until the System Administrator clears the files.

You should use a temporary project for:

- Simple searches. For example, a temporary project is useful if you want to look something up quickly, or run an occasional BLAST search.

To find out more about running a temporary project see section 2.2 “Temporary Projects” (p. 41).

2.1.2 Permanent Projects

Permanent projects are used within a permanent SRS user account. There may be one or more (up to 99) projects within any such account. All your queries and views are stored in a project or projects. The fact that the projects are part of a user account, means that they remain available for you to use in

the future, whenever you choose to return to that user account. User accounts (and the permanent projects within them) can also be password protected, allowing you to restrict access.

You should work in a permanent project if any of the following apply:

- You or a colleague will want to return to a project at a later time.
- You want to be able to move your work from one project to another.
- You want to keep a safe record of all your projects.
- You want to recall any of your previous permanent or saved projects.

To find out more about running a permanent project see section 2.3 “Permanent Projects” (p. 43).

2.1.3 Starting a Project

Whether you decide to use a temporary project or a permanent project, you will begin your project from the **Start** page (see Figure 2.1).

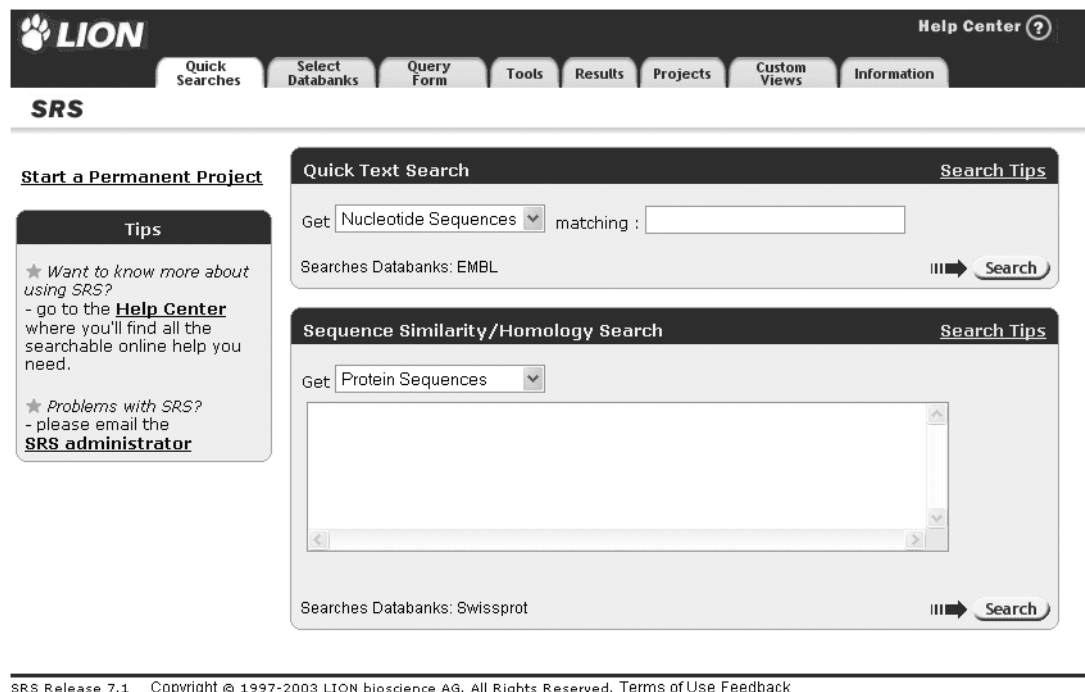


Figure 2.1 The **Start** page.

The **Start** page contains links that allow you to start a permanent project, do a quick search (if you have the relevant databanks installed on your installation) or access the online help files. You can also start a temporary project from here, simply by using the quick search options, or clicking on any of the tabs.

If you choose to start a permanent project, you can return to projects in an existing account or open a new account.

There are also links which allow you to:

- Look up the SRS on-line documentation that is available on your server.
- Contact your SRS Administrator.

The SRS on-line documentation consists of HTML and PDF versions of the available documentation. This includes this User Guide and online page by page SRS help (the SRS Online Guide) as well as documents designed to help SRS Administrators, e.g. the SRS Administrator's Guide, the Icarus Guide, Icarus Quick Reference - Classes & Commands, etc.

2.2 Temporary Projects

Temporary projects should only be used for temporary work. Temporary project data will be removed whenever your System Administrator performs a routine clean-up operation.

2.2.1 Starting a Temporary Project

A temporary project is started from the **Start** page. By default, a temporary project will be started for you as soon as you start to perform tasks within SRS:

1. For example, click on the **Quick Searches** tab. This takes you to the **SRS Quick Search** page. You will see the temporary project ID in the place where the option of starting a permanent project appeared on the **Start** page.



Figure 2.2 Typical temporary project ID, as shown on the **SRS Quick Search** page.

2.2.2 Managing Temporary Projects

If you click the **Projects** tab on the navigation bar, the **Project Manager** page will be displayed (see Figure 2.3).

The screenshot shows the SRS Project Manager interface. At the top, there is a navigation bar with the LION logo and a Help Center icon. Below the navigation bar, there are several tabs: Quick Searches, Select Databanks, Query Form, Tools, Results, **Projects**, Custom Views, and Information. The main content area is titled "SRS Project Options" and "Contents of your temporary project".

SRS Project Options

Save to desktop:

Open from desktop:

Contents of your temporary project

Queries		Views
Name	Query Expression	Name
Q7	(((swissprot-Organism:homo*) & [swissprot-Organism:sapiens*]) [swissprot-Organism:homo sapiens*]) & (((swissprot-Description:kinase*) & [swissprot-Description:inhibitor*]) [swissprot-Description:kinase inhibitor*])	Embl_View Swiss_View Swiss_Emdl_View
Q6	[libs={embl swissprot}-Organism:toad*]	
Q5	[swissprot-ID:rat*]	
Q4	[swissprot-Description:cellulase*]	
Q3	[embl-Description:cellulase*]	
Q2	[embl-Description:kinase*]	
Q1	[swissprot-AllText:cyclase*]	

Figure 2.3 Project Manager for temporary projects.

The **Project Manager** page provides you with the necessary tools to manage your SRS projects. Here you can save projects and open previously saved projects. After completing your project you may want to save your work to disk. (The next time you access SRS, unsaved temporary project work is unlikely to be available from the server.) Saving is also useful for moving temporary work to your permanent project history list, or permanent project work to a temporary project, or if you want to share project work with another user.

Unlike permanent projects, you cannot rename, delete, or switch between temporary projects, nor can you share queries and views. This means that any queries or views created for

one project will usually have to be recreated from scratch if you wish to use them with another (temporary) project.

The **Project Manager** page is described in section 2.4 “Using the Project Manager” (p. 47).

2.3 Permanent Projects

Permanent projects exist within SRS user accounts. There are two types of account: secure and non-secure.

- A secure user account gives you an httpd password. This ensures that the account and the permanent projects within it are available only to those authorized to access it.
- A non-secure user account also uses a login procedure to identify the account that you wish to open, but does not enforce any kind of access control. Anyone who knows the account name can view the account and the permanent projects within it.

A permanent project, regardless of the type of account used, stores all your project data in a single location from where it can be recalled.

Within your user account you can swap between projects and move work from one project to another, as well as retrieving saved projects, e.g. from other accounts to which you have access or from temporary projects that have been saved.

Note: Check with your SRS Administrator about your site policy for permanent projects. They may need to create an account for you.

Note: It is important that spaces and other special characters are not used in usernames (or passwords) because some systems do not handle them properly. Use an underscore, or start new words with a capital letter instead.

2.3.1 Starting a Permanent Project

This section describes how to open a permanent project by entering your user account. SRS will automatically create a new permanent project within your account if you are entering it for the first time. If you are re-entering an existing account, then the projects within that account will be available to you.

To start or return to a permanent project:

1. Click the **Start a Permanent Project** link on the **Start** page.

Secure or Non-Secure?

When you start a permanent project, your system will determine whether it is set up for secure or non-secure accounts. This was determined when SRS was installed, and you cannot choose at this point whether to use a secure or a non-secure account.

The system will display a **Log in** dialog, allowing you to log in to your account. Typical dialogs for secure and non-secure accounts are given in Figure 2.4 and Figure 2.5 respectively.


Note: The dialog boxes shown throughout this manual were generated in the Windows32 version of Internet Explorer 5. The dialog boxes you see may look different, depending on your system and web browser.



Figure 2.4 SRS secure (password protected) account **Log in** dialog box.



Figure 2.5 SRS non-secure account **Log in** dialog box.

2. For secure account access complete the login information. The web browser will prompt you to give a **User Name** and **Password**. You may need to ask your SRS Administrator for these because they may not be the same as your system account ID and password. For non-secure account access enter your **user name** at the prompt.
3. Click  .

In either case the **Project Manager** page is displayed. From here you can continue working with the most recently used project, switch to another project, or create a new project. This is described in section 2.3.2 “Managing Permanent Projects” (p. 46) and section 2.4 “Using the Project Manager” (p. 47).

2.3.2 Managing Permanent Projects

If you followed the steps outlined in section 2.3.1 “Starting a Permanent Project” (p. 44), the **Project Manager** page will be displayed (see Figure 2.6).

The **Project Manager** page provides you with the necessary tools to manage your SRS projects. It allows you to move queries or views between projects, save and open projects, delete unwanted projects, start new projects, and swap to any other project currently on your account.

The **Project Manager** page is described in section 2.4 “Using the Project Manager” (p. 47).

Reset

Session
Account Name: helenp
Session: project1

project1 Options
Save to desktop: **Save**
Rename project:
project1 **Rename**
Delete project: **Delete**

Other Projects
Create a project:
New Project
Open from desktop:
Browse...
Open
Switch projects:
project1 **Switch**

Contents of project1

Queries		Views
Name	Query Expression	Name
<input type="checkbox"/> Q3	[[swissprot-AllText:alcohol*] & [swissprot-AllText:dehydrogenase*]]	<input type="checkbox"/> Embl_View
<input type="checkbox"/> Q2	[swissprot-AllText:dehydrogenase*]	
<input type="checkbox"/> Q1	[swissprot-AllText:alcohol*]	

Figure 2.6 SRS Project Manager page for permanent projects.

2.4 Using the Project Manager

The Project Manager page is used for project administration. From this page you can create a new project, switch to another project, copy information between projects, open saved projects, save projects, rename projects and delete projects.

These functions are described below.

Note: Some functions, namely creating a new project, switching to other projects, deleting and renaming projects, are not available for temporary projects.

2.4.1 Creating a New Project

The **Project Manager** page allows you to create new permanent projects within an account and move between them (see also section 2.4.2 “Switching to Another Project”, p. 48).

Note: This feature is not available for temporary projects.

To create a new project:

1. Click the  button.


This will start a new project and display the **Select Databanks To Search** page.

2.4.2 Switching to Another Project

The **Project Manager** page allows you to move between projects that have been saved in your account.

Note: This feature is not available for temporary projects.

To return to a project that you have already saved:

1. From the **Project Manager** page, select the project you want to resume from the drop-down list in the **Other Projects** box.
2. Click the  button.
The project you have just selected is now displayed on the **Project Manager** page. You can work with this project as usual.

2.4.3 Copying Project Information

You can copy information such as views and queries from one project to another.

Note: This option will only be available if you have more than one project in your account. This option is not available for temporary projects.

To copy project information:

1. Make sure you are in the project from which you want to copy items (see section 2.4.2 “Switching to Another Project”, p. 48). Select the items you wish to copy by ticking the check boxes beside those items.

Contents of *project1*

Queries		Views
Name	Query Expression	Name
<input checked="" type="checkbox"/> Q3	[(swissprot-AllText:alcohol*) & [swissprot-AllText:dehydrogenase*)]	<input checked="" type="checkbox"/> Embl_View
<input type="checkbox"/> Q2	[swissprot-AllText:dehydrogenase*]	
<input type="checkbox"/> Q1	[swissprot-AllText:alcohol*]	

Figure 2.7 Choosing items to be copied.

2. Choose the project to which you want the items to be copied using the drop-down list in the **Project Options** box:

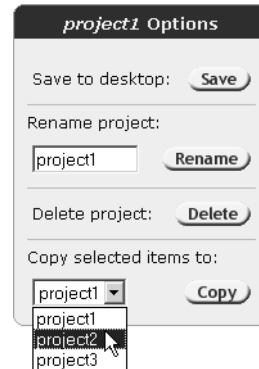


Figure 2.8 Choosing a project to which to copy items.

3. Click on the **Copy** button to copy the items.

This will copy the selected queries and views to the chosen project, and switch to using that project.

2.4.4 Saving a Project

If you want to share a project with other users, or if you want to move a temporary project to a permanent project list, you first need to save a file containing the project work to your local disk.

To save a project:

1. Make it the current project.
(See section 2.4.2 “Switching to Another Project”, p. 48.)
2. Click the **Save** button, in the **Project Options** box.

3. Select **Save this file to disk** when prompted.

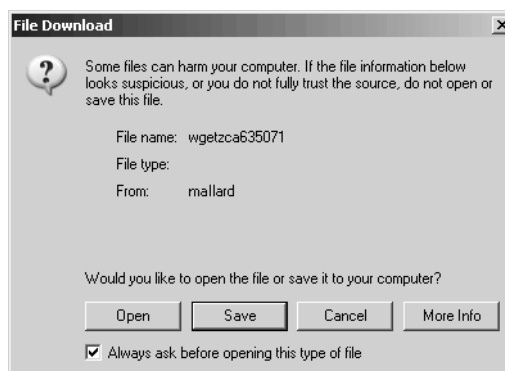


Figure 2.9 File Download dialog.

4. Click **Save**.
5. In the **Save As** dialog box, give the project file a name.

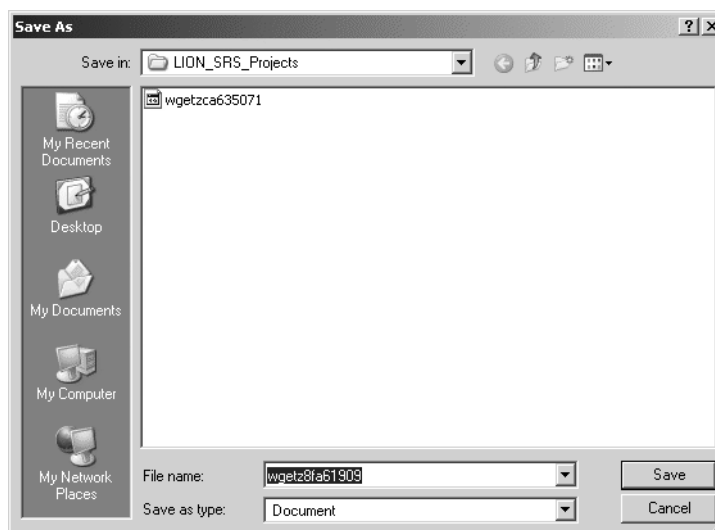


Figure 2.10 Save As dialog.


Note: It is better not to include spaces and other special characters in names as some systems do not handle them prop-

erly. Use an underscore, or start new words with a capital letter instead.

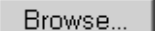
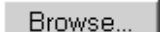

6. Click .

The file is now saved and the system is ready to save another project, if you wish.

2.4.5 Opening a Saved Project

You can open a saved project (e.g. from a different account (to which you have access), or a temporary project that was saved) using the  button in the **Other Projects** box on the **Project Manager** page.

To open a project from your local disk:

1. Type the file name in the text box beside the  button, or use the  button to help locate your project.
2. Click the  button.
This will open the project, and you can start working with the project as usual.


2.4.6 Deleting a Project

When you no longer need a project you might want to delete it. This may become necessary if the number of projects for a particular permanent project account nears 99. (The upper

limit for the number of projects available to a permanent project account is 99.)

Note: This feature is not available for temporary projects.

To delete a project:

1. Set the project that you want to delete as the current project.
(See section 2.4.2 “Switching to Another Project”, p. 48.)
2. Click the  button.

2.4.7 Renaming a Project


Default project names (***project1***, ***project2***, ***projectN***) can be personalized to make it easier to keep track of the work you have performed in each project.

Note: This feature is not available for temporary projects.

To rename a project:

1. Type a new name in the **Rename project:** text box in the **Project Options** box.

Note: It is better not to include spaces and other special characters in names as some systems do not handle them properly. Use an underscore, or start new words with a capital letter instead.

2. Click the  button.
The page will be refreshed showing the new project name.

The screenshot displays the LION SRS interface. At the top, there is a navigation bar with the LION logo on the left and a 'Help Center' link with a question mark icon on the right. Below the logo, a series of tabs are visible: 'Quick Searches', 'Select Databanks', 'Query Form', 'Tools', 'Results', 'Projects', 'Custom Views', and 'Information'. The 'Projects' tab is currently selected.

Below the navigation bar, the main content area is titled 'SRS'. On the left side, there is a 'Reset' button. Below it, a 'Session' box shows 'Account Name: helenp' and 'Session: project2'. Underneath, the 'project2 Options' panel is open, showing several actions: 'Save to desktop:' with a 'Save' button; 'Rename project:' with a text input field containing 'Transit Pepti' and a 'Rename' button; 'Delete project:' with a 'Delete' button; and 'Copy selected items to:' with a dropdown menu set to 'project2' and a 'Copy' button. A mouse cursor is hovering over the 'Rename' button, and a tooltip with the text 'Rename' is visible.

On the right side, the 'Contents of project2' section is displayed. It contains two tables: 'Queries' and 'Views'. The 'Queries' table has columns 'Name' and 'Query Expression' and lists two items: 'Q2 ([swissprot-FtKey:transit*] > parent)' and 'Q1 ([embl-FtKey:transit_peptide*] > parent)', each with an unchecked checkbox. The 'Views' table has a 'Name' column and lists two items: 'SwissProt_View' and 'Embl_View', each with an unchecked checkbox.

Figure 2.11 Personalized project names.

QUERYING WITH SRS

One of the greatest advantages of using SRS is its ability to search databanks, and the range of methods available to do so. This chapter introduces you to the ways in which you can perform a query using SRS.

By the end of this chapter you will have learned:

- What a search term is and how to construct one.
- How to perform a **Quick Search**.
- How to use the **Standard Query Form** and **Extended Query Form**.
- What subentries are and the advantages of using them in queries.
- How to search for entries that relate to a specific expression or phrase (**Expression Query**).
- How to sort your query results.
- How to browse the index.

3.1 Search Terms

Whichever search method you use, you will need to use some sort of search term. Search terms used in SRS can be categorized as follows:

- Single-word search (see section 3.1.1 “Single Word Searches”, p. 56).
- Multiple-word phrases (see section 3.1.2 “Multiple-Word Searches”, p. 57).
- Numbers and dates (see section 3.1.3 “Numbers and Dates”, p. 57).
- Regular expressions (see section 3.1.4 “Regular Expressions”, p. 59).
- Wildcards (see section 3.1.5 “Wildcards”, p. 59).

The rest of this section explains how search terms are constructed. The individual search methods are detailed later in this chapter (see section 3.2 “Quick Searches”, p. 60, section 3.3 “Query Forms”, p. 62, section 3.5 “Expression Queries”, p. 79 and section 3.7 “Browse Index”, p. 86).

3.1.1 Single Word Searches

When you search a databank for a single word, such as “reductase”, in a single field, the results of your query will be a list of entries that included that word in the index for the selected datafield.

3.1.2 Multiple-Word Searches

You can search for a phrase having more than one word, such as **aldehyde reductase**, in several ways. For example, if the phrase is enclosed in quotation marks (“**aldehyde reductase**”), SRS will search for the complete string, only returning exact matches. If the phrase is not enclosed in quotation marks (**aldehyde reductase**), SRS will search for each word separately — combining the results in some fashion. The default is to require an entry to contain each of the words in the phrase so that a search for **aldehyde reductase** (without quotation marks) would find entries that have either **aldehyde** and **reductase** or the complete string **aldehyde reductase**.

You can make the relationship between the words explicit by including an operator in the string (see section 8.3.4 “Operators”, p. 168). For example, you could search for **aldehyde & reductase** (AND), or **aldehyde | reductase** (OR), or **aldehyde ! reductase** (BUTNOT).

3.1.3 Numbers and Dates

SRS uses numerical entries to allow you to search for information such as dates, or entries with a specified length.

Numerical entries can be combined into expressions using various operators. This allows you to search within particular ranges. Typically these operators are: less than; less than or equal; greater than; and greater than or equal. These are created by combining two simpler operators, namely the

colon, :, and the exclamation mark, !. The colon means greater than or less than depending on which side of your number it lies. The exclamation mark indicates that the number to the right of the exclamation mark is to be excluded. The exclamation mark can be regarded as “not” or “not equal to”. It is probably easier to demonstrate this with some examples:

12:15

Greater than or equal to 12, but less than or equal to 15.

12:

Greater than or equal to 12, with no specific upper limit.

!12:

Greater than but not equal to 12, with no specific upper limit.

:12

Less than or equal to 12, with no specific lower limit.

:!12

Less than but not equal to 12, with no specific lower limit.

When using the **Extended Query Form** (see section 3.3.3 “Using the Extended Query Form”, p. 65) such a range search is often simplified for you, so that you do not need to use the query language syntax (see Figure 3.1). Using this method allows you to select the operators from a drop-down list.

The screenshot shows a web-based query form with several fields and dropdown menus. The fields are: 'Entry Creation Date' (with a 'between' dropdown, values 11, Mar, 2002, 9, Mar, 2003), 'LastUpdated' (with a 'select' dropdown, values 1, Jan, YYYY, 1, Jan, YYYY), and 'Sequence Length' (with a 'select on' dropdown, values >=, 212, <=, 400). There are checkboxes to the right of each field. Below these is a 'References subentry fields' section with a 'References' dropdown. At the bottom, there is a 'View results using:' field with a dropdown showing '*NameEntry' and a 'Search' button.

Figure 3.1 Extended Query Form showing a numerical data query being set up, using the drop-down lists.

See chapter 8, “SRS Query Language”, (p. 155) for more details on the syntax for numerical searches.

3.1.4 Regular Expressions

Regular expression searches are useful when you wish to search for alternative spellings, or words with the same root, but different suffixes, etc. Regular expressions allow you to use a combination of characters, along with regular expression characters, and get a list of matching entries as your result. You always need to include the forward slash (/) character at the start and end of the regular expression string. For example, “/^phos/” will find all words beginning with “phos” (e.g., phosphate, phosphorylase), “/ase\$/” will find words ending in “ase” (e.g., kinase, phosphatase). (See section “Regular Expressions” (p. 159) for an explanation.)

3.1.5 Wildcards

The SRS query language also uses the familiar “*” and “?” wildcards. This is usually much simpler than using regular

expressions for basic searches. For example, “cell*ase” would find all words starting with “cell” and ending in “ase” (e.g., cellobiase, cellobiohydrolase, cellulase).

3.2 Quick Searches

3.2.1 SRS Quick Search Page

Depending on the databanks which are available in your SRS installation, you may be able to perform a small number of Quick Searches from the **SRS Quick Search** page. Typically you will have the following searches available:

- Quick Text Search
- Sequence Similarity/Homology Search

These allow you to make some pre-formatted searches without needing to understand SRS in great detail.

For more information, see the help files for the SRS Quick Search page, (“Querying - SRS Quick Search”, in the *Online Help pages*).

3.2.2 About Quick Search

Quick Search works by searching all datafields of type **text** in the selected databanks, and is an alternative method of generating query results fast without using the query forms.

If you do not get the results you expect, then a slight change to the search phrase may help to target your search and thus may yield improved results. Alternatively, try one of the other search methods, as these allow you to specify your area of interest in more detail.

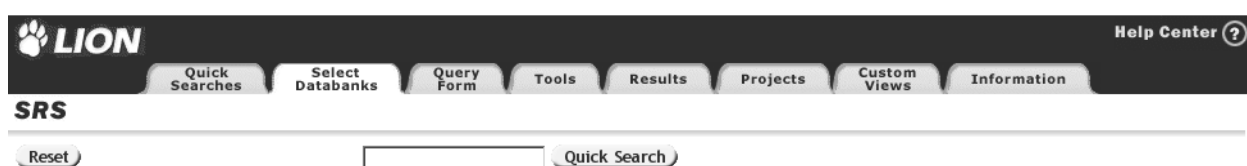


Figure 3.2 The location of **Quick Search** on the **Select Databanks To Search** page.

Note:The **Quick Search** option combines all the fields with a data type of **Text** using the **OR** operator.

3.2.3 Using Quick Search

Quick Search allows you to search with a minimum number of steps from selecting the databanks to viewing the results.

1. Tick the check box(es) to the left of the databank(s) you wish to search, on the **Select Databanks To Search** page,.
2. Enter the search term in the text box beside the **Quick Search** button.
Use a suitable word or expression (see section 3.1 “Search Terms”, p. 56).
3. Click the **Quick Search** button.

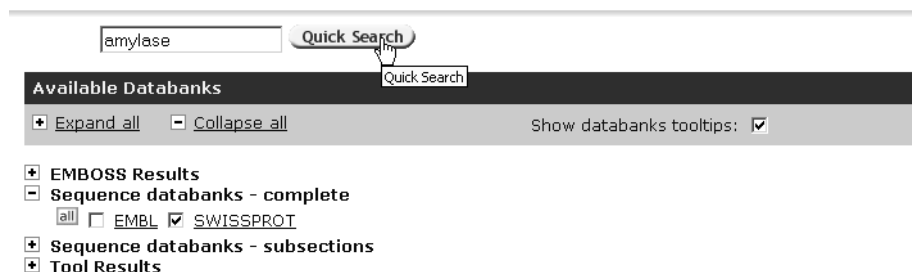


Figure 3.3 Quick Search of the SWISS-PROT databank for “amylase”.

3.3 Query Forms

3.3.1 About Query Forms

Results of queries can usually be refined by adding to the information used for the search. If you are able to supply a larger range of information about your area of interest, then SRS can target your query more precisely. The query forms allow you to enter specified information about your subject, in various fields. There are two types of query form: the **Standard Query Form** and **Extended Query Form**. These are described in the remainder of this section.

3.3.2 Using the Standard Query Form

The **Standard Query Form** allows you to enter up to four separate search terms and search against up to four different datafields simultaneously.

LION Help Center ?

Quick Searches Select Databanks Query Form Tools Results Projects Custom Views Information

SRS

Reset search SWISSPROT

Search Options

Combine search terms with: & (AND)

Use wildcards

Get results of type: Entry

Result Display Options

View results using: SeqSimpleView or Create a view

Show 30 results per page

Fields you can search

Your search terms

In a single field, you can separate multiple values by &, |, !

Fields you can search	Your search terms
AllText	
AllText	
AllText	
AllText	

Create a view

Select the fields you want displayed in your view and choose the format

Choose 1 or more fields: ID, Accession Number, Primary Accession Number, Description, Gene Name, Keywords, Entry Creation Date

Display As: Table List


Sequence Format: swiss

Tips: To do more advanced queries, use the **Extended Query** Form.

Figure 3.4 Standard Query Form.

1. From the **Select Databanks To Search** page, select the databank you want to search, by ticking the check box to the left of a databank name.
2. Click the **Standard Query Form** button from the **Query forms** box. This will display the **Standard Query Form**, Figure 3.4.
3. Enter the search phrase(s). See section 3.1 “Search Terms” (p. 56). In each of the four search rows, specify the field to be searched using the left-hand drop-down list and enter the

search word or phrase in the corresponding right-hand text box. You do not have to use all four rows, but greater detail may help refine your search.

4. Use **Combine search terms with:** drop-down list to specify the method by which the various search terms should be combined: **& (AND)**, **| (OR)** or **! (BUTNOT)**.
5. Choose the type of entries you wish to retrieve, using the **Get results of type:** drop-down list. Leave it set at **Entry** for this example.
6. Choose a view using the **View results using:** drop-down list, or choose **Create a view**. (See chapter 5, “Views”, p. 109, for more information on views.)
7. Specify the number of entries to display per page, using the drop-down list.
8. Click the  button.

SRS

[Reset](#) search [EMBL](#)

Search Options

Combine search terms with:

Use wildcards

Get results of type:

Fields you can search	Your search terms
In a single field, you can separate multiple values by &, , !	
<input type="text" value="Keywords"/>	<input type="text" value="kinase"/>
<input type="text" value="Entry Creation Date"/>	<input type="text" value="01-Jan-2002"/>
<input type="text" value="AllText"/>	<input type="text"/>
<input type="text" value="AllText"/>	<input type="text"/>

Result Display Options

View results using:

or

Create a view

Show results per page

Create a view

Select the fields you want displayed in your view and choose the format

Choose 1 or more fields:

Display As: Table List

Sequence Format:

Tips

To do more advanced queries, use the [Extended Query](#) Form.

Figure 3.5 Standard Query Form showing a query for EMBL for entries with the keyword, “kinase”, that were added to the databank after January 1, 2002.

Note: If you want to know more about any of the fields, choose that field from the drop-down list, and click the **i** beside it.

3.3.3 Using the Extended Query Form

The **Extended Query Form** lists all the common datafields and allows you to enter search terms for as many of these fields as you want.

SRS

[Reset](#) search EMBL

Search Options	Fields you can search	Your search terms	Create a view
Combine search terms with: <input type="text" value="& (AND)"/>	In a single field, you can separate multiple values by &, , !		<input type="button" value="Search"/>
Use wildcards <input checked="" type="checkbox"/>	<u>AllText</u>	<input type="text"/>	<input type="checkbox"/>
Get results of type: <input type="text" value="Entry"/>	<u>ID</u>	<input type="text"/>	<input type="checkbox"/>
	<u>Division</u>	<input type="checkbox"/> est <input type="checkbox"/> fun <input type="checkbox"/> gss <input type="checkbox"/> htc <input type="checkbox"/> htg <input type="checkbox"/> hum <input type="checkbox"/> inv <input type="checkbox"/> mam <input type="checkbox"/> mus <input type="checkbox"/> org <input type="checkbox"/> phg <input type="checkbox"/> pln <input type="checkbox"/> pro <input type="checkbox"/> rod <input type="checkbox"/> sts <input type="checkbox"/> syn <input type="checkbox"/> unc <input type="checkbox"/> vrl <input type="checkbox"/> vrt	<input type="checkbox"/>
	<u>Accession Number</u>	<input type="text"/>	<input type="checkbox"/>
	<u>Primary Accession Number</u>	<input type="text"/>	<input type="checkbox"/>
	<u>SeqVersion</u>	<input type="text"/>	<input type="checkbox"/>
	<u>Molecule</u>	<input type="checkbox"/> circular dna <input type="checkbox"/> circular rna <input type="checkbox"/> dna <input type="checkbox"/> rna <input type="checkbox"/> xxx	<input type="checkbox"/>
	<u>Description</u>	<input type="text"/>	<input type="checkbox"/>
	<u>Keywords</u>	<input type="text"/>	<input type="checkbox"/>
	<u>Organism Name</u>	<input type="text"/>	<input type="checkbox"/>
	<u>Taxon</u>	<input type="text"/>	<input type="checkbox"/>
	<u>Organelle</u>	<input type="text"/>	<input type="checkbox"/>
	<u>Comment</u>	<input type="text"/>	<input type="checkbox"/>
	<u>Entry Creation Date</u>	<input type="text" value="select"/> <input type="text" value="1"/> <input type="text" value="Jan"/> <input type="text" value="YYYY"/> <input type="text" value="1"/> <input type="text" value="Jan"/> <input type="text" value="YYYY"/>	<input type="checkbox"/>
	<u>LastUpdated</u>	<input type="text" value="select"/> <input type="text" value="1"/> <input type="text" value="Jan"/> <input type="text" value="YYYY"/> <input type="text" value="1"/> <input type="text" value="Jan"/> <input type="text" value="YYYY"/>	<input type="checkbox"/>
	<u>Sequence Length</u>	<input type="text" value=">="/> <input type="text"/> <input type="text" value="<="/> <input type="text"/>	<input type="checkbox"/>

Result Display Options

View results using: SeqSimpleView Create a view using selected fields

Sequence Format:

Show results per page

Tips

You can also use the **Standard Query** Form.

Figure 3.6 Part of the **Extended Query Form**.

1. From the **Select Databanks To Search** page, select the databank(s) you wish to search by ticking the check box(es) to the left of the databank name(s).
2. Select **Extended Query Form** button from the **Search Options** box. This will display the **Extended Query Form**.
3. Enter the search phrase(s) (see section 3.1 “Search Terms”, p. 56) in the text boxes beside the relevant fields.

Note: You do not have to complete all the fields, but the greater the detail, the more refined your search will be.

4. Use **Combine search terms with:** drop-down list to specify the method by which the various search terms should be combined: **& (AND)**, **| (OR)** or **! (BUTNOT)**.
5. Choose the type of entries you wish to retrieve, using the **Get results of type:** drop-down list. Leave it set at **Entry** for this example.
6. Choose a view using the **View results using:** drop-down list, or choose **Create a view**. (See chapter 5, “Views”, p. 109, for more information on views.)
7. Choose the **Sequence format:** using the drop-down menu.
8. Specify the number of entries to display per page, using the drop-down list.
9. Click the **Search** button.

<u>Description</u>	histamine	<input checked="" type="checkbox"/>
<u>Keywords</u>		<input checked="" type="checkbox"/>
<u>Organism Name</u>	human	<input type="checkbox"/>
<u>Taxon</u>		<input type="checkbox"/>
<u>Organelle</u>		<input type="checkbox"/>
<u>Comment</u>		<input type="checkbox"/>
<u>Entry Creation Date</u>	after 1 Mar 2002 1 Jan YYYY	<input checked="" type="checkbox"/>
<u>LastUpdated</u>	select 1 Jan YYYY 1 Jan YYYY	<input checked="" type="checkbox"/>
<u>Sequence Length</u>	>= <=	<input checked="" type="checkbox"/>

Figure 3.7 Extended Query Form querying EMBL for entries about “histamine” in humans, that were added to the databank after 1st March, 2002. The search uses the **Description**, **Organism** and **Entry Creation Date** fields.

3.4 Subentries

3.4.1 About Subentries

Subentries are parts of data entries that have an internal structure, which needs to be conserved when searching a database. A typical example of this might be a databank whose entries contain multiple literature references, or one that contains features (e.g. within a sequence) in which the relationship between the features is important.

3.4.2 Use of Subentries

Subentries are particularly useful for queries where a relationship between items is important.

Suppose a user wants to search for references written jointly by authors Smith and Jones, but does not want to find papers written by only one of those authors.

In a case like this it becomes important that the user can search for data where Smith and Jones occur within the same literature reference, whilst excluding entries that only contain papers written by each author separately, with no papers authored jointly by Smith & Jones.

A search of entries, even where it is specified that both Smith and Jones must appear within an entry will reveal all entries containing references by Smith & Jones jointly, and those which contain separate papers by Smith and by Jones.

A search using subentries, on the other hand will allow the user to specify a relationship between the two authors, i.e. to specify that there must be at least one reference that is written jointly.

Using subentries allows each literature reference to become a specific subentry: in effect creating a mini-entry within an entry.

The next section will take you through a typical subentry query.

3.4.3 Querying Using Subentries

This worked example will take you through using subentries to make queries using both of the **Query Forms**.

Searching for Entries which Reference Papers that include those by Smith and by Jones using the Standard Query Form

In this example entries will be retrieved that reference papers by Smith and Jones. For an entry to be retrieved there must be papers whose authors include Smith and Jones, but Smith and Jones need not be co-authors of the same paper.

1. From the **Select Databanks To Search** page, select the databank you wish to search by ticking the check box to the left of the databank name, e.g. SWISS-PROT.
2. Select **Standard Query Form** button from the **Search Options** box. This will display the **Standard Query Form**.

- Use the first drop-down datafield list to select an appropriate subentry field, e.g. **Reference:Authors**.

Note:Datafields for subentry fields are shown in the drop-down lists using a special format to identify them. In many cases, they are the same as for SWISS-PROT, i.e. a subentry name (e.g. **Reference**) followed by a field within that subentry (e.g. **Authors**); the two parts are separated by a colon (:). Hence **Reference:Authors** will search for authors fields within subentries.

- Enter **smith** in the first text box, beside where you selected **Reference:Authors**. The search is case insensitive, so it does not matter whether smith is written with a lower or upper case S.
- Repeat the process for the next line of information, using the name **jones** rather than smith.

Fields you can search	Your search terms
Reference: Authors	smith
Reference: Authors	jones
AllText	
AllText	

Figure 3.8 Standard Query Form showing a query for entries, using the **References:Authors** subentry fields.

- Leave **Combine searches with:** set to **& (AND)**.
- Leave **Get entries of type:** set to **Entry**. This will cause entire entries to be retrieved. Setting this option to one of the other (subentry) fields, e.g. **Reference** will tell SRS to

retrieve each reference (which fulfils the search criteria) as a separate entity.

8. Click the **Search** button.

Query "[([swissprot-Authors:smith*] > parent) & ([swissprot-Authors:jones*] > parent))" found 1336 entries next

SWISSPROT	Accession	Description	SeqLength
<input type="checkbox"/> SWISSPROT:AATM_LUPAN	P26563	Aspartate aminotransferase-P2, mitochondrial precursor (EC 2.6.1.1) (Transaminase A) (Fragment).	454
<input type="checkbox"/> SWISSPROT:AMYG_YEAST	P08019	Glucoamylase, intracellular sporulation-specific (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase) (1,4-alpha-D-glucan glucohydrolase).	549
<input type="checkbox"/> SWISSPROT:AMYH_YEAST	P08640	Glucoamylase S1/S2 precursor (EC 3.2.1.3) (Glucan 1,4-alpha- glucosidase) (1,4-alpha-D-glucan glucohydrolase).	1367
<input type="checkbox"/> SWISSPROT:AR16_YEAST	P40518	ARP2/3 complex 16 kDa subunit (P16-ARC).	154
<input type="checkbox"/> SWISSPROT:AXL2_YEAST	P38928	AXL2 protein precursor (SRO4 protein).	823
<input type="checkbox"/> SWISSPROT:BNR1_YEAST	P40450	BNI1 related protein 1.	1375
<input type="checkbox"/> SWISSPROT:COPE_YEAST	P40509	Coatomer epsilon subunit (Epsilon-coat protein) (Epsilon-COP).	296
<input type="checkbox"/> SWISSPROT:DA81_YEAST	P21657	Transcriptional activator protein DAL81 (Regulatory protein UGA35).	970

Figure 3.9 Query Result page showing the results of a query for entries, using the **References:Authors** subentry fields, using two separate text boxes to search on the **Standard Query Form**.

Using two separate text boxes for the subentries search retrieves all entries containing references by Smith, and by Jones, including those where there are no papers which were co-authored by Smith and Jones. Figure 3.10 shows an entry from the above search.

References	
1.	Tarver,A.P.; Clark,D.P.; Diamond,G.; Russell,J.P.; Erdjument-Bromage,H.; Tempst,P.; Cohen,K.S.; Jones,D.E.; Sweeney,R.W.; Wines,M.; Hwang,S.; Bevins,C.L.; Enteric beta-defensin: molecular cloning and characterization of a gene with inducible intestinal epithelial cell expression associated with Cryptosporidium parvum infection. Infect. Immun. 66:1045 (1998)
	Medline 98147718
	PubMed 9488394
	Position sequence from n.a.
	Comment tissue=small intestine;
2.	Selsted,M.E.; Tang,Y.-Q.; Morris,W.L.; McGuire,P.A.; Novotny,M.J.; Smith,W.; Henschen,A.H.; Cullor,J.S.; Purification, primary structures, and antibacterial activities of beta-defensins, a new family of antimicrobial peptides from bovine neutrophils. J. Biol. Chem. 268:6641 (1993)
	Medline 93203264
	PubMed 8454635
	Position sequence of 15-67.
	Comment strain=hereford; tissue=neutrophils;

Figure 3.10 Entry page showing that some entries reference only papers that are not co-authored by Smith & Jones. (SWISS-PROT accession number P46161.)

In effect, this search looks for reference subentries that contain smith or jones. It then takes the entries for each of the references matching the criteria, to create two lists of entries: one for smith and one for jones. It then combines those, seeking only those which are in both lists.

Searching for Entries which Reference Papers that are Co-authored by Smith & Jones using the Standard Query Form

In the above example, entries were retrieved that referenced papers by Smith and Jones, but without concern for whether Smith & Jones were co-authors of any papers.

If you wish to search for entries that reference papers that are always co-authored by Smith & Jones, you should modify your search, so that rather than searching in two separate text

boxes on the **Standard Query Form**, you search for both authors in a single text box.

1. From the **Select Databanks To Search** page, select a databank to search, e.g. SWISS-PROT and go to the **Standard Query Form**.
2. Use the first drop-down datafield list to select an appropriate subentry field, e.g. **Reference:Authors**.
3. In the text box beside the datafield you have used, type **smith & jones**.

Note: Make sure you use the & sign which tells SRS that you want to look for both authors combined using the boolean operator & (AND), rather than | (OR), or ! (NOT).

Reset search SWISSPROT

Search Options	Fields you can search	Your search terms
Combine search terms with: & (AND)	Reference: Authors	smith & jones
Use wildcards <input checked="" type="checkbox"/>	AllText	
Get results of type: Entry	AllText	
	AllText	

Figure 3.11 Standard Query Form showing a query for entries, using the **References:Authors** subentry fields, searching for papers which are co-authored by Smith & Jones.

4. Leave **combine searches with** set to **& (AND)**, and **retrieve entries of type** set to **Entry**.
5. Click the **Search** button.

Query "[swissprot-Authors:smith*] & [swissprot-Authors:jones*] > parent " found 1214 entries next

SWISSPROT	Accession	Description	SeqLength
<input type="checkbox"/> SWISSPROT:AATM_LUPAN	P26563	Aspartate aminotransferase-P2, mitochondrial precursor (EC 2.6.1.1) (Transaminase A) (Fragment).	454
<input type="checkbox"/> SWISSPROT:AMYG_YEAST	P08019	Glucosylase, intracellular sporulation-specific (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase) (1,4-alpha-D-glucan glucohydrolase).	549
<input type="checkbox"/> SWISSPROT:AMYG_YEAST	P08640	Glucosylase S1/S2 precursor (EC 3.2.1.3) (Glucan 1,4-alpha- glucosidase) (1,4-alpha-D-glucan glucohydrolase).	1367
<input type="checkbox"/> SWISSPROT:AR16_YEAST	P40518	ARP2/3 complex 16 kDa subunit (P16-ARC).	154
<input type="checkbox"/> SWISSPROT:AXL2_YEAST	P38928	AXL2 protein precursor (SRO4 protein).	823
<input type="checkbox"/> SWISSPROT:BNR1_YEAST	P40450	BNR1 related protein 1.	1375
<input type="checkbox"/> SWISSPROT:COPE_YEAST	P40509	Coatomer epsilon subunit (Epsilon-coat protein) (Epsilon-COP).	296

Figure 3.12 Query Result page showing the results of a query for entries, using the **References:Authors** subentry fields, using a single text box with two search terms, to search on the **Standard Query Form**.

In essence, this search looks for reference subentries that contain smith or jones. It then combines these to create a list of those that contain both Smith and Jones in a single reference subentry. It then takes the entries for each of the references matching the criteria. This creates a single list of entries, which contain references to papers which are co-authored by Smith and Jones.

Searching for Subentries which Reference Papers that are Co-authored by Smith & Jones using the Standard Query Form

If you wish to retrieve the subentry fields that are found, rather than the parent entries, the searches can be repeated with the **retrieve entries of type** set to **Reference**.

1. Repeat one of the above searches, but set **retrieve entries of type to Reference**.

Reset search SWISSPROT

Search Options	Fields you can search	Your search terms
Combine search terms with: <input type="text" value="& (AND)"/>	In a single field, you can separate multiple values by &, , ! Search	
Use wildcards <input checked="" type="checkbox"/>	<input type="text" value="Reference: Authors"/>	<input type="text" value="smith"/>
Get results of type: <input type="text" value="Reference"/>	<input type="text" value="Reference: Authors"/>	<input type="text" value="jones"/>
	<input type="text" value="AllText"/>	<input type="text"/>
	<input type="text" value="AllText"/>	<input type="text"/>

Figure 3.13 Standard Query Form showing a query for subentries, using the **References:Authors** fields.

This search will retrieve the subentries that match the criteria rather than the parent entries. Otherwise, the search behaves similarly to the above, so that searching for the two terms in a single text box, combined with **&** will produce reference subentries that contain only co-authored papers, whilst searching for the two terms using separate text boxes will produce reference subentries which contain papers by both authors, but will not check whether they are co-authors.

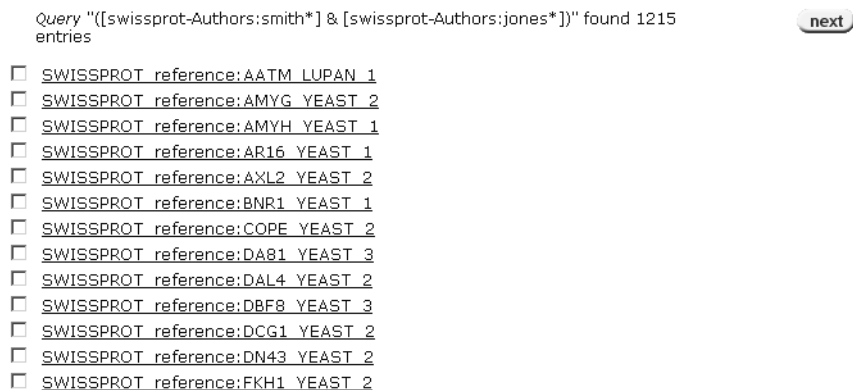


Figure 3.14 Query Result page showing the results of a query for subentries, using the **References:Authors** subentry fields, using the **Standard Query Form**.

Searching for Subentries which Reference Papers that are Co-authored by Smith & Jones using the Extended Query Form

Querying subentry fields on the **Extended Query Form** is similar to that on the **Standard Query Form**. However, because the **Extended Query Form** does not allow you to search the same field using two separate text boxes, it is not possible to repeat the type of search done in “Searching for Entries which Reference Papers that include those by Smith and by Jones using the Standard Query Form” (p. 69).

1. From the **Select Databanks To Search** page, select a databank to search, e.g. SWISS-PROT and go to the **Extended Query Form**.
2. Scroll down the **Extended Query Form** until you find the section marked, **Reference subentry fields**. Enter **smith & jones** into the **Authors** text box.

Reference subentry fields Reference ▾

View results using: *Names only* Search

<u>Authors</u>	<input type="text" value="smith & jones"/>	<input checked="" type="checkbox"/>
<u>Title</u>	<input type="text"/>	<input checked="" type="checkbox"/>
<u>Journal</u>	<input type="text"/>	<input checked="" type="checkbox"/>
<u>VolumeNo</u>	>= <input type="text"/> <= <input type="text"/>	<input checked="" type="checkbox"/>
<u>FirstPage</u>	>= <input type="text"/> <= <input type="text"/>	<input checked="" type="checkbox"/>
<u>Year</u>	>= <input type="text"/> <= <input type="text"/>	<input type="checkbox"/>
<u>MedlineID</u>	<input type="text"/>	<input type="checkbox"/>
<u>PubMedID</u>	<input type="text"/>	<input type="checkbox"/>
<u>RefPosition</u>	<input type="text"/>	<input type="checkbox"/>
<u>RefCommentCode</u>	<div style="border: 1px solid black; padding: 2px; display: inline-block;"> plasmid species strain tissue transposon </div>	<input type="checkbox"/>
<input checked="" type="radio"/> or <input type="radio"/> and		
<u>RefComment</u>	<input type="text"/>	<input type="checkbox"/>

Figure 3.15 Part of the **Extended Query Form** showing a query for subentries, using the **References:Authors** fields.

3. Choose whether you wish to retrieve entire entries, or the **Reference** subentries.

Reference subentry fields Reference ▾

View results using: *Names only*
 Entry
 Reference

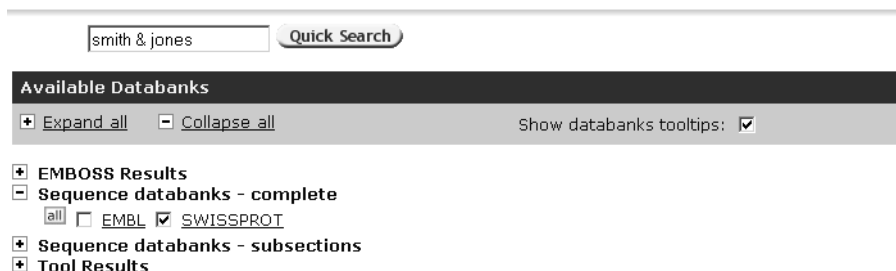
Figure 3.16 Choosing whether to retrieve entire entries or the **Reference** subentries.

4. Submit the query as usual.

Searching for Subentries which Reference Papers that are Co-authored by Smith & Jones using Quick Search

The quick search option will always return complete entries. You can, nevertheless, search for entries using the subentries as search items, however, as Quick Search simply searches all the text fields in an entry, you cannot control the field within which the search text occurs.

1. From the **Select Databanks To Search** page, select a databank to search, e.g. SWISS-PROT.
2. Enter **smith & jones** in the **Quick Search** text box.
3. Click the **Quick Search** button.



smith & jones Quick Search

Available Databanks

Expand all Collapse all Show databanks tooltips:

- EMBOSS Results
- Sequence databanks - complete
 - all EMBL SWISSPROT
- Sequence databanks - subsections
- Tool Results

Figure 3.17 Quick Search of SWISS-PROT for smith & jones.

In this case, you cannot tell SRS that you want only to search the subentries, and you have no way of indicating that you wish to search only for papers that are co-authored by Smith & Jones. The results will be entries which contain papers by both authors, without any control over whether the papers are co-authored. The same results could be achieved using the **Alltext** fields in either the **Standard** or **Extended Query Forms**.

Query "[[swissprot-ALLTEXT:smith*] & [swissprot-ALLTEXT:jones*]]" found 1336 entries next

SWISSPROT	Accession	Description	SeqLength
<input type="checkbox"/> SWISSPROT:AATM_LUPAN	P26563	Aspartate aminotransferase-P2, mitochondrial precursor (EC 2.6.1.1) (Transaminase A) (Fragment).	454
<input type="checkbox"/> SWISSPROT:AMYG_YEAST	P08019	Glucoamylase, intracellular sporulation-specific (EC 3.2.1.3) (Glucan 1,4-alpha-glucosidase) (1,4-alpha-D-glucan glucohydrolase).	549
<input type="checkbox"/> SWISSPROT:AMYH_YEAST	P08640	Glucoamylase S1/S2 precursor (EC 3.2.1.3) (Glucan 1,4-alpha- glucosidase) (1,4-alpha-D-glucan glucohydrolase).	1367
<input type="checkbox"/> SWISSPROT:AR16_YEAST	P40518	ARP2/3 complex 16 kDa subunit (P16-ARC).	154
<input type="checkbox"/> SWISSPROT:AXL2_YEAST	P38928	AXL2 protein precursor (SRO4 protein).	823
<input type="checkbox"/> SWISSPROT:BNR1_YEAST	P40450	BNI1 related protein 1.	1375
<input type="checkbox"/> SWISSPROT:COPE_YEAST	P40509	Coatomer epsilon subunit (Epsilon-coat protein) (Epsilon-COP).	296

Figure 3.18 Quick Search of SWISS-PROT for smith & jones.

3.5 Expression Queries

3.5.1 Searches using a Query Expression

You can make queries using a query expression (Figure 3.19) which you will find on the **Manage your Query Results** page. You can also use this method to combine, link, or refine the results of existing queries.

Search using a query expression: Search

Figure 3.19 Expression text box.

3.5.2 Using Expression Queries

1. On the **Manage your Query Results** page, enter your query into the **Search using a query expression:** text box.
2. Click on the **Search** button.

For example, to search for all entries that satisfy two existing queries, **Q1** and **Q2**, you could type:



A screenshot of a search input field with a light gray border. The text 'Q1 & Q2' is entered in the field. To the right of the text are three small, light gray square buttons: a left-pointing arrow, a right-pointing arrow, and a square button with a downward-pointing arrow.

This will return a new listing, **Q3**, that has all entries that are in both **Q1** and **Q2**.

If you wanted to search for the entries in query, **Q3**, that had links to the SWISS-PROT databank, you would type:



A screenshot of a search input field with a light gray border. The text 'Q3 < SWISSPROT' is entered in the field. To the right of the text are three small, light gray square buttons: a left-pointing arrow, a right-pointing arrow, and a square button with a downward-pointing arrow.

Typical operators allow you to combine searches using the standard logical functions (& (AND), | (OR), ! (BUT NOT)), and to look for links between result sets. For more help on linking using expressions, see section 4.5 “Expression Linking” (p. 104). See Table 8.5, page 168 for a list of the SRS query language operators that may be used.

This method of searching is very powerful because many detailed queries can be created using the SRS Query Language. (See chapter 8, “SRS Query Language”, p. 155 for more information.)

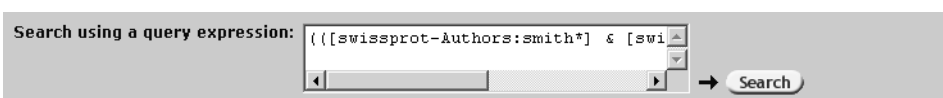
Using Expression Queries to Search Subentries

You can use Expression Queries to search for subentries. This requires a knowledge of the correct SRS query language. However, until you are confident enough to create your own queries, you can find out the correct query syntax by looking back at the **Query Result** page that you have previously generated. For example, use the text generated in “Searching for Entries which Reference Papers that are Co-authored by Smith & Jones using the Standard Query Form” (p. 72).

1. From the **Select Databanks To Search** page, select a databank to search, e.g. SWISS-PROT.
2. Click the Results tab to display the Manage your Query Results page. Enter the following text in the **Expression Query** text box:

(([swissprot-Authors:smith*] & [swissprot-Authors:jones*]) > parent)

The text should all be on one line, and the text box will scroll as you type. You can scroll back to the left to check what you have written, but it is often difficult to see the whole query. You could try typing the query in a suitable text editor, so that you can see it, and then using copy and paste to insert it into the **Expression Query** text box.



3.6 Sorting Results

3.6.1 What is Sorting?

Many queries will yield a large number of results. Sometimes, it is useful to be able to sort these to help you select the most relevant results. You could sort by almost any of the available fields. However, the fields which are available for sorting will be limited to those selected for the installation of SRS that you are using. In addition, a few fields are not appropriate for sorting. Some typical sortable fields are shown in Figure 3.20.

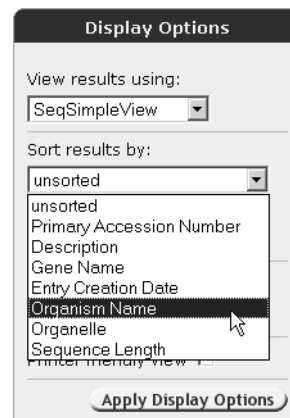


Figure 3.20 Display Options box, showing drop-down list of sortable fields on a **Query Result** page.

3.6.2 Sorting a Set of Results

1. First choose a set of results, e.g. by making a search, and view them on a **Query Result** page.

LION Help Center ?

Quick Searches | Select Databanks | Query Form | Tools | Results | Projects | Custom Views | Information

SRS

Reset Query "[swissprot-Description:kinase*]" found 3741 entries next

Apply Options to:

selected results only

unselected results only

Result Options

Launch analysis tool:

BlastP Launch

Show tools relevant to these results: Tools

Link to related information: Link

Save results: Save

Display Options

View results using:

SeqSimpleView

Sort results by:

unsorted

ascending

descending

Show 30 results per page

Printer friendly view

Apply Display Options

SWISSPROT	Accession	Description	SeqLength
<input type="checkbox"/> SWISSPROT:143F_MOUSE	P11576	14-3-3 protein eta (Protein kinase C inhibitor protein-1) (KCIP-1).	245
<input type="checkbox"/> SWISSPROT:143G_BOVIN	P29359	14-3-3 protein gamma (Protein kinase C inhibitor protein-1) (KCIP-1).	246
<input type="checkbox"/> SWISSPROT:143Z_MOUSE	P35215	14-3-3 protein zeta/delta (Protein kinase C inhibitor protein-1) (KCIP-1) (Mitochondrial import stimulation factor S1 subunit).	245
<input type="checkbox"/> SWISSPROT:143Z_SHEEP	P29361	14-3-3 protein zeta/delta (Protein kinase C inhibitor protein-1) (KCIP-1).	245
<input type="checkbox"/> SWISSPROT:AAIP_WHEAT	Q02066	Abcisic acid-inducible protein kinase (EC 2.7.1.-) (Fragment).	332
<input type="checkbox"/> SWISSPROT:AAK1_PIG	Q09136	5'-AMP-activated protein kinase, catalytic alpha-1 chain (EC 2.7.1.-) (AMPK alpha-1 chain) (63 kDa subunit) (AMPK) (Fragments).	132
<input type="checkbox"/> SWISSPROT:AAK2_PIG	Q28948	5'-AMP-activated protein kinase, catalytic alpha-2 chain (EC 2.7.1.-) (AMPK alpha-2 chain) (Fragment).	129
<input type="checkbox"/> SWISSPROT:AAKB_PIG	P80387	5'-AMP-activated protein kinase, beta-1 subunit (AMPK beta-1 chain) (AMPKb) (40 kDa subunit) (Fragment).	122
<input type="checkbox"/> SWISSPROT:AAKC_RAT	Q9QZH4	5'-AMP-activated protein kinase, beta-2 subunit (AMPK beta-2 chain).	271
<input type="checkbox"/> SWISSPROT:AAKG_BOVIN	P58108	5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain) (AMPKg).	330
<input type="checkbox"/> SWISSPROT:AAKG_PIG	Q09138	5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain) (AMPKg) (38 kDa subunit) (Fragments).	133
<input type="checkbox"/> SWISSPROT:AAKG_RAT	P80385	5'-AMP-activated protein kinase, gamma-1 subunit (AMPK gamma-1 chain) (AMPKg).	330
<input type="checkbox"/> SWISSPROT:AAKI_HUMAN	Q9UG19	5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain) (AMPK gamma3).	464
<input type="checkbox"/> SWISSPROT:AAKI_PIG	Q9MYP4	5'-AMP-activated protein kinase, gamma-3 subunit (AMPK gamma-3 chain) (AMPK gamma3).	464
<input type="checkbox"/> SWISSPROT:ACEK_ECOLI	P11071	Isocitrate dehydrogenase kinase/phosphatase (IDH kinase/phosphatase) (EC 2.7.1.116) (EC 3.1.3.-).	578
<input type="checkbox"/> SWISSPROT:AFC2_ARATH	P51567	Protein kinase AFC2 (EC 2.7.1.-).	427
<input type="checkbox"/> SWISSPROT:AFC3_ARATH	P51568	Protein kinase AFC3 (EC 2.7.1.-).	400
<input type="checkbox"/> SWISSPROT:AFSK_STRGR	P54742	Serine/threonine protein kinase afsK (EC 2.7.1.-).	807

Figure 3.21 Query Result page, showing the results from a search of the SWISS-PROT **Description** field, for the word **kinase**.

2. Choose the **Organism Name** field from the **Sort Results by:** drop-down list.

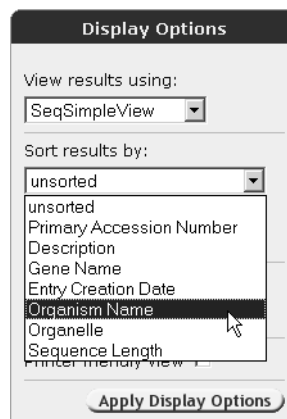


Figure 3.22 Display Options area showing drop-down list of available sort fields.

Note: Entries for which the chosen sort field is missing or null will be listed at the bottom of a sorted list whether sorting was in ascending or descending order.

3. For this example, choose to sort in an **ascending** order.

Note: Unsorted results will be listed in **ascending** order. It is not possible to reverse this order for unsorted results.

4. You might also like to apply a different view to the results, perhaps one that shows the field on which you are sorting. For this example choose **SwissView**.

Display Options

View results using:

Sort results by:

ascending
 descending

Show results
per page

Printer friendly view

[Apply Display Options](#)

Figure 3.23 Choose the order for sorting and a view to display the results.

5. Initiate sorting by clicking on the [Apply Display Options](#) button. The result of the above sort is shown in Figure 3.24.

Query "[swissprot-Description:kinase*]" found 3741 entries [next](#)

SWISSPROT	Accession	Description	GeneName	Organism	Authors
<input type="checkbox"/> SWISSPROT:ABL_MLVAB	P00521	Tyrosine-protein kinase transforming protein ABL (EC 2.7.1.112).	V-ABL	Abelson murine leukemia virus.	Reddy,E.P. Smith,M.J. Srinivasan,A. Groffen,J. Heisterkamp,N. Reynolds,F.H. Jr. Stephenson,J.R.
<input type="checkbox"/> SWISSPROT:KKA6_ACIBA	P09885	Aminoglycoside 3'-phosphotransferase (EC 2.7.1.95) (Kanamycin kinase, type VI) (Neomycin-kanamycin phosphotransferase, type VI) (APH(3')VI).	APHA-6	Acinetobacter baumannii.	Martin,P. Jullien,E. Courvalin,P.
<input type="checkbox"/> SWISSPROT:PPK_ACIBA	Q9X4M8	Polyphosphate kinase (EC 2.7.4.1) (Polyphosphoric acid kinase) (ATP-polyphosphate phosphotransferase).	PPK	Acinetobacter baumannii.	Gavigan,J.A. Marshall,L.M. Dobson,A.D.W.
<input type="checkbox"/> SWISSPROT:PTK_ACIJO	O52788	Tyrosine-protein kinase ptk (EC 2.7.1.112).	PTK	Acinetobacter johnsonii.	Grangeasse,C. Doublet,P. Vaganay,E. Vincent,C. Deleage,G. Duclos,B. Cozzone,A.J. Riberty,M.

Figure 3.24 Query Result page, showing the results from a search of the SWISS-PROT **Description** field, for the word "kinase", sorted according to the **Organism Name** field and displayed using the **SwissView** view.

3.7 Browse Index

3.7.1 About Browsing Indices

In addition to the other search methods, you can also browse the indices for a search term.

3.7.2 Browsing Indices

You can browse indices from the **Field Information** page for the datafield type (e.g., **Description**, **SeqLength**) that interests you. See section 3.7.3 "Getting to the Field Information Page" (p. 88) for more details of how to get to this page.

The screenshot shows the LION SRS interface. At the top, there is a navigation bar with the LION logo and a Help Center link. Below this is a menu with options: Quick Searches, Select Databanks, Query Form, Tools, Results, Projects, Custom Views, and Information. The main content area is titled 'SRS' and 'Field Name: Description'. A description text states: 'This is probably the best data field for searching an entry you don't know very much about; however, you can't expect to find all entries of a class, since often different conventions are used for naming enzymes, organisms, genes, etc..'. Below this is a table titled 'Data-fields in SRS' with columns: Databank, Name, Print Name, Short Name, Type, No. of Keys, No. of Entry References, Indexing Date, and Status. The table contains one row for SWISSPROT. At the bottom, there is a 'Browse Index' section with a search form: 'List values that match * [input] and occur in at least [input] entries' followed by a 'List Values' button.

Databank	Name	Print Name	Short Name	Type	No. of Keys	No. of Entry References	Indexing Date	Status
SWISSPROT	Description	Description	des	index	0	0		see member databank

Figure 3.25 Field Information page for the **Description** field index for the SWISS-PROT databank.

1. On the **Field Information** page for the appropriate databank and datafield, enter your search term, using wildcards as appropriate.

Note: In contrast to other query methods, implicit wildcards are not appended automatically to the search term when browsing indices, and must be specified explicitly.

2. Click the **List Values** button.
This will take you to the **Browse Index** page.

The screenshot shows the LION SRS interface. At the top, there is a navigation bar with the LION logo and a Help Center icon. Below the navigation bar are several tabs: Quick Searches, Select Databanks, Query Form, Tools, Results, Projects, Custom Views, and Information. The main content area is titled 'SRS' and includes a 'Reset' button. On the left, there is a 'Browse Options' sidebar with 'More Values' and 'Make Query' buttons. The main area displays two tables of search results:

Values in SWISSNEW	
Value	No of Entries
<input type="checkbox"/> serca1	4
<input type="checkbox"/> serca2	8
<input type="checkbox"/> serca3	3

Values in SWISSPROTRELEASE	
Value	No of Entries
<input type="checkbox"/> serca1	6
<input type="checkbox"/> serca2	8
<input type="checkbox"/> serca3	5

Figure 3.26 Browsing the **Description** field index on the **Browse Index** page.

3. Tick the check boxes beside the resultant terms that best match your needs and click the **Make Query** button. This returns entries that match the selected terms.

3.7.3 Getting to the Field Information Page

There are a number of routes you can use to get to the **Field Information** page. The common routes are detailed here.

From the Query Forms

On the **Standard Query Form** there is an **i** icon beside each field box. Choose the field of interest from the drop-down list, and click on the **i** icon.





Fields you can search	Your search terms
In a single field, you can separate multiple values by &, , !	
 Organism Name	<input type="text"/>
 AllText	<input type="text"/>
 AllText	<input type="text"/>
 AllText	<input type="text"/>

Figure 3.27 Standard Query Form showing use of the  icon to access the **Field Information** page.

On the **Extended Query Form** each of the datafields is a hyperlink to the **Field Information** page. Simply click on the hyperlink for the field of interest.

Fields you can search	Your search terms	Create a view
In a single field, you can separate multiple values by &, , !		
AllText	<input type="text"/>	<input type="checkbox"/>
ID	<input type="text"/>	<input type="checkbox"/>
Accession Number	<input type="text"/>	<input type="checkbox"/>
Primary Accession Number	<input type="text"/>	<input type="checkbox"/>
Description	<input type="text"/>	<input type="checkbox"/>
Gene Name	<input type="text"/>	<input type="checkbox"/>
Keywords	<input type="text"/>	<input type="checkbox"/>
Entry Creation Date	select 1 Jan YYYY 1 Jan YYYY	<input type="checkbox"/>
LastSequenceUpdate	select 1 Jan YYYY 1 Jan YYYY	<input type="checkbox"/>
LastAnnotationUpdate	select 1 Jan YYYY 1 Jan YYYY	<input type="checkbox"/>

Figure 3.28 Extended Query Form showing hyperlinks for datafields.

From the List Of Databanks Page

The **Information** tab takes you to the **List of Databanks** page. Here you will see a list of the available databanks. Clicking on the hyperlink for a databank will take you to its

Databank Information page. This contains a section listing the datafields for the relevant databank.

Data-fields in SRS							
Field Name	Short Name	Type	No. of Keys	No. of Entry References	Indexing Date	Status	
AllText	all	group	0	0		not indexed	
ID	id	id	101602	101602	25-Feb-2003	ok	
Accession Number	acc	index	119193	119300	25-Feb-2003	ok	
Primary Accession Number	pac	index	101602	101602	25-Feb-2003	ok	
Description	des	index	58091	578507	25-Feb-2003	ok	
Gene Name	gen	index	83823	145372	25-Feb-2003	ok	
Keywords	key	index	840	370433	25-Feb-2003	ok	
Entry Creation Date	crd	date	40	101602	25-Feb-2003	ok	
LastSequenceUpdate	lsu	date	40	101602	25-Feb-2003	ok	
LastAnnotationUpdate	lau	date	39	101602	25-Feb-2003	ok	
Organism Name	org	index	10800	236202	25-Feb-2003	ok	
Taxon	tax	index	4834	940067	25-Feb-2003	ok	
NCBI TaxId	txi	num	7101	104471	25-Feb-2003	ok	
Organelle	ogn	index	559	11670	25-Feb-2003	ok	
ProteinID	prd	index	160681	163655	25-Feb-2003	ok	
Checksum	cks	index	101503	101602	25-Feb-2003	ok	

Figure 3.29 Part of the **Databank Information** page for the SWISSPROTRELEASE databank showing datafield hyperlinks.

Click on one of the datafield hyperlinks (e.g., **Description**) to go to its **Field Information** page (see Figure 3.25).

Note: SWISSPROT is a virtual databank. This means it is made up from other databanks, e.g. SWISSPROT (Release) and SWISSPROT (Updates). If you look at the Databank Information page for SWISSPROT you will not see all the information that is shown in Figure 3.29, if you want all the detail, you need to look at the individual component databanks.

LINKS TO ADDITIONAL DATA

After you have queried a databank using SRS, you may find that you want to refine your search further. For example, you may want to find further information on the entries returned by your initial search, e.g., by looking for related entries in another databank. Alternatively you may have queried a databank that returned several entries relating to your query, but you only want to know about the entries that relate to a specific protein. During this chapter you will learn:

- How to use SRS links to find entries which are related to each other in more than one databank.

4.1 What is a Link?

A databank entry may contain references to other databanks, and vice versa. In SRS these relationships are known as links and can be used to extend a query across multiple databanks. Thus you can obtain all the entries in one databank that are linked to an entry (or entries) in another databank.

From a user perspective there are two types of link: hypertext links and index links (query links).

Hypertext links are links between entries which are displayed as hypertext. These are hardcoded into SRS and you can use them whenever you wish. They are useful for examining entries that are referenced directly from entries.

Index links are built into the SRS indices at the same time as databanks are added. They allow you to construct queries using relationships between databanks. They require SRS to search through entries or indices in other databanks, looking for matches.

It is assumed that you are already familiar with hypertext links, so only a limited demonstration of them is given in this chapter (see section 4.2 “Hypertext Links”, p. 92). The remainder of this chapter is devoted to explaining index links.

4.2 Hypertext Links

Some entries have hypertext links which allow you to access linked entries directly. Example 4.1 shows such an entry.

Example 4.1 Accessing a hypertext linked entry.

1. Query EMBL for entries whose description field contains the word “kinase”.

EMBL entries will be displayed on the **Query Result** page.

The screenshot shows the LION SRS interface. At the top, there is a navigation bar with 'LION' logo, 'Help Center', and tabs for 'Quick Searches', 'Select Databanks', 'Query Form', 'Tools', 'Results', 'Projects', 'Custom Views', and 'Information'. Below this is the 'SRS' header and a search bar containing the query "[embl-Description:kinase*]" with 73271 entries found. The main content area is divided into three panels on the left and a table of results on the right.

Apply Options to:

- selected results only
- unselected results only

Result Options

Launch analysis tool: BlastN **Launch**

Show tools relevant to these results: **Tools**

Link to related information: **Link**

Save results: **Save**

Display Options

View results using: SeqSimpleView

Sort results by: Accession Number

- ascending
- descending

Show 30 results per page

Printer friendly view

Apply Display Options

EMBL	Accession	Description	SeqLength
<input type="checkbox"/> EMBL:A00229	A00229	Artificial sequence for kinased linker R140	24
<input type="checkbox"/> EMBL:A01166	A01166	Artificial sequence derived from yeast phosphoglycerate kinase (PGK) gene	1207
<input type="checkbox"/> EMBL:A01168	A01168	Artificial sequence for kinased linker R140	24
<input type="checkbox"/> EMBL:A15367	A15367	phosphoglycerate kinase gene	425
<input type="checkbox"/> EMBL:A16530	A16530	Synthetic pyruvate kinase gene (SEQ ID NO: 1).	3602
<input type="checkbox"/> EMBL:A16753	A16753	tie receptor tyrosine kinase seq ID No 3	3845
<input type="checkbox"/> EMBL:A16754	A16754	tie receptor tyrosine kinase (alternatively spliced to lack one of the epidermal growth factor homology domains) seq ID No 4	3713
<input type="checkbox"/> EMBL:A21640	A21640	protein kinase gene	1403
<input type="checkbox"/> EMBL:A22127	A22127	partial protein kinase protein of ILTV	656
<input type="checkbox"/> EMBL:A22428	A22428	Kinased linker P140	24
<input type="checkbox"/> EMBL:A26374	A26374	M.musculus HSV thymidine kinase 5' non coding region	60
<input type="checkbox"/> EMBL:A27171	A27171	Tyrosine kinase receptor gene	2966
<input type="checkbox"/> EMBL:A29799	A29799	S.cerevisiae pyruvate kinase gene 5'terminus	49
<input type="checkbox"/> EMBL:A29801	A29801	Pyruvate kinase gene altered	49
<input type="checkbox"/> EMBL:A32502	A32502	Synthetic NDP kinase protein primer	17
<input type="checkbox"/> EMBL:A32505	A32505	D.discoideum NDP kinase gene	530
<input type="checkbox"/> EMBL:A32507	A32507	Synthetic N-terminal fusion NDP kinase gene	184
<input type="checkbox"/> EMBL:MMA00144	AA000144	mg31a09.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone IMAGE:425368 5' similar to gb:M74149 Mouse creatine kinase B gene, complete cds (MOUSE);, mRNA sequence.	600
<input type="checkbox"/> EMBL:MMA00177	AA000177	mg33h05.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone IMAGE:425625 5' similar to gb:Z49877 M.musculus syk mRNA for protein-tyrosine kinase (MOUSE);, mRNA sequence.	474
<input type="checkbox"/> EMBL:MMA00178		mg33h08.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone IMAGE:425631 5' similar to gb:U07358	

Figure 4.1 The Query Result page, with results sorted by Accession Number (see section 3.6 “Sorting Results”, p. 82).

Linked data in the entry:

- Click on the hypertext link beside an entry to view the entire entry.

The screenshot shows the LION interface for an EMBL entry. The top navigation bar includes 'Quick Searches', 'Select Databanks', 'Query Form', 'Tools', 'Results', 'Projects', 'Custom Views', and 'Information'. Below this, there are links for 'Text Entry' and 'EmblEntry'. A 'Reset' button is located on the left side.

The main content area is divided into several sections:

- Entry Information:** Shows 'Entry from: EMBL'.
- Entry Options:** Includes a 'Launch analysis tool:' section with a dropdown menu set to 'BlastN' and a 'Launch' button. Below it, there is a 'Link to related information:' section with a 'Link' button, and a 'Save entry:' section with a 'Save' button. At the bottom, there is a 'View:' section with a 'Printer Friendly' button.
- Go to:** A horizontal list of links: 'General', 'Description', 'References', 'Cross-references', 'Features', and 'Sequence'.
- General Information:** A table with the following data:

Entry Name	EMBL:A21640
Molecule Type	DNA
Sequence Length	1403
Entry Division	VRL
Accession #	A21640;
Sequence Version	A21640.1
Creation Date	22-JUL-1994
Modification Date	22-JUL-1994
- Description:** A table with the following data:

Description	protein kinase gene
Keywords	
Organism	Pseudorabies virus
Organism Classification	Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Varicellovirus.
Organelle	
- References:** A list of references, with the first one being:
 - MUTANT PSEUDORABIES VIRUS, AND VACCINES CONTAINING THE SAME** Patent number WO9102795-A/3, 07-MAR-1991. Position 1-1403
- Database Cross-references:** A table with the following data:

GOA	P24381; P24381.
SWISS-PROT	P24381; KR1_PRVN3.

Figure 4.2 The **Entry** page for the EMBL entry with accession number A21640.

- Scroll through the entry until you find a hyperlink (e.g. in the entry shown in Figure 4.2 there is a hypertext link to **SWISS-PROT**). This will have a unique identifier (accession number) — “P24381” in figures 4.2 and 4.3.

Database Cross-references	
GOA	P24381 ; P24381 .
SWISS-PROT	P24381 ; KR1_PRVN3.

Figure 4.3 Database Cross-references entry, showing a hypertext link to **SWISS-PROT**.

If you do not find a link in your chosen entry, try choosing a different entry from the **Query Result** page.

4. Click the hypertext link to view the **Entry** page for the SWISS-PROT entry.

4.3 Index Links

4.3.1 Searching for Data using Index Links

It is assumed that you have arrived at the **Link** page (Figure 4.4). Use section 4.4 “Getting to the Link Page” (p. 99) to help you navigate here if necessary.

LION Help Center ?

Quick Searches | Select Databanks | Query Form | Tools | Results | Projects | Custom Views | Information

SRS

Reset Find entries related to current query: "[EMBL-ID:A21640]" in other databanks

Link Options

Select the **databanks** you want to search for **related information**

Find **related** entries

Refine Query - **show** only **results** with **related entries**

Show only results **without** related entries

Search

Databanks Available to Link to

Expand all Collapse all

To Parent Entry

EMBOSS Results

Sequence databanks - complete

all EMBL SWISSPROT

Sequence databanks - subsections

SeqRelated

Tool Results

Display Options

Show results per page

Figure 4.4 The **Link** page for linking initiated from the **Manage your Query Results** and **Query Result** pages.

Note: There are two forms of the **Link** page. If you have initiated the links from either the **Manage your Query Results** or **Query Result** pages then you will see the **Link** page as shown in Figure 4.4. If you have initiated linking from the **Entry** page then the **Link Options** are not available (see Figure 4.5).

Figure 4.5 The **Link** page for linking initiated from the **Entry** page.

1. Choose the databank in which you wish to search for links, by ticking the check box to the left of it.
2. Choose the appropriate **Link Option** (see section 4.3.2 “Link Options” p. 98).
3. Click the **Search** button to display the results.

EMBL	Accession	Description	SeqLength
<input checked="" type="checkbox"/> EMBL:A21640	A21640	protein kinase gene	1403

Figure 4.6 Typical **Query Result** page for a linking request.

4.3.2 Link Options

Unless you have initiated your search for links from the **Entry** page there are several options for choosing the set of entries for which you wish to search. These are listed on the **Link** page in the **Link Options** box.

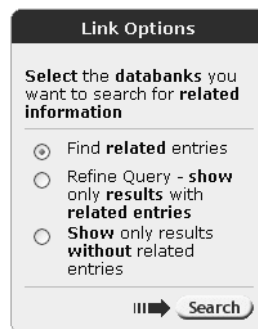


Figure 4.7 Link Options.

The meanings for these options are as follows:

Find related entries.

This returns entries from other databanks which have links with entries in the current query.

Refine Query - show only results with related entries.

This limits the query so that it includes only the entries (from the original query) which are linked to all of the selected databanks.

Show only results without related entries.


This limits the query so that it includes only the entries (from the original query) which do not have links to the specified databanks.

4.4 Getting to the Link Page

It is possible to initiate a linking operation from many pages. These include the **Manage your Query Results** page (see section 4.4.1 “Finding Links from the Manage your Query Results Page” p. 99), the **Query Result** page (see section 4.4.2 “Finding Links from the Query Result Page” p. 101) and the **Entry** page (see section 4.4.3 “Finding Links from the Entry Page” p. 103).

4.4.1 Finding Links from the Manage your Query Results Page

This method is useful for finding links for complete queries. You can search for links from a single query or from multiple queries.

There are two ways to find links for your queries from the **Manage your Query Results** page. The first and most common method is to tick the check box that corresponds to a query set and click the  button.

Results Options

Options are **applied** to **selected** queries only

Save results:

Delete results:

Combine queries with: & (AND)

Find related info:

Results Display Options

View results using: * Complete entries *

Show results per page

Result History

Name	Type	Total No.	From	No.	Query Expression	Comment
<input type="checkbox"/> Q8	select	1	EMBL	1	[EMBL-ID:A21640]	<input type="text"/>
<input type="checkbox"/> Q7	select	1	EMBL	1	[EMBL-ID:A21640]	<input type="text"/>
<input type="checkbox"/> Q6	select	1	EMBL	1	[EMBL-ID:A21640]	<input type="text"/>
<input type="checkbox"/> Q5	query	1	EMBL	1	[embl-ALLTEXT:a21640]	<input type="text"/>
<input type="checkbox"/> Q4	query	73271	EMBL	73441	[embl-Description:ki]	<input type="text"/>
<input type="checkbox"/> Q3	query	50	BLASTX	50	[BLASTX-JobName:temp]	<input type="text"/>
<input type="checkbox"/> Q2	workflow	3	EMBL	3	[EMBL-alltext:mnkx*]	<input type="text"/>
<input type="checkbox"/> Q1	workflow	586050	EMBL	586050	[EMBL-alltext:kinase]	<input type="text"/>


Figure 4.8 The **Manage your Query Results** page.

This takes you to the **Link** page (see section 4.3 “Index Links” p. 95) from where you can complete your search for linked data.

The second option is to use the Query Expression text box at the top of the **Manage your Query Results** page. This option is discussed later (see section 4.5 “Expression Linking” p. 104).

4.4.2 Finding Links from the Query Result Page

This method is useful for searching for links from individual entries or groups of entries.

It is assumed that you have made your query as normal. From the **Query Result** page, choose those entries which are of interest and click the  button to display the **Link** page (see section 4.3 “Index Links” p. 95).

Note: You can choose whether to search for links to your results based only on the selected entries/results (**selected only**) or based only on the entries/results that are not selected (**unselected only**) items, by setting the **Apply Options To:** options. This is particularly useful if you wish to find links for a large proportion of your results. The **Apply Options To:** options are usually set to **unselected results only**.

Example 4.2 Initiating a link from the Query Result page.

1. Query EMBL for entries whose description field contains the word “kinase”.
EMBL entries will be displayed on the **Query Result** page.

LION Help Center ?

Quick Searches Select Databanks Query Form Tools Results Projects Custom Views Information

SRS

Reset Query "[embl-ALLTEXT:kinase*]" found 586050 entries next

Apply Options to:

selected results only

unselected results only

Result Options

Launch analysis tool:

BlastN

Show tools relevant to these results:

Link to related information:

Save results:

Display Options

View results using:

SeqSimpleView

Sort results by:

unsorted

ascending

descending

Show 30 results per page

Printer friendly view

EMBL	Accession	Description	SeqLength
<input type="checkbox"/> EMBL:AA901535	AA901535	NCM5B1T3 Mycelial Neurospora crassa cDNA clone NM5B1 5' end, mRNA sequence.	347
<input type="checkbox"/> EMBL:AA901632	AA901632	NCM4G1T3 Mycelial Neurospora crassa cDNA clone NM4G1 5' end, mRNA sequence.	536
<input type="checkbox"/> EMBL:AA901739	AA901739	NCP5C11T7 Perithecial Neurospora crassa cDNA clone NP5C11 3' end.	473
<input type="checkbox"/> EMBL:AA901740	AA901740	NCM6F2T3 Mycelial Neurospora crassa cDNA clone NM6F2 5' end, mRNA sequence.	442
<input type="checkbox"/> EMBL:AA901762	AA901762	NCM1C7T7 Mycelial Neurospora crassa cDNA clone NM1C7 3' end, mRNA sequence.	580
<input type="checkbox"/> EMBL:AA901817	AA901817	NCC1G8T7 Conidial Neurospora crassa cDNA clone NC1G8 3' end, mRNA sequence.	477
<input type="checkbox"/> EMBL:AA901858	AA901858	NCC1E9T7 Conidial Neurospora crassa cDNA clone NC1E9 3' end, mRNA sequence.	486
<input type="checkbox"/> EMBL:AA901859	AA901859	NCC2D1T7 Conidial Neurospora crassa cDNA clone NC2D1 3' end, mRNA sequence.	609
<input type="checkbox"/> EMBL:AA901875	AA901875	NCP4B7T7 Perithecial Neurospora crassa cDNA clone NP4B7 3' end.	555
<input type="checkbox"/> EMBL:AA901876	AA901876	NCC4C8T7 Conidial Neurospora crassa cDNA clone NC4C8 3' end, mRNA sequence.	583
<input type="checkbox"/> EMBL:AA901892	AA901892	NCC1A1T7 Conidial Neurospora crassa cDNA clone NC1A1 3' end, mRNA sequence.	493
<input type="checkbox"/> EMBL:AA901893	AA901893	NCC1D8T7 Conidial Neurospora crassa cDNA clone NC1D8 3' end, mRNA sequence.	514
<input type="checkbox"/> EMBL:AA901894	AA901894	NCC1F2T7 Conidial Neurospora crassa cDNA clone NC1F2 3' end, mRNA sequence.	530
<input type="checkbox"/> EMBL:AA901895	AA901895	NCC2H6T7 Conidial Neurospora crassa cDNA clone NC2H6 3' end, mRNA sequence.	524
<input type="checkbox"/> EMBL:AA901896	AA901896	NCC3D6T7 Conidial Neurospora crassa cDNA clone NC3D6 3' end, mRNA sequence.	494
<input type="checkbox"/> EMBL:AA901987	AA901987	NCC4E1T7 Conidial Neurospora crassa cDNA clone NC4E1 3' end, mRNA sequence.	548
<input type="checkbox"/> EMBL:AF056142	AF056142	AF056142 rice blast fungus mRNA from nitrogen-starved mycelial culture Pyricularia grisea cDNA clone PU40 similar to protein kinase C, mRNA sequence.	647
<input type="checkbox"/> EMBL:AI391970	AI391970	NCC1F3T3 Conidial Neurospora crassa cDNA clone NC1F3 5' similar to casein kinase II, alpha catalytic subunit (a serine/threonine-protein kinase), mRNA sequence.	552

Figure 4.9 The Query Result page.

2. Select the entries of interest, by ticking the check boxes beside them.
3. Click on the button. This will display the **Link** page.
(see section 4.3 "Index Links" p. 95).

4.4.3 Finding Links from the Entry Page

This method is useful for linking from individual entries.

It is assumed that you have made your query as normal. When the **Query Result** page appears, choose one of the entries and click on its hyperlink to display the **Entry** page for that entry. Note that you can usually view such a page using **Text Entry** format or in a databank-specific format, e.g., **SwissEntry** format.

The screenshot shows the LION Bioscience interface. At the top, there is a navigation bar with the LION logo and a Help Center link. Below the navigation bar, there are tabs for Quick Searches, Select Databanks, Query Form, Tools, Results, Projects, Custom Views, and Information. The current page is titled "Text Entry | EmblEntry".

The main content area displays "Entry 1 of 586050 from Query 9" with a "Next Entry" button. On the left side, there are two panels: "Entry Information" and "Entry Options".

Entry Information:

- Entry from: [EMBL](#)

Entry Options:


- Launch analysis tool: BlastN
- Link to related information:
- Save entry:
- View:

Entry Details:

```

ID AA901535 standard; RNA; EST; 347 BP.
XX
AC AA901535;
XX
SV AA901535.1
XX
DT 15-APR-1998 (Rel. 55, Created)
DT 11-JAN-2001 (Rel. 66, Last updated, Version 4)
XX
DE NCM5B1T3 Mycelial Neurospora crassa cDNA clone NM5B1 5' end, mRNA sequence.
XX
KW EST.
XX
OS Neurospora crassa
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariales;
OC Sordariaceae; Neurospora.
XX
RN [1]
RP 1-347
RX MEDLINE; 97435549.
RA Nelson M.A., Kang S., Braun E.L., Crawford M.E., Dolan P.L., Leonard P.M.,
RA Mitchell J., Armijo A.M., Bean L., Blueyes E., Cushing T., Errett A.,
RA Fleharty M., Gorman M., Judson K., Miller R., Ortega J., Pavlova I.,
RA Perea J., Todisco S., Trujillo R., Valentine J., Wells A.,
RA Werner-Washburne M., Yazzie S., Natvig D.O.;
RT "Expressed sequences from conidial, mycelial, and sexual stages of
RT Neurospora crassa";
RL Fungal Genet. Biol. 21:348-363 (1997).
  
```

Figure 4.10 A typical **Entry** page shown using **Text Entry** format.

Click the  button. This will display the **Link** page for single entry linking (see section 4.3 “Index Links” p. 95).

4.5 Expression Linking

The Query Expression text box (e.g., on the **Manage your Query Results** page) is a useful alternative to the linking methods discussed earlier in this chapter. It can be used to search for a link between two or more sets of results or between a set of results and a databank.

Note: Using the Query Expression text box allows you to search for links without going through the **Link** page.

4.5.1 Expression Linking Procedure

You might be interested in the entries from a set of DNA sequences (in **Q1**) that are linked to a set of protein sequences (in **Q2**).

If there are links between entries in the queries, you have the choice of returning the entries in **Q1** that have links with **Q2** (which is what is wanted here) or of returning the entries in **Q2** that have links with **Q1**. (See also the section 4.5.2 “Linking Operators” p. 105.)

The Query Expression text box is shown in Figure 4.11. You should enter your expression (e.g., “Q1 < Q2”) as shown.

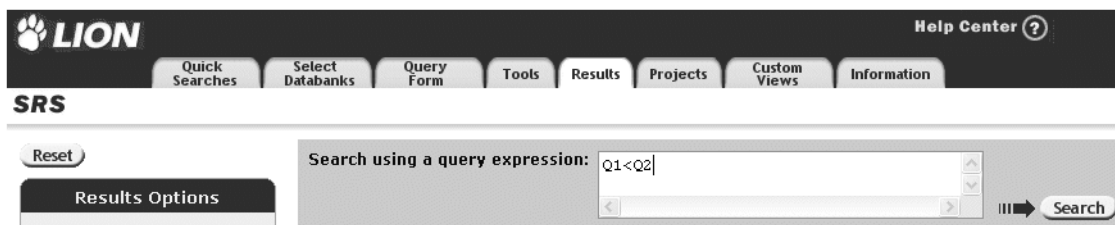


Figure 4.11 Expression linking.

To link complete query sets using expression linking:

1. Type in the name of the query set to be linked (“Q1”), the type of query (e.g., “<” or “>”, see section 4.5.2, “Linking Operators”, p. 105, for more details) and the databank or second query (“Q2”) to which the set should link. For instance:

Q1 < Q2

will list all the entries in **Q1** that are linked to **Q2**.

Q1 > Q2

will list all the entries in **Q2** that are linked to **Q1**. See the section 4.5.2 “Linking Operators” (p. 105).

2. Click the **Search** button.

4.5.2 Linking Operators

Expression linking uses linking operators to describe the nature of the links between the specified queries or

databanks. Typical operators are the < and > which indicate a link:

<

Entries in the set or databank to the left of the operator are returned if they have a link to any entries in the set or databank to the right of the operator.

>

Entries in the set or databank to the right of the operator are returned if they have a link to any entries in the set or databank to the left of the operator.

You can combine linking (< and >) and logical (&, | and !) operators to build up more complex queries (see the examples in Table 4.1).

Table 4.1 Linking operations.

Operators	Example	Returns Entries in:
<	Q1 < Q2	In Q1 that link to Q2.
>	Q1 > Q2	In Q2 that link to Q1.
< &	Q1 < Q2 & Q3	In Q1 that link to Q2 <u>and</u> Q3.
<	Q1 < Q2 Q3	In Q1 that link to Q2 <u>or</u> Q3.
< !	Q1 < Q2 ! Q3	In Q1 that link to Q2 <u>but not</u> Q3.

4.5.3 Expression Linking Examples

If you have a set of EMBL entries in a query, **Q3**, which you wish to search for links with the SWISS-PROT databank, type:


```
Q3 < swissprot
```



This will show the EMBL entries from the original query that have links to the SWISS-PROT databank.

If you would rather see the SWISS-PROT entries that the above operation returned, turn the linking operator around so that it points towards SWISS-PROT:

```
Q3 > swissprot
```



This returns the SWISS-PROT entries that have links with the entries in **Q3**.

CHAPTER

5

VIEWS

SRS includes a number of predefined views that you can use to change the way in which your results are displayed. In addition, you can create your own views. During this chapter you will learn:

- About the predefined views available in SRS.
- How to work with views.
- How to create and use your own views.

5.1 What is a View?

SRS uses views to determine which data types are shown. A view may include only a few data types, or it may display all available types of data.

In addition to the predefined views which are part of SRS you can create your own views. A view can be defined to work with a specific databank, or with several databanks. The data in a view could be from one or more databanks.

For convenience, views can be categorized as “predefined views” and “project views” (see section 5.3 “Creating Views”, p. 123). Predefined views are the default views that are available to everybody using an SRS server. Some of these are designed to work with specific databanks and will only be available when they are appropriate. Project views are views that you define, using the **View Manager** page (see section 5.3.2 “Creating Views using the View Manager Pages”, p. 129).

5.2 Using Views

Views can be applied from any page that creates or holds a query. The following list identifies the pages where views can be created or applied to data sets.

- The **Standard Query** or **Extended Query Forms**.
- The **Query Result** page (which contains the results of a query).

- The **Entry** page (which contains the details of a single entry).
- The **Manage your Query Results** page.
- The **Launch** page (from which you can launch tools).

The views drop-down list (available on pages such as the **Manage your Query Results** page) contains the views that are currently available (see Figure 5.1). These will change, according to factors such as which views are installed on your system, which databank you are using and what, if any, views you have created in the current project.

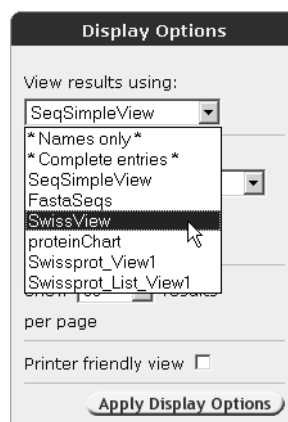


Figure 5.1 Typical drop-down list of available views.

5.2.1 How to Apply Views

This section gives instructions for applying views to your query results. In the first example, you will begin by querying a databank and setting the query options for views so that the query results are displayed with your chosen view. The other

examples show how to apply your view to an already existing set of results.

Applying a View From the Standard Query or Extended Query Forms

Whenever you query a databank using either of the query forms you can also choose (from a drop-down list) a view to be used to display the query results. This example shows you how to apply a view from the **Standard Query Form**. On the **Extended Query Form** the contents of the **Result Display Options** box are slightly different, but the basic steps are similar

To apply a view to a SWISSPROT query from the **Standard Query Form**:

1. On the **Select Databanks To Search** page, select the SWISSPROT databank using the check box beside it and click the **Standard Query Form** button. (See chapter 3, "Querying with SRS", p. 55, for more information.)
2. Enter **oxidase** in the text box. The default **AllText** datafield is fine.
3. In the **Result Display Options** box, select a view from the **View results using:** drop-down list (see Figure 5.2).

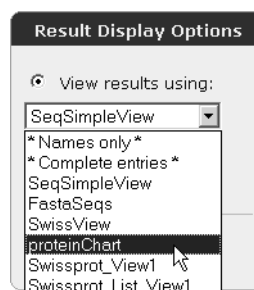


Figure 5.2 Choosing a predefined view for results from the Standard Query Form.

4. Click on the **Search** button.
Complete any further information, e.g. if a separate window pops up.

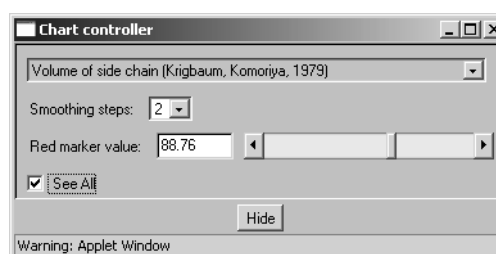


Figure 5.3 Chart controller window for proteinChart view.

5. This will display the **Query Result** page with the results of your query displayed using the selected view (see Figure 5.4).

LION Help Center ?

Quick Searches Select Databanks Query Form Tools Results Projects Custom Views Information

SRS

Reset Query "[swissprot-AllText:oxidase*]" found 1728 entries next

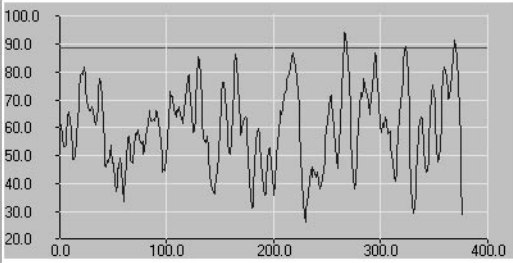
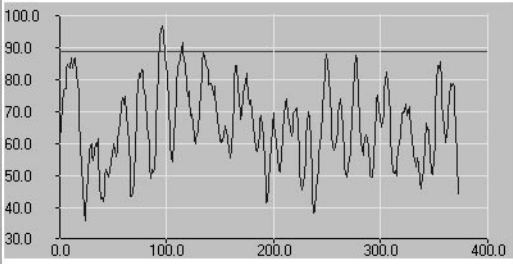
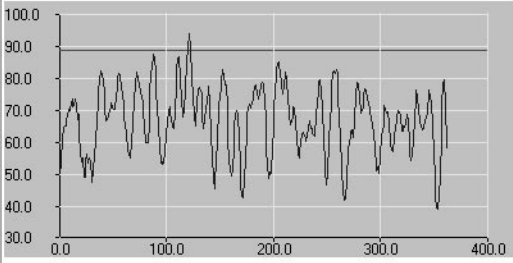
Apply Options to:	SWISSPROT	Description	ProtSequence
<input type="radio"/> selected results only <input checked="" type="radio"/> unselected results only	<input type="checkbox"/> SWISSPROT:2NPD_NEUCR	2-nitropropane dioxygenase precursor (EC 1.13.11.32) (Nitroalkane oxidase) (2-NPD).	
Result Options Launch analysis tool: <input type="text" value="BlastP"/> <input type="button" value="Launch"/> Show tools relevant to these results: <input type="button" value="Tools"/> Link to related information: <input type="button" value="Link"/> Save results: <input type="button" value="Save"/>	<input type="checkbox"/> SWISSPROT:2NPD_WILMR	2-nitropropane dioxygenase (EC 1.13.11.32) (Nitroalkane oxidase) (2-NPD).	
Display Options View results using: <input type="text" value="proteinChart"/> Sort results by: <input type="text" value="unsorted"/> <input checked="" type="radio"/> ascending <input type="radio"/> descending Show <input type="text" value="30"/> results per page Printer friendly view <input type="checkbox"/> <input type="button" value="Apply Display Options"/>	<input type="checkbox"/> SWISSPROT:ACC3_LYCES	1-aminocyclopropane-1-carboxylate oxidase homolog (Protein E8).	

Figure 5.4 The Query Result page using the selected view.

From the Query Result Page

Querying, linking, and launching applications always returns the **Query Result** page. This page contains a list of the

entries that meet the requirements specified in your query. If you do not select a view for the results, when the operation is run, a default view is used.

To change the view on the **Query Result** page:

In the **Display Options** box you will see a drop-down list below the text **View results using:**, (see Figure 5.5). The drop-down list displays the name of the view currently applied.

The screenshot shows the LION SRS interface. At the top, there is a navigation bar with 'LION' logo and 'Help Center' link. Below it are tabs for 'Quick Searches', 'Select Databanks', 'Query Form', 'Tools', 'Results', 'Projects', 'Custom Views', and 'Information'. The main header reads 'SRS' and 'Query "[swissprot-AllText:oxidase*]" found 1728 entries'. On the left, there are three panels: 'Apply Options to:' (selected results only/unselected results only), 'Result Options' (Launch analysis tool: BlastP, Show tools relevant to these results: Tools, Link to related information: Link, Save results: Save), and 'Display Options' (View results using: proteinChart, *Names only*, *Complete entries*, SeqSimpleView, FastaSegs, SwissView, proteinChart, Swissprot_View1, Swissprot_List_View1, per page, Printer friendly view, Apply Display Options). The main content area is a table with three columns: SWISSPROT, Description, and ProtSequence. The table contains three entries, each with a checkbox, a SWISSPROT ID, a description, and a line graph showing the protein sequence.

SWISSPROT	Description	ProtSequence
<input type="checkbox"/> SWISSPROT:2NPD_NEUCR	2-nitropropane dioxygenase precursor (EC 1.13.11.32) (Nitroalkane oxidase) (2-NPD).	
<input type="checkbox"/> SWISSPROT:2NPD_WILMP	2-nitropropane dioxygenase (EC 1.13.11.32) (Nitroalkane oxidase) (2-NPD).	
<input type="checkbox"/> SWISSPROT:ACC3_LYCES	1-aminocyclopropane-1-carboxylate oxidase homolog (Protein E8).	

Figure 5.5 Selecting the view on the **Query Result** page.

1. Select the view you want to use from the drop-down list.
2. Click the **Apply Display Options** button.

The **Query Result** page will be refreshed so that it displays the results using the selected view.

Note: If you select a number of entries prior to clicking the **Apply Display Options** button then only those entries will be shown when you click **Apply Display Options**.

Applying a View From the Entry Page

For most of the **Entry** pages you will come across, you will have a simple choice of viewing the entry in a text format or a databank-specific format; in a small number of cases (where there is no databank entry format) you may have the option of using a view from the drop-down views menu. The use of the drop-down view menu is the same as on other SRS pages (see, for example the section “From the Query Result Page”, p. 114). The example below describes only the first method.

1. You need to be on an **Entry** page, so start with a **Query Result** page, and click the entry name hyperlink for one of the entries (see Figure 5.6).


SWISSPROT	Accession	Description	SeqLength
<input type="checkbox"/> SWISSPROT:2NPD_NEUCR	Q01284	2-nitropropane dioxygenase precursor (EC 1.13.11.32) (Nitroalkane oxidase) (2-NPD).	378
<input type="checkbox"/> SWISSPROT:2NPD_WILMB	Q12723	2-nitropropane dioxygenase (EC 1.13.11.32) (Nitroalkane oxidase) (2-NPD).	374
<input type="checkbox"/> SWISSPROT:ACC3_LYCES	P10967	1-aminocyclopropane-1-carboxylate oxidase homolog (Protein EB).	363
<input type="checkbox"/> SWISSPROT:A11M_YEAST	P03875	COX1/OXI3 intron 1 protein.	834
<input type="checkbox"/> SWISSPROT:A12M_YEAST	P03876	Putative COX1/OXI3 intron 2 protein.	789
<input type="checkbox"/> SWISSPROT:A13M_YEAST	P03877	Putative COX1/OXI3 intron 3 protein.	378

Figure 5.6 Click the hyperlink for an entry.

- Assuming that the current databank uses databank-specific and text formats, the **Entry** page will display the entry using the databank-specific format (see Figure 5.7).

The screenshot displays the LION Bioscience website interface. At the top, there is a navigation bar with the LION logo and a Help Center icon. Below this is a menu with options: Quick Searches, Select Databanks, Query Form, Tools, Results, Projects, Custom Views, and Information. The main content area shows the entry page for SWISSPROT:A11M_YEAST. The page is titled 'Text Entry | SwissEntry' and includes navigation links for 'Previous Entry' and 'Next Entry'. The entry information is presented in a structured format with sections for 'Entry Information', 'Entry Options', 'General Information about the Entry', 'Description and Origin of the Protein', and 'References'. The 'Entry Options' section includes a 'Launch analysis tool' dropdown set to 'BlastP' with a 'Launch' button, a 'Link to related information' button, a 'Save entry' button, and a 'Printer Friendly' view option. The 'General Information' section lists the entry name, accession number, creation date, and last update. The 'Description and Origin of the Protein' section provides keywords, a description, gene name, organism source, and classification. The 'References' section lists two scientific papers with links to Medline, PubMed, and their positions.

Entry Information

Entry from: 

Entry Options

Launch analysis tool:

Link to related information:

Save entry:

View:

General Information about the Entry

Entry name	SWISSPROT:A11M_YEAST
Prim. accession #	P03875
Created	Release 1, 21-JUL-1986
Last sequence update	Release 1, 21-JUL-1986
Last annotation update	Release 35, 1-NOV-1997

Description and Origin of the Protein

Keywords	Mitochondrion;
Description	cox1/oxi3 intron 1 protein.
Gene name(s)	ai1.
Organism source	Saccharomyces cerevisiae (Baker's yeast).
Organism classification	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NCBI Taxonomy ID	4932;
Organelle	Mitochondrion;

References

- Bonitz,S.G.; Coruzzi,G.; Thalenfeld,B.E.; Tzagoloff,A.; Macino,G.; **Assembly of the mitochondrial membrane system. Structure and nucleotide sequence of the gene coding for subunit 1 of yeast cytochrome oxidase.** J. Biol. Chem. 255:11927 (1980)
 Medline [81069885](#)
 PubMed [6254986](#)
 Position: sequence from n.a.
 Comment: strain=d273-10b;
- de Zamaroczy,M.; Bernardi,G.; **The primary structure of the mitochondrial genome of Saccharomyces cerevisiae -- a review.** Gene 47:155 (1986)
 Medline [87163488](#)
 PubMed [3549452](#)
 Position: sequence from n.a.

Figure 5.7 Entry page using **SwissEntry** format.

- The available viewing options are shown across the top of the page:

[Text Entry](#) | SwissEntry

- The option currently being displayed will be shown, with the alternative entry shown as a hyperlink. Click on the **Text Entry** hyperlink to view the entry using a text only format (see Figure 5.8).

LION Help Center ?

Quick Searches | Select Databanks | Query Form | Tools | Results | Projects | Custom Views | Information

Text Entry | [SwissEntry](#)

[Reset](#) [Previous Entry](#) Entry 6 of 33 from [Query 1](#) [Next Entry](#)

Entry Information

Entry from: 

Entry Options

Launch analysis tool:
 [Launch](#)

Link to related information:
[Link](#)

Save entry: [Save](#)

View: [Printer Friendly](#)

```

ID  AI1M YEAST  STANDARD;  PRT;  834 AA.
AC  P03875;
DT  21-JUL-1986 (Rel. 01, Created)
DT  21-JUL-1986 (Rel. 01, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last annotation update)
DE  COX1/OXI3 intron 1 protein.
GN  AI1.
OS  Saccharomyces cerevisiae (Baker's yeast).
OG  Mitochondrion.
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC  Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX  NCBI_TaxID=4932;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=D273-10B;
RX  MEDLINE=81069885; PubMed=6254986;
RA  Bonitz S.G., Coruzzi G., Thalenfeld B.E., Tzagoloff A., Macino G.;
RT  "Assembly of the mitochondrial membrane system. Structure and
RT  nucleotide sequence of the gene coding for subunit 1 of yeast
RT  cytochrome oxidase.";
RL  J. Biol. Chem. 255:11927-11941(1980).
RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=87163488; PubMed=3549452;
RA  de Zamaroczy M., Bernardi G.;
RT  "The primary structure of the mitochondrial genome of Saccharomyces
RT  cerevisiae -- a review.";
RL  Gene 47:155-177(1986).
CC  -!- MISCELLANEOUS: THIS PROTEIN IS CODED IN GROUP-II INTRON 1 OF OXI3
CC  (COX1).
CC  -!- SIMILARITY: TO GROUP II INTRON MATURASES.

```

Figure 5.8 Entry page using **Text Entry** format.

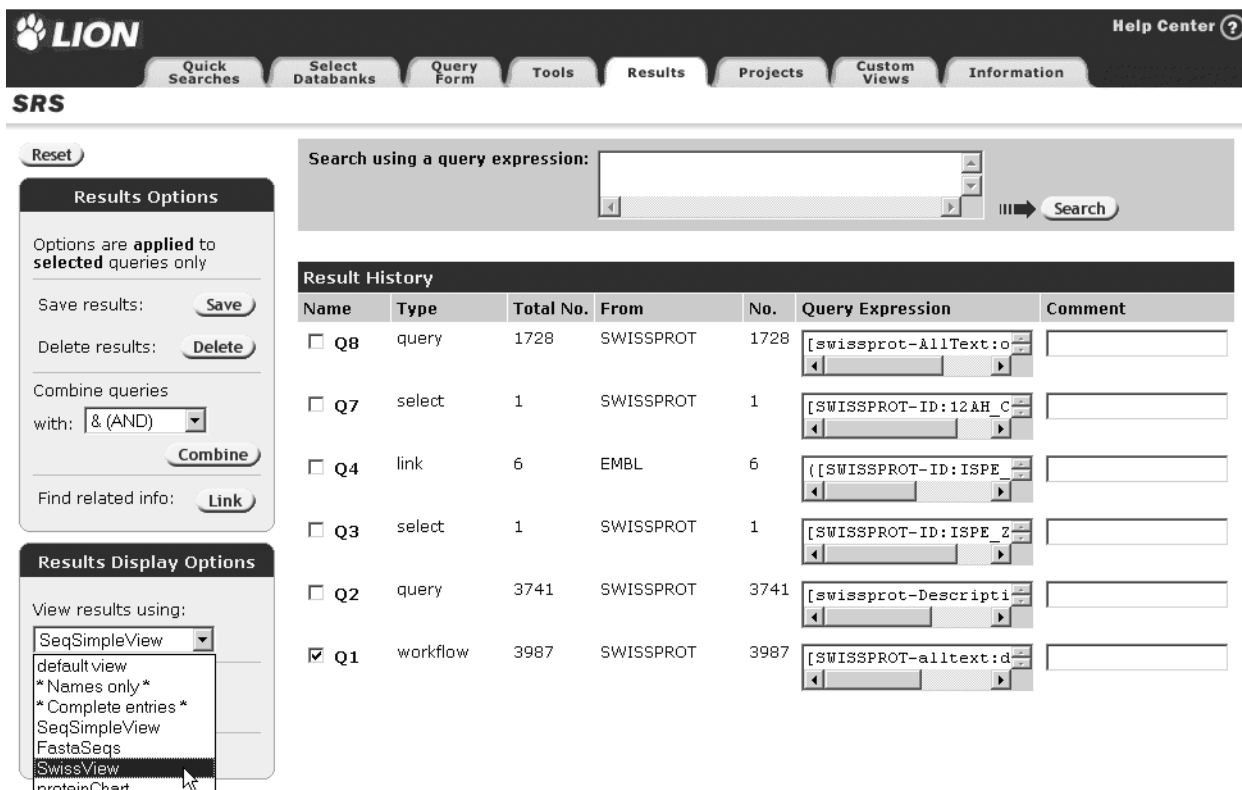
Applying a View From the Manage your Query Results Page

To apply a view from the **Manage your Query Results** page:

1. Select the query to which you want to apply the view, by ticking the check box beside it.

Note: Only one query can be displayed at a time. If more than one query is chosen, only one will be displayed.

2. Select the required view from the drop-down list in the options area (see Figure 5.9).



The screenshot shows the LION SRS interface. At the top, there is a navigation bar with tabs for Quick Searches, Select Databanks, Query Form, Tools, Results, Projects, Custom Views, and Information. A 'Reset' button is located in the top left. The main search area contains a text input field for a query expression and a 'Search' button. On the left, there are two panels: 'Results Options' and 'Results Display Options'. The 'Results Options' panel includes buttons for 'Save results:', 'Delete results:', 'Combine queries with: & (AND)', and 'Find related info:'. The 'Results Display Options' panel has a dropdown menu for 'View results using:' with a list of options: SeqSimpleView, default view, *Names only*, *Complete entries*, SeqSimpleView, FastaSeqs, SwissView (highlighted), and proteinChart. On the right, a 'Result History' table displays a list of queries.

Name	Type	Total No.	From	No.	Query Expression	Comment
<input type="checkbox"/> Q8	query	1728	SWISSPROT	1728	[swissprot-AllText:o	
<input type="checkbox"/> Q7	select	1	SWISSPROT	1	[SWISSPROT-ID:12AH_C	
<input type="checkbox"/> Q4	link	6	EMBL	6	[[SWISSPROT-ID:ISPE	
<input type="checkbox"/> Q3	select	1	SWISSPROT	1	[SWISSPROT-ID:ISPE_Z	
<input type="checkbox"/> Q2	query	3741	SWISSPROT	3741	[swissprot-Descripti	
<input checked="" type="checkbox"/> Q1	workflow	3987	SWISSPROT	3987	[SWISSPROT-alltext:d	

Figure 5.9 Selecting the required view.

3. Click the **Rerun Query** button.

Applying a View From the Launch Page for Interactive Tools

Note:For analysis tools, run via a batch queue, the option to set a view is not available on the **Launch** page, but you can

apply a view to the results after they have been generated. Follow the instructions given in sections “From the Query Result Page” (p. 114) and “Applying a View From the Entry Page” (p. 117).

To choose a view for the results of a tool other than the default view:

1. From the **Query Result** page choose the entries against which you want to run the tool, or choose **unselected only**, from the **Apply Options to:** box.

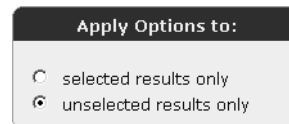


Figure 5.10 Using the option buttons to choose **unselected only**.

2. Select an analysis tool from the **Result Options** box. Click the **Launch** button to display the **Launch** page. See chapter 6, “Analysis Tools”, (p. 137) for more information on launching tools.
3. Find and select the view you want from the **View results using:** drop-down list (see Figure 5.11).

Launch - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites Media History Mail Print Edit Discuss

Address http://klee:7313/srs71bin/cgi-bin/wgetz

LION Help Center ?

Quick Searches Select Databanks Query Form Tools Results Projects Custom Views Information

SRS

Reset

Result Display Options

View results using:

Blast_View

* Names only *

* Complete entries *

BlastAlignment

Blast_View

Show results automatically:

Parameter set options

Save current parameter set

BlastX More Info...

Job name: temp Database to search: SWISSPROT (Updates) Launch

EMBL: AA901739

begin 1 end 473

1	11	21	31	41	51
ttgtggatttccatagctgaggccaggtctcgcgttgggttgtgcaaagccccaggc					
61	71	81	91	101	111
cgatcccagtcctcctctgtcggacttgttaaccttcacagtcataatcgccgtcc					
121	131	141	151	161	171
tcaacgccaccggcctccttgagcgcctgagagaaaagtctcggaggtagtcgtaatcg					

Strand of query sequence to use: Both Codon Translation table: Standard Genetic Code

Figure 5.11 Choosing a view to display results from one of the analysis tools.

4. Launch the tool as usual (see section 1.7 “Using Analysis Tools”, p. 20). The results of the analysis will be displayed on the **Query Result** page using the selected view.

5.3 Creating Views

Sometimes the predefined SRS views may not show the information that you want to see. To allow you to view the particular information you require, you can create your own views using the **View Manager** page or from the **Standard**

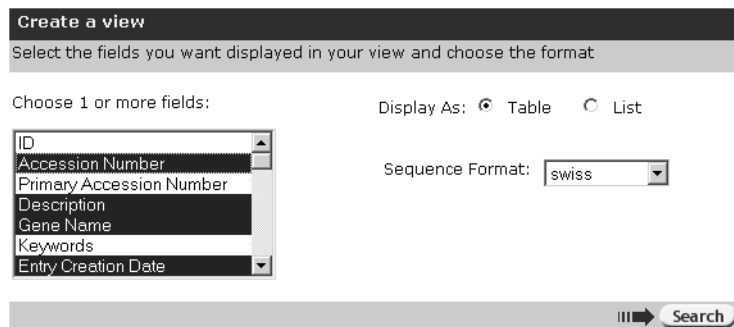
Query Form or **Extended Query Form**. These are referred to as “project views”.

5.3.1 Creating Views from the Query Forms

Both the **Standard Query Form** and the **Extended Query Form** provide a mechanism for specifying the datafields to be displayed on the **Query Result** page that results from the current query. This section describes how this is done.

Creating Views from the Standard Query Form

The bottom half of the **Standard Query Form** has a section entitled **Create a view** which allows you to specify the datafields that are to be used to display the results of your query (see Figure 5.12). You can also choose whether to display the results using a table or a list view, and the sequence format that is used.



The screenshot shows a web interface for creating a view. At the top, there is a dark header with the text "Create a view" and a subtitle "Select the fields you want displayed in your view and choose the format". Below this, there are two main sections. On the left, under the heading "Choose 1 or more fields:", there is a list box containing the following items: ID, Accession Number, Primary Accession Number, Description, Gene Name, Keywords, and Entry Creation Date. On the right, there are two options: "Display As:" with radio buttons for "Table" (selected) and "List", and "Sequence Format:" with a dropdown menu set to "swiss". At the bottom right of the form is a "Search" button with a magnifying glass icon.

Figure 5.12 Specifying the datafields to be used to display results on the **Standard Query Form**.

To design your own view:

1. Select the datafields you want to display from the **Choose 1 or more fields:** list (see Figure 5.12).

Note: You can usually use the **Shift** or **Control** keys to select multiple fields, but refer to your browser's documentation for help, if required.

2. Choose whether you wish to display the results using a table or a list view, using the option buttons.

Display As: Table List

Figure 5.13 Table and list view option buttons.

3. Choose the sequence format using the drop-down list.

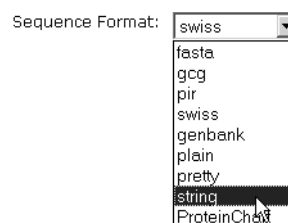


Figure 5.14 Sequence format drop-down list.

4. Complete the query as normal.
The results will be displayed on the **Query Result** page as specified.

LION Help Center ?

Quick Searches Select Databanks Query Form Tools Results Projects Custom Views Information

SRS

Reset Query "[swissprot-ALLTEXT:amylase*]" found 99 entries next

Apply Options to:

selected results only
 unselected results only

Result Options

Launch analysis tool:
BlastP

Show tools relevant to these results:

Link to related information:

Save results:

Display Options

View results using:
SeqSimpleView

Sort results by:
unsorted

ascending
 descending

Show 30 results per page

Printer friendly view

SWISSPROT	Accession	Description	SeqLength
<input type="checkbox"/> SWISSPROT:AM3D_ORYSA	P27933	Alpha-amylase isozyme 3D precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).	435
<input type="checkbox"/> SWISSPROT:AM3E_ORYSA	P27934	Alpha-amylase isozyme 3E precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).	437
<input type="checkbox"/> SWISSPROT:AMC1_ORYSA	P27940	Alpha-amylase isozyme C precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase) (Isozyme 1B).	383
<input type="checkbox"/> SWISSPROT:AMY1_AEDAE	P53354	Alpha-amylase I precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).	737
<input type="checkbox"/> SWISSPROT:AMY1_AERHY	P22630	Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).	464
<input type="checkbox"/> SWISSPROT:AMY1_DICTH	P09961	Alpha-amylase 1 (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).	685
<input type="checkbox"/> SWISSPROT:AMY1_ORYSA	P17654	Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase) (Isozyme 1B).	428
<input type="checkbox"/> SWISSPROT:AMY2_DICTH	P14898	Alpha-amylase 2 (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).	562
<input type="checkbox"/> SWISSPROT:AMY3_DICTH	P14899	Alpha-amylase 3 (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).	498
<input type="checkbox"/> SWISSPROT:AMY3_HORVU	P04747	Alpha-amylase type B isozyme precursor (EC 3.2.1.1) (1,4-alpha-D- glucan glucanohydrolase) (Clone PHV19) (Fragment).	368
<input type="checkbox"/> SWISSPROT:AMYB_DROME	P81641	Alpha-amylase B precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).	494
<input type="checkbox"/> SWISSPROT:AMYR_BACS8	P17692	Raw-starch-digesting amylase precursor (EC 3.2.1.1) (1,4-alpha-D- glucan glucanohydrolase).	713
<input type="checkbox"/> SWISSPROT:AMYR_DROPS	O18552	Alpha-amylase-related protein precursor (EC 3.2.1.1).	494
<input type="checkbox"/> SWISSPROT:AMYR_DROSU	O18420	Alpha-amylase-related protein precursor (EC 3.2.1.1).	494
<input type="checkbox"/> SWISSPROT:AMY_BACAM	P00692	Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).	514
<input type="checkbox"/> SWISSPROT:AMY_BACCI	P08137	Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).	528
<input type="checkbox"/> SWISSPROT:AMY_BACME	P20845	Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).	520
<input type="checkbox"/> SWISSPROT:AMY_BACST	P06279	Alpha-amylase precursor (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase).	549

Figure 5.15 Query Result page, showing a table view.

SRS Query "[swissprot-AllText:amylase*]" found 99 entries

Apply Options to:

selected results only
 unselected results only

Result Options

Launch analysis tool:
 BlastP

Show tools relevant to these results:

Link to related information:

Save results:

Display Options

View results using:
 SWISSPROT

Sort results by:
 unsorted

ascending
 descending

Show 30 results per page

Printer friendly view

[SWISSPROT:AM3D_ORYSA](#)

AC P27933;
 DE Alpha-amylase isozyme 3D precursor (EC 3.2.1.1) (1.4-alpha-D-glucan
 DE glucanohydrolase).
 GN AMY1.3 OR AMY3D.
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

[SWISSPROT:AM3E_ORYSA](#)

AC P27934;
 DE Alpha-amylase isozyme 3E precursor (EC 3.2.1.1) (1.4-alpha-D-glucan
 DE glucanohydrolase).
 GN AMY1.4 OR AMY3E.
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

[SWISSPROT:AMC1_ORYSA](#)

AC P27940;
 DE Alpha-amylase isozyme C precursor (EC 3.2.1.1) (1.4-alpha-D-glucan
 DE glucanohydrolase) (Isozyme 1B).
 GN AMYC OR AMY1B.
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

[SWISSPROT:AMY1_AEDAE](#)

AC P53354;
 DE Alpha-amylase I precursor (EC 3.2.1.1) (1.4-alpha-D-glucan
 DE glucanohydrolase).
 GN AMY1 OR AMY I.
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)

[SWISSPROT:AMY1_AERHY](#)

AC P22630;
 DE Alpha-amylase precursor (EC 3.2.1.1) (1.4-alpha-D-glucan
 DE glucanohydrolase).
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)

[SWISSPROT:AMY1_DICTH](#)

AC P09961;
 DE Alpha-amylase precursor (EC 3.2.1.1) (1.4-alpha-D-glucan
 DE glucanohydrolase).
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last annotation update)

Figure 5.16 Query Result page, showing a list view.

Creating Views from the Extended Query Form

On the right of the Data Area is a column entitled **Create a view** (see Figure 5.17). This contains check boxes (one for each datafield) which are used to tell SRS to include a

datafield in the results. You do not have to enter a search term in a field to have it included in the results. The drop-down list in the **Result Display Options** box allows you to choose whether to create a table or list view.

Search Options

Combine search terms with:

Use wildcards

Get results of type:

Result Display Options

View results using:

or

Create a view using select:

Sequence Format:

Show results per page

Tips

You can also use the **Standard Query** Form.

Fields you can search	Your search terms	Create a view
AllText	amylase	
ID		<input checked="" type="checkbox"/>
Accession Number		<input checked="" type="checkbox"/>
Primary Accession Number		<input type="checkbox"/>
Description		<input type="checkbox"/>
Gene Name		<input checked="" type="checkbox"/>
Keywords		<input checked="" type="checkbox"/>
Entry Creation Date	select 1 Jan YYYY 1 Jan YYYY	<input type="checkbox"/>
LastSequenceUpdate	select 1 Jan YYYY 1 Jan YYYY	<input checked="" type="checkbox"/>
LastAnnotationUpdate	select 1 Jan YYYY 1 Jan YYYY	<input checked="" type="checkbox"/>
Organism Name		<input checked="" type="checkbox"/>
Taxon		<input type="checkbox"/>
NCBI TaxId	>= <=	<input type="checkbox"/>
Organelle		<input type="checkbox"/>
ProteinID		<input type="checkbox"/>
Checksum		<input type="checkbox"/>
DbName		<input type="checkbox"/>

or
 and

Figure 5.17 Select the fields to include on the **Extended Query Form**.

To create a view from the **Extended Query Form**:

1. Enter your search term(s) as usual.
See chapter 3, “Querying with SRS”, (p. 55), for details.
2. Use the drop-down list to choose whether the view will display as a table or list view.
3. Tick the check box to the right of any fields that you want to display in the results.
4. Continue with your query as usual.

Typical results for table and list views are shown in Figures 5.15 and 5.16 respectively.

5.3.2 Creating Views using the View Manager Pages

If the built-in views do not do what you need and the options on the **Query Forms** are insufficient, you can use the **View Manager** pages to create views.

Note: There are two **View Manager** pages, referred to here as **View Manager** page 1 and **View Manager** page 2.

Click the **Views** tab to go to the **View Manager** pages. You can control the following elements of any views that you create:

- The name of the view (**View Manager** page 1).
- Whether the view displays as a table or list format (**View Manager** page 1).
- Whether the fields show on **View Manager** page 2 will include all available fields, or just the common ones (**View Manager** page 1).

- The list of databanks for which the view is defined (and for which the view is therefore available). These are referred to as “root databanks” (**View Manager** page 1).
- The list of databanks that may have a link with the entries (**View Manager** page 1).
- The list of fields from the root databanks and the linked databanks that should be displayed (**View Manager** page 2).
- The format of some fields such as **Sequences** (**View Manager** page 2).

In selecting the root databanks for a view, you specify two important things. Firstly, the view will only be available for queries and result sets based on these databanks. Secondly, the available fields from the root databanks will be those that are common across all of the root databank selections. Therefore, it is sensible to select databanks that have similar fields when you are creating a new view.

To Create a View

1. Click on the **Custom Views** tab. This will take you to the **View Manager** page 1.
2. Name the view by entering a suitable name in the **View name** text box.

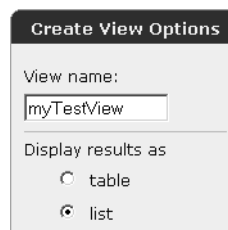


Figure 5.18 Create View Options box, showing naming of a view and the table and list view options.

Note: It is better not to include spaces and other special characters in names as some systems do not handle them properly. Use an underscore, or start new words with a capital letter instead.

3. Choose whether the new view will be displayed using a table or list format.
4. Choose whether you wish to choose from all available fields in the selected databanks, or just the common ones.

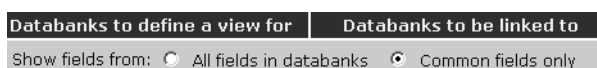


Figure 5.19 Selecting fields.

5. Select the databanks for which the view will be available, using the left-hand column of databanks, headed **Databanks to define a view for**. These are the root databanks for the view.
6. Select the databanks to which you wish to link, using the right-hand column of databanks, headed **Databanks to be linked to**.

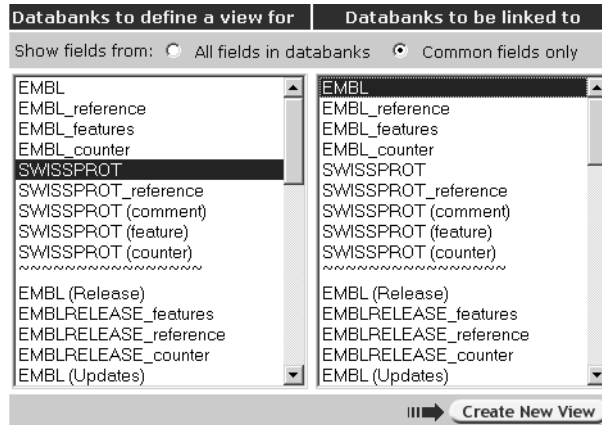


Figure 5.20 Selected databanks.

7. Click on the **Create New View** button (see Figure 5.21).

SRS

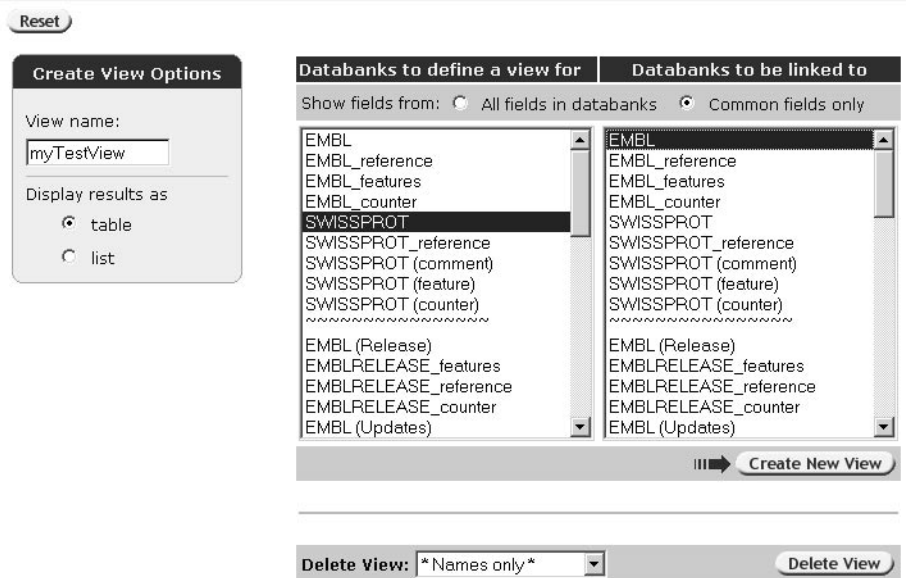


Figure 5.21 View Manager page 1.

SRS will display the **View Manager** page 2 on which you can select the specific fields for your view. By default (it can be changed on **View Manager** page 1) the top box contains the fields which are common to all of the root databanks selected on the previous page. There are additional boxes below this for each linked databank you chose. Using these check boxes, select the fields that you would like to see in your view.

8. Select the fields you would like to display in your view by ticking the check boxes beside the relevant fields.

Create View Options

Select the datafields you want **displayed in your view** using the checkboxes.

View name: myTestView

Save view:

Datafields for your primary databanks

SWISSPROT

<input type="checkbox"/> ID	<input type="checkbox"/> AccNumber	<input type="checkbox"/> PrimAccNumber	<input checked="" type="checkbox"/> Description
<input type="checkbox"/> GeneName	<input type="checkbox"/> Keywords	<input type="checkbox"/> DateCreated	<input type="checkbox"/> LastSequenceUpdate
<input type="checkbox"/> LastAnnotationUpdate	<input type="checkbox"/> Organism	<input type="checkbox"/> Taxon	<input type="checkbox"/> NCBI_TaxId
<input type="checkbox"/> Organelle	<input type="checkbox"/> ProteinID	<input type="checkbox"/> Checksum	<input type="checkbox"/> DbName
<input type="checkbox"/> DBxref	<input type="checkbox"/> SeqLength	<input type="checkbox"/> Sequence	fasta

Fields of subentry **Reference**

<input type="checkbox"/> Authors	<input type="checkbox"/> Title	<input type="checkbox"/> Journal	<input type="checkbox"/> VolumeNo
<input type="checkbox"/> FirstPage	<input type="checkbox"/> Year	<input type="checkbox"/> MedlineID	<input type="checkbox"/> PubMedID
<input type="checkbox"/> RefPosition	<input type="checkbox"/> RefCommentCode	<input type="checkbox"/> RefComment	

Fields of subentry **Comment**

<input type="checkbox"/> CommentType	<input type="checkbox"/> Comment
--------------------------------------	----------------------------------

Fields of subentry **Feature**

<input type="checkbox"/> FtKey	<input type="checkbox"/> FtLength	<input type="checkbox"/> FtDescription
--------------------------------	-----------------------------------	--

Fields of subentry **Counter**

<input type="checkbox"/> CountedItem	<input type="checkbox"/> CountedN
--------------------------------------	-----------------------------------

Datafields from databanks to be linked to

EMBL

<input type="checkbox"/> ID	<input type="checkbox"/> Division	<input type="checkbox"/> AccNumber	<input type="checkbox"/> PrimAccNumber
<input type="checkbox"/> SeqVersion	<input type="checkbox"/> Molecule	<input checked="" type="checkbox"/> Description	<input type="checkbox"/> Keywords
<input type="checkbox"/> Organism	<input type="checkbox"/> Taxon	<input type="checkbox"/> Organelle	<input type="checkbox"/> Comment
<input type="checkbox"/> DateCreated	<input type="checkbox"/> LastUpdated	<input type="checkbox"/> SeqLength	<input type="checkbox"/> Link
<input type="checkbox"/> Sequence	fasta		

Fields of subentry **References**

<input type="checkbox"/> ID	<input type="checkbox"/> Authors	<input type="checkbox"/> Title	<input type="checkbox"/> Journal
<input type="checkbox"/> VolumeNo	<input type="checkbox"/> FirstPage	<input type="checkbox"/> Year	<input type="checkbox"/> MedlineID
<input type="checkbox"/> PubMedID	<input type="checkbox"/> RefPosition		

Fields of subentry **Features**

<input type="checkbox"/> ID	<input type="checkbox"/> AccNumber	<input type="checkbox"/> FtKey	<input type="checkbox"/> FtQualifier
<input type="checkbox"/> ProteinID	<input type="checkbox"/> Gene	<input type="checkbox"/> FtDescription	<input type="checkbox"/> FtLength
<input type="checkbox"/> Sequence	gCg		

Fields of subentry **Counters**

<input type="checkbox"/> ID	<input type="checkbox"/> CountedItem	<input type="checkbox"/> CountedN
-----------------------------	--------------------------------------	-----------------------------------

Use explicit link:

Use view to display entries:

Display only number of linked entries only

Figure 5.22 Select the datafields to be included in your view.

Note: Each field name is a hypertext link to its **Field Information** page (see section on the “Information Pages - Field Information” in the *Online Help*). So if you want to know more about a particular field, use these hypertext links to find out.

9. Click on the  button.

Your new view can now be used for any of the databanks indicated in the root databanks (see step 5, above). Try your new view for a query on one or more of the root databanks.

CHAPTER

6

ANALYSIS TOOLS

Various analysis tools can be run on queries within SRS.
This chapter describes how to:

- Use the tools that are available within SRS.

6.1 Introducing Analysis Tools

SRS analysis tools are bioinformatics programs that use a databank query as input. The output file from an analysis tool is indexed in the same way as any other databank. This enables users to store and query their analyses.

The results from one analysis tool can be used by another. Hits in a sequence databank from a sequence similarity search (e.g., one of the BLAST programs) can be selected and aligned. The resulting alignment can then be used for further analysis.

6.2 Accessing Analysis Tools

Analyses using tools usually start with a databank query (although it is possible to use sequences from other sources). The **Query Result** page includes a **Launch analysis tool:** drop-down list that lists the most commonly used available tools. This can be used to access the analysis tools in the drop-down list. (Select a tool from the drop-down list and click the **Launch** button.) Alternatively, the **Tools** button will take you to the **Tool Select** page. This lists all the tools that can use the current query data as input.

You can also get to the **Tool Select** page by clicking on the **Tools** tab. This will list all of the applications that are available on your SRS installation. However, you will have to enter your own sequence data on the **Launch** page.

Section 1.7 “Using Analysis Tools” (p. 20) explains how to start a tool. The remainder of this chapter gives further

information on the use of tools, and further details of the differing behavior of interactive and batch tools.

6.3 Launch Page

6.3.1 Tool Parameters

LION Help Center ?

Quick Searches Select Databanks Query Form Tools Results Projects Custom Views Information

SRS

Reset

Result Display Options

View results using:
Blast_View

Show 30 results per page

Show results automatically:

Parameter set options

Save current parameter set as:

BlastX More Info...

Job name: temp Database to search: SWISSPROT (Updates) Launch

EMBL: MMNKX

begin 1

```
1 11 21 31 41 51
cgaggggcccaggtacaggggcaggcaggcctgagctctgtagaggaggtactcgggggt
61 71 81 91 101 111
gtgtcagagcagcacaggtcaccagagctcttgagcaaaaagaggagacctcaagactat
121 131 141 151 161 171
cttcccatggccaagagagaaagctttcagggcacaggcaagaggaggagggccaccgag
```

end 1972

Strand of query sequence to use: Both Codon Translation table
Standard Genetic Code

Output Options

Number of hits and alignments to show 50

Number of best hits from a region to keep 100

Search Parameters

Filter query sequence

Scoring matrix BLOSUM62

The E value 10.000000

word size Default

Perform gapped alignment

Cost to open a gap Default

Cost to extend a gap Default


Figure 6.1 Launch page for an analysis tool which will be run interactively (see also section 6.4 “Interactive and Batch Tools”, p. 140).

For sequence similarity searches, there is usually a choice of databanks to be searched. In many cases using the other default values on the **Launch** page should provide adequate results. However, the values can be changed freely should the need arise. SRS will display an error message if an invalid value is given.

The full sequence is displayed (providing the analysis tool has been accessed from a valid query, rather than direct from the **Tools Select** page). This means that you can edit it, usually by changing the begin or end position. Editing the contents of the sequence itself is also possible.

6.4 Interactive and Batch Tools

6.4.1 Introduction

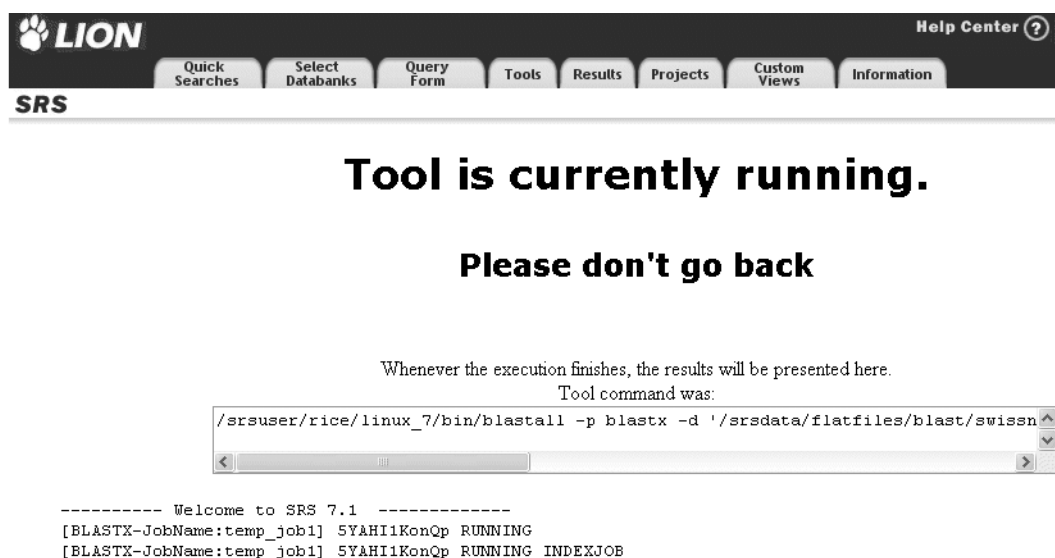
When you click on the  button, the chosen analysis tool will either be run interactively (you wait for the results to be displayed) or will be submitted to a batch queue. The **Launch** page indicates which method is to be used. Figure 6.1 (above) shows the **Launch** page for a tool which is to be run interactively. Figure 6.4 (below) shows a typical batch queue message from a **Launch** page for a tool which has been sent to a batch queue.

Note: The choice of whether tools are run interactively or submitted to batch queues depends on how SRS has been set up.

The next sections describe these methods (see sections 6.4.2 “Interactive Tools”, p. 141 and 6.4.3 “Batch Tools”, p. 142).

6.4.2 Interactive Tools

Tools that do not take long to run are usually run interactively. This means they will be started immediately, and run while you wait. The **Tool Invocation** page (Figure 6.2) will be displayed while the analysis is running.



The screenshot shows the SRS (Sequence Retrieval System) interface. At the top, there is a navigation bar with the LION logo on the left and a Help Center icon on the right. Below the logo is the text 'SRS'. The navigation bar contains several buttons: 'Quick Searches', 'Select Databanks', 'Query Form', 'Tools', 'Results', 'Projects', 'Custom Views', and 'Information'. The main content area displays the message 'Tool is currently running.' in large, bold, black text. Below this message is the instruction 'Please don't go back' in bold black text. Further down, there is a text box that says 'Whenever the execution finishes, the results will be presented here.' followed by 'Tool command was:' and a scrollable text area containing the command: '/srsuser/rice/linux_7/bin/blastall -p blastx -d '/srsdata/flatfiles/blast/swissn'. Below the scrollable area, there is a status bar with the following text: '----- Welcome to SRS 7.1 -----', '[BLASTX-JobName:temp_job1] SYAHI1KonQp RUNNING', and '[BLASTX-JobName:temp_job1] SYAHI1KonQp RUNNING_INDEXJOB'.

Figure 6.2 Tool Invocation page.

The **Query Result** page (Figure 6.3) will appear automatically on completion unless the **show results automatically** check box has been unticked.

BLASTX	Query	Search Database	Hit
<input type="checkbox"/> BLASTX:temp_job1_swissnew_1_HMX_STRPU	EMBL:MMNKX	swissnew	>sp Q26656 HMX_STRPU
<input type="checkbox"/> BLASTX:temp_job1_swissnew_2_HK32_MOUSE	EMBL:MMNKX	swissnew	>sp P97503 HK32_MOUSE
<input type="checkbox"/> BLASTX:temp_job1_swissnew_3_HK32_HUMAN	EMBL:MMNKX	swissnew	>sp P78367 HK32_HUMAN
<input type="checkbox"/> BLASTX:temp_job1_swissnew_4_HM1D_DROAN	EMBL:MMNKX	swissnew	>sp P22544 HM1D_DROAN
<input type="checkbox"/> BLASTX:temp_job1_swissnew_5_HMX2_COTJA	EMBL:MMNKX	swissnew	>sp P23410 HMX2_COTJA
<input type="checkbox"/> BLASTX:temp_job1_swissnew_6_NK2E_XENLA	EMBL:MMNKX	swissnew	>sp P42583 NK2E_XENLA
<input type="checkbox"/> BLASTX:temp_job1_swissnew_7_HMX2_CHICK	EMBL:MMNKX	swissnew	>sp P28362 HMX2_CHICK
<input type="checkbox"/> BLASTX:temp_job1_swissnew_8_SAX1_MOUSE	EMBL:MMNKX	swissnew	>sp P42580 SAX1_MOUSE
<input type="checkbox"/> BLASTX:temp_job1_swissnew_9_HMBP_DROME	EMBL:MMNKX	swissnew	>sp P22809 HMBP_DROME

Figure 6.3 Query Result page showing the results of an analysis using BLASTX.

6.4.3 Batch Tools

6.4.3.1 Introduction to Batch Tools

Tools which are likely to take a while to run are usually set to run as batch jobs. SRS does not wait while batch jobs are run.

Instead you can carry on with further queries, etc. and return to look at the results of your analysis when it is complete. If a job will be run from the batch queue this is indicated on the **Launch** page prior to launching (Figure 6.4).

The screenshot shows the SRS interface for the BlastN tool. At the top, there is a navigation bar with 'LION' logo and 'Help Center (?)'. Below it are tabs for 'Quick Searches', 'Select Databanks', 'Query Form', 'Tools', 'Results', 'Projects', 'Custom Views', and 'Information'. The main content area is titled 'SRS' and contains a 'Reset' button and a 'BlastN' section. The 'BlastN' section has a 'Job name' field with 'temp' and a 'Database to search' dropdown menu set to 'EMBL (Updates)'. A 'Launch' button is visible. Below these fields, the EMBL accession number 'AA901762' is shown, along with 'begin' (1) and 'end' (580) fields. A sequence viewer displays the DNA sequence: 'tgtataaccaccctcacaaacgatggcatggtacgcccgaacctagcaaaggcccacaa61 71 81 91 101 111 ggggtcccatgtcaccggctaagagcatctccccgggcatccacgagacggagg121 131 141 151 161 171 tagagggaaaaccactgggtcgtcggtccgtcgaccaaggcaccacctctccccgttc'. A 'Strand of query sequence to use' dropdown is set to 'Both'. To the left of the BlastN section, there are 'Job Options' and 'Parameter set options' panels. The 'Job Options' panel includes a note about LSF batch queuing and a dropdown menu for 'blast-m sisley(batch)'. The 'Parameter set options' panel has a 'Save current parameter set as:' field.

Figure 6.4 Typical SRS batch queuing message from a **Launch** page, for a run of BLASTN.


6.4.3.2 Tool Invocation Page for Batched Tools


After the **Launch** button is clicked, the **Tool Invocation** page will be shown.

The screenshot shows the SRS 'Tool Invocation' page. At the top, there is a navigation bar with 'LION' logo and 'Help Center (?)'. Below it are tabs for 'Quick Searches', 'Select Databanks', 'Query Form', 'Tools', 'Results', 'Projects', 'Custom Views', and 'Information'. The main content area is titled 'SRS' and displays the message: 'Tool was submitted to Queue:blast -m sisley(batch). Tool command: blastall -p blastn -d '/srsdata/flatfiles/blast/emblnew' -i ./temp_blastn_1_in'. Below the message is a text box containing the command: 'blastall -p blastn -d '/srsdata/flatfiles/blast/emblnew' -i ./temp_blastn_1_in'. At the bottom, there is a link: 'Use Batch job status page to view the results'.

Figure 6.5 Tool Invocation page for batch processing.

6.4.3.3 Accessing the Results

The results from an analysis can be reached from the **Job Status** page (see Figure 6.6). This page is reached by clicking on the [results](#) hyperlink or on the  icon on the **Tool Invocation** page.



Job Options

Options are **applied** to **selected** jobs only

Delete jobs:

View job results using:

* Complete entries *



Run job again with different options:

List of Batch Jobs

Job Name	Status	Start Date	Results from	Result Set	Queue Name
<input type="checkbox"/> temp_blastn_1	<input checked="" type="checkbox"/>	18-Mar-2003 12:40	BLASTN	Q2: 50	priority -m klee(batch)

SRS Release 7.1 Copyright © 1997-2003 LION bioscience AG. All Rights Reserved. [Terms of Use](#) [Feedback](#)

Figure 6.6 Job Status page.

When a job has completed, the job name becomes a link to the results of your analysis, and the hourglass  becomes an hourglass with a green tick . Clicking on the hyperlink for the job that has just run, e.g. [temp_blastn_1](#) will display the **Entry** page containing the results.

Query "[BLASTN-JobName:temp_job1]" found 17 entries

BLASTN	Query	Search Database	Hit
<input type="checkbox"/> BLASTN:temp_job1_emb1new 1 AC104140	EMBL:AA901632	emblnew	>emb AC104140 AC104140
<input type="checkbox"/> BLASTN:temp_job1_emb1new 2 AX305845	EMBL:AA901632	emblnew	>emb AX305845 AX305845
<input type="checkbox"/> BLASTN:temp_job1_emb1new 3 AP001551	EMBL:AA901632	emblnew	>emb AP001551 AP001551
<input type="checkbox"/> BLASTN:temp_job1_emb1new 4 AX305600	EMBL:AA901632	emblnew	>emb AX305600 AX305600
<input type="checkbox"/> BLASTN:temp_job1_emb1new 5 AF443205	EMBL:AA901632	emblnew	>emb AF443205 AF443205
<input type="checkbox"/> BLASTN:temp_job1_emb1new 6 SPBC8D2	EMBL:AA901632	emblnew	>emb AL022072 SPBC8D2
<input type="checkbox"/> BLASTN:temp_job1_emb1new 7 AY069346	EMBL:AA901632	emblnew	>emb AY069346 AY069346
<input type="checkbox"/> BLASTN:temp_job1_emb1new 8 AY064020	EMBL:AA901632	emblnew	>emb AY064020 AY064020
<input type="checkbox"/> BLASTN:temp_job1_emb1new 9 AC023741	EMBL:AA901632	emblnew	>emb AC023741 AC023741
<input type="checkbox"/> BLASTN:temp_job1_emb1new 10 AC023681	EMBL:AA901632	emblnew	>emb AC023681 AC023681
<input type="checkbox"/> BLASTN:temp_job1_emb1new 11 NC18F11	EMBL:AA901632	emblnew	>emb AL670011 NC18F11
<input type="checkbox"/> BLASTN:temp_job1_emb1new 12 CEF07A5	EMBL:AA901632	emblnew	>emb Z72506 CEF07A5
<input type="checkbox"/> BLASTN:temp_job1_emb1new 13 AY069602	EMBL:AA901632	emblnew	>emb AY069602 AY069602
<input type="checkbox"/> BLASTN:temp_job1_emb1new 14 AY053423	EMBL:AA901632	emblnew	>emb AY053423 AY053423
<input type="checkbox"/> BLASTN:temp_job1_emb1new 15 AP003407	EMBL:AA901632	emblnew	>emb AP003407 AP003407
<input type="checkbox"/> BLASTN:temp_job1_emb1new 16 AL646060	EMBL:AA901632	emblnew	>emb AL646060 AL646060
<input type="checkbox"/> BLASTN:temp_job1_emb1new 17 AC091221	EMBL:AA901632	emblnew	>emb AC091221 AC091221







Figure 6.7 Entry page showing open .

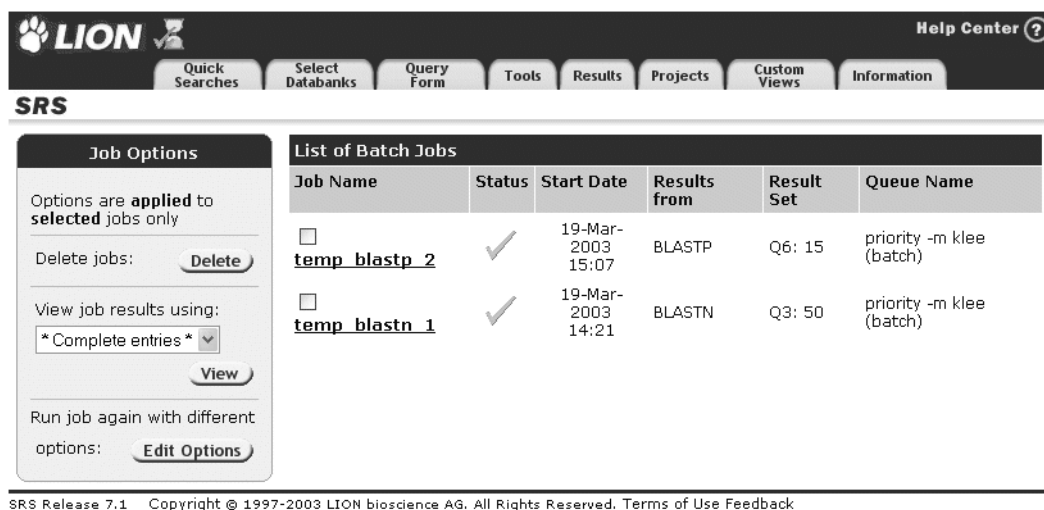
Alternatively, it is possible to select a job by ticking the check box beside it, selecting a view and clicking the **View** button. This will display the results on a **Query Result** page.

6.4.3.4 Job Status

The **Job Status** page (see Figure 6.8) lists all batch jobs in

the current project and shows their current status. (The most recent job is at the top of the list.)

- Completed jobs are shown with a  icon.
- Jobs still running are shown with a  icon.
- Waiting jobs (i.e. those which are waiting to be run) are shown with a  icon.
- Jobs that have been submitted to the queue are shown with a  icon.
- Jobs which have produced no meaningful results (e.g. those for which the search sequence produced no appropriate results or those which have crashed) are shown with a  icon.
- Jobs which have unspecified problems may be shown with a  icon.



Job Options

Options are **applied** to **selected** jobs only



Delete jobs:

View job results using:

Complete entries

Run job again with different options:



List of Batch Jobs

Job Name	Status	Start Date	Results from	Result Set	Queue Name
<input type="checkbox"/> temp_blastp_2		19-Mar-2003 15:07	BLASTP	Q6: 15	priority -m klee (batch)
<input type="checkbox"/> temp_blastn_1		19-Mar-2003 14:21	BLASTN	Q3: 50	priority -m klee (batch)

SRS Release 7.1 [Copyright](#) © 1997-2003 LION bioscience AG. All Rights Reserved. [Terms of Use](#) [Feedback](#)

Figure 6.8 Job Status page showing multiple jobs.

6.4.3.5 Accessing the Job Status Page

The  icon appears on SRS pages when batch jobs exist, and can be used to access the **Job Status** page. It changes to , when all jobs have completed.

CHAPTER

7

COMMAND LINE SRS

The SRS command line interface is called `getz`. This chapter:

- Introduces `getz`.
- Explains the available `getz` options.

7.1 getz

The SRS command line is available through a UNIX shell window, and uses a program called `getz`. You may need to contact your SRS Administrator to set this up.

Using `getz` you can query databanks from the command line. For further information on creating `getz` queries, see section 7.1.1 “`getz` Usage” (p. 151) and chapter 8, “SRS Query Language”, (p. 155).

Here are a few examples:

To retrieve a list of entries from the SWISS-PROT databank that have “azurin” mentioned in the description, you might enter:

```
getz "[swissprot-des:azurin]"
```

If you want to perform the same search as above, but retrieve the complete entries, use the `-e` argument:

```
getz -e "[swissprot-des:azurin]"
```

To retrieve the sequence for the entries using the FASTA format, type:

```
getz -f seq -sf fasta "[swissprot-des:azurin]"
```

To specify that the results should be sorted (and the sort direction) include `-sort` and `-sortDir` arguments in your command.

```
getz "[swissprot-id:p*]" -sort "sl acc" -sortDir "01"
```

Section 7.1.1 “getz Usage”, p. 151 explains more about the arguments.

7.1.1 getz Usage

The following table lists the options available for command line SRS. These options can be used to refine your getz queries. Alternatively, to find out more about getz, use the command:

```
getz -help
```

Table 7.1 Common getz options.

Option	Default	Function
-help		Help with getz.
-e	FALSE	Prints the entire entry.
-t	FALSE	Copy the complete text (annotation) part of the entry..
-d	FALSE	Copy the data (e.g. sequence) part of the entry.
-i	FALSE	Print the tokens that would be generated for indexing.
-f <string>	""	Include fields in entry list.
-vf <string>	""	List of fields that will be placed into a table view.
-w	FALSE	Appends a wildcard to each search word.
-lb <n>	0	Number of first entry in set to be viewed.

Table 7.1 Common getz options. (Continued)

Option	Default	Function
-ll <n>	0	Number of entries to be viewed in one go.
-lv	FALSE	List all values that match the query.
-lvf	FALSE	List all values that match the query, plus the number of entries for each match.
-lmin <n>	0	List only values that occur at least the specified number of times. Use together with <i>-lv</i> .
-c	FALSE	Report the number of entries that were found, but not the entries themselves.
-info	FALSE	Prints info about the specified databank.
-libs	FALSE	Prints a list of all active databanks.
-view <string>	""	Name of view to be used when displaying entries.
-rs <string>	""	String of one or more characters to separate records in view.
-cs <string>	""	String of one or more characters to separate columns in view.
-sf <string>	""	Format of sequence output file.
-af <string>	""	Format of sequence alignment output file.

Table 7.1 Common getz options. (Continued)

Option	Default	Function
-html	FALSE	Select HTML format for output.
-ascii	TRUE	Select ASCII format for output.
-off	FALSE	Accesses the off-line version of a databank.
-id <string>	""	The user ID, or filename associated with a WWW session.
-sort <string>	""	The name of the field on which to sort the query.
-sortDir <string>	""	Ascending (0) or descending sort.

8

SRS QUERY LANGUAGE

In SRS many actions, e.g., retrieval commands, logical operations with sets obtained from previous queries, links between entries from different databanks, or a combination, can be expressed using the SRS query language. This chapter describes the language and gives examples of its use. During this chapter you will learn:

- About searching in indices.
- The general syntax of the SRS query language.
- About using logical and link operators.
- About entries and subentries.
- About storing intermediate results in sets.

8.1 Introduction

In earlier chapters you have seen how to use the SRS **Query Forms**, **Link** page, etc., to create a query. In addition to this method, you can type your query using the SRS query language. There are various places where this can be done, including, the **Quick Search** text box on the **Select Databanks To Search** page, for example.

The SRS query language syntax is also used internally by SRS whenever you make any query (e.g., using the query forms, linking, etc.), so you will see examples of them on SRS pages that contain the results of any such queries, (e.g., at the top of the **Query Result** page).

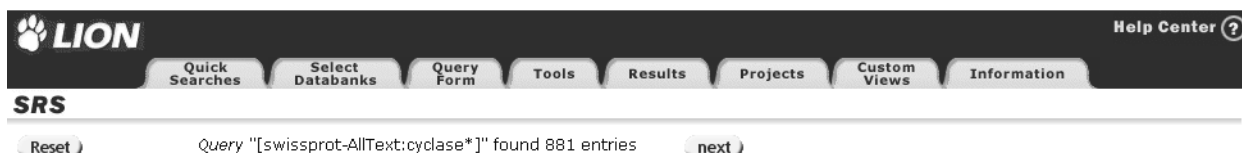


Figure 8.1 Query Result page, showing a query of the **AllText** field of the SWISS-PROT databank, for **cyclase**.

In addition to simple searches of single databanks you can search multiple databanks, link searches, etc.

8.2 Searching in Indices

8.2.1 Introduction

Probably the simplest form of the SRS query language syntax is that used for simple searches in indices. Index searches include searches for simple strings, searches for numbers and ranges of numbers, as well as searches for dates. This section covers the various forms of index search.

8.2.2 General Syntax

An index search must specify within square brackets: the databank or databank group name, the index or index group name, and a search expression. The two names must be separated by a hyphen (-), and be separated from the search expression by, either a colon (:), (string search, see section 8.2.3 “Search Strings”, p. 158), or a hash (#), (range search). Range searches can be performed only in indices of the types num and real (see section 8.2.4 “Searching Using Numerical Ranges”, p. 161 and section 8.2.5 “Searching for Dates”, p. 163).

Either the field name (e.g., `description`) or its abbreviation (`des`) can be used as the index name. All strings, including the search words, are case-insensitive. For example:

```
[pir-des:elastase]
```

searches for the string “elastase” in the `des` (description) field of the protein databank, PIR.

8.2.3 Search Strings

A search string may be a single search word, or several words separated by logical operator(s) (see section 8.3.4 “Operators”, p. 168). Parentheses may be used to create a group which will be treated as a single operand (see Example 8.1 Search strings). Wildcards and regular expressions may also be used (see section “Wildcards”, p. 159 and section “Regular Expressions”, p. 159).

Example 8.1 Search strings

To search the `keywords` field of the EMBL databank for “insulin” you might enter:

```
[embl-key:insulin]
```

To search the `description` field of the EMBL databank for entries which include “acetylchol*” and “receptor”, but remove any entries that contain “muscarinic” you might enter:

```
[embl-des:(acetylchol*&receptor)!muscarinic]
```

To search the `authors index` field of the SWISS-PROT databank to look for entries containing “sanger,f*” but not “coulson,a*”, you might use a query like:

```
[swissprot-aut:sanger,f*!coulson,a*]
```

Wildcards

Wildcards are useful if, for example, you wish to search for a group of words (e.g., all words starting with “cell” and ending with “ase”) or if it is unclear how a word is spelt in a databank.

SRS uses two types of wildcard:

- * Matches zero or more characters of any value.
- ? Matches one character of any value.

Any number of wildcards can be placed anywhere in a search word.

Note: Placing a wildcard at the start of a word or string may increase the response time because all words in the index have to be checked against your string.

Regular Expressions

In addition to the use of wildcards, it is also possible to enter regular expressions directly. Regular expressions must appear within forward slashes (/).

Some characters (^\$.[]()*+?) have a special meaning — these must be prefixed with a backslash (\) to indicate that the specified character is to be matched literally.

Tables 8.1 and 8.2 respectively, list typical regular operands and examples of their use.

Table 8.1 Regular expression operators.

Operator	Meaning
^	A caret (^) is used to mark the start of a string. For example, /^phos/ will find all words beginning with phos (e.g., phosphate).
\$	A dollar sign (\$) is used to mark the end of a string. For example, /ase\$ will find all words ending with “ase” (e.g., kinase).
.	A dot indicates any single character.
[...]	Characters enclosed in square brackets are regarded as a set, any of which can be matched. For instance, [()] matches an opening or closing parenthesis. Character ranges can be specified using a hyphen (-), e.g., [0-9] matches any single digit. A caret (^) in front of the character set (after the opening square bracket) negates the character set: e.g., [^0-9] matches any non-digit character.
(...)	A series of pattern elements enclosed in parentheses is regarded as comprising a single element.
*	An asterisk indicates that the preceding group may be repeated zero or more times.

Table 8.1 Regular expression operators. (Continued)

Operator	Meaning
+	A plus sign indicates that the preceding group may be repeated one or more times.
?	A question mark indicates that the preceding character or group of characters occurs one or zero times.

Table 8.2 Examples of regular expressions.

Expression	Meaning
<code>/^j..\$/</code>	This expression finds all three-character strings that start with j.
<code>/^5[0-9][0-9][0-9]\$/</code>	This expression finds all four-digit numbers that start with 5.
<code>/^nif[a-e]\$/</code>	This expression finds the gene names nifa, nifb, nifc, nifd, nife.
<code>/^mue?lller\$/</code>	This expression finds both muller and mueller.

Note: Searches with regular expressions are sometimes slow since all the words in the index have to be searched.

8.2.4 Searching Using Numerical Ranges

In a numerical index (whether it contains integers or reals) it is possible to search numerical ranges. A numerical index is only possible where there is a one-to-one relationship

between entry and value (e.g., sequence length, creation date, resolution).

A range can be specified using a single value or by two values separated by a colon (:). The value on the left must be smaller than the value on the right. To exclude a value from the range, put an exclamation mark (!), in front of it. The absence of a number on the left indicates that the search should start at the minimum value in the index. Similarly, an absent value on the right indicates that the search should include values up to the maximum for that index.

Table 8.3 Examples of queries on an index of the sequence length.

Written Range	Meaning
400	All sequences with a length of exactly 400.
400:500	All sequences with lengths between 400 and 500.
400:	All sequences with lengths greater than 400.
:500	All sequences with lengths less than 500.
400:!500	All sequences with lengths between 400 and 500, excluding 500.
:	A range from the minimum value to the maximum value, i.e., <u>all</u> sequences.

Combining Ranges

Ranges can be combined using logical operators. For instance either:

```
300:!500 | !600:700
```

or

```
300:700 ! 500:600
```

would retrieve the same set of sequences, i.e., all sequences from 300 to 500, excluding 500, and all sequences from 600 to 700, excluding 600.

8.2.5 Searching for Dates

Searches for dates can be made using one of the two special formats recognized by the SRS query language. These are: YYYYMMDD or DD-MMM-YYYY. For example,

```
20020619  
19-Jun-2002
```

Dates can be used within ranges in the same way as other numbers. For example,

```
[swissprot-date#20010415:20020414]  
[swissprot-date#15-APR-2001:14-APR-2002]
```

8.2.6 Searching Multiple Databanks

As well as allowing you to search a field of a single databank, the SRS query language allows you to search multiple databanks in a single query expression. This is done using a list of databank names, enclosed in curly brackets, to replace the single databank name seen in earlier examples. The names in the list must be separated by spaces. For example:

```
[{swissprot swissnew sptrembl}-des:kinase]
```

searches for the word, “kinase”, in the **Description** index of the SWISS-PROT, SWISSNEW and SPtrEMBL databanks.

It is often convenient to give a name to a group of databanks so that that name can be used later in the query rather than repeating the list of names. For instance:

```
[dbs={swissprot swissnew sptrembl}-des:kinase]  
&[dbs-org:human]
```

creates the group, `dbs`, which combines the three databanks SWISS-PROT, SWISSNEW and SPtrEMBL, and then uses the group name, `dbs`, to replace the search name in the second part of the search.

Note: It is better not to include spaces and other special characters in names as some systems do not handle them properly. Use an underscore, or start new words with a capital letter instead.

8.3 Combining Searches

8.3.1 Introduction

The earlier parts of this chapter dealt with simple index searches. The SRS query language can also be used to create more complex queries. These take the form of expressions and are constructed using operands (see section 8.3.3 “Operands”, p. 167) and operators (see section 8.3.4 “Operators”, p. 168).

8.3.2 General Syntax

Queries using the SRS query language take the general form:

```
operand operator operand ...
```

(See section 8.3.3 “Operands”, p. 167 and section 8.3.4 “Operators”, p. 168 for more information on operands and operators, respectively.)

For example:

```
enzyme > pdb
```

where `enzyme` and `pdb` are operands specifying the databanks, Enzyme and PDB respectively, and `>` is an operator telling SRS to search for links between the databanks, and keep only those entries which belong to the databank on the right. The result of this query is a list of

entries in the PDB databank that have links to the Enzyme databank.

Combinations may also include index searches. For instance:

```
[swissprot-des:kinase] > pdb
```

will create a list of all the entries in the PDB databank that have links to the results of the index search:

```
[swissprot-des:kinase]
```

See section 8.2 “Searching in Indices” (p. 157), for an explanation of this sub-search.

The above examples are fairly trivial, but the SRS query language allows you to build up more complex queries using the various operators. In addition expressions can be grouped using parentheses so that they are treated as a single entity. You will see examples of searches and ways of combining them, throughout this chapter.

8.3.3 Operands

Operands are the items upon which the expression is performed, e.g., the name of a databank, or a set, etc.

Table 8.4 Typical SRS operands.

Operand	Example	Meaning
Databank name	EMBL	Each databank has a unique name.
Set name	Q1	SRS gives each query a name which can be used when you want to perform an operation on the set of results from that query.
Index search	[embl- des: kinase]	This command initiates a search in one or more indices, of one or more databanks.
Expression	(Q1&Q2)	If an expression is enclosed in parentheses, it is treated as a single operand. Parentheses can be nested to any degree.
Parent	parent	This is a special operand that allows the conversion of a set of subentries into a set of entries (see section 8.4 “Entries and Subentries”, p. 173).

8.3.4 Operators

Operators tell SRS what to do with the operands, e.g., search for links between two operands. Table 8.5 shows a list of available operators. For more information on the types of operator see section “Logical Operators” (p. 169) and section “Link Operators” (p. 169). For information on the use of operators, see section 8.3.5 “Use of Operators to Combine Search Items” (p. 171).

Table 8.5 SRS query language operators.

Operator	Type	Meaning
	Logical	OR.
&	Logical	AND.
!	Logical	BUTNOT. This operator may need to be escaped in UNIX, using "!".
>	Link	Link, keeping items in the set to the right of the >.
<	Link	Link, keeping items in the set to the left of the >.
>^	Link	Get subtree defined by left operand (hierarchical links).
>_	Link	Get leaf entries of the subtree defined by left operand (hierarchical links).

Logical Operators

The logical operators, OR ($|$), AND ($\&$) and BUTNOT ($!$), can be used to combine search words in an index search, or to combine sets in a query.

The following figure illustrates the effects of the three operators in an expression of the form, A operator B.

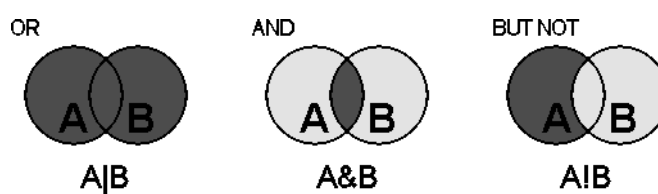


Figure 8.2 Logical operators in SRS.

Logical operations can only be performed between sets of the same type. It is not possible, for instance, to combine a set of entries and a set of subentries (see section 8.4 “Entries and Subentries”, p. 173) using logical operators. In such cases, an additional link operation must be specified (see section 8.4.1 “Links with Sets Containing Subentries”, p. 174).

Link Operators

Link operators are unique to the SRS query language. The two link operators, $<$ and $>$, allow sets of data from different databanks to be combined.

Figure 8.3 shows two databanks, A and B, in which some entries in A have links to entries in B. These links are processed to build link indices that provide the basis for the

link operation. The figure shows the results of two searches for links between sets A and B, using the operators, < and >.

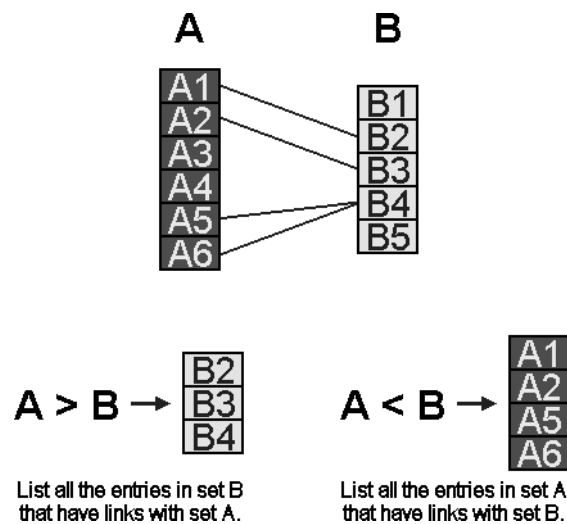


Figure 8.3 Linked Data.

Links are not usually bidirectional; however, the link indices in SRS are used bidirectionally. For instance:

A > B

This retrieves those entries in B that are linked to entries in A.

A < B

This retrieves those entries in A that are linked to entries in B.

8.3.5 Use of Operators to Combine Search Items

Combining queries allows you to refine your search results. This can be done using logical operators (OR, AND, BUTNOT, see section “Logical Operators”, p. 169) or link operators (see section “Link Operators”, p. 169). Note that link operators take precedence over logical operators.

Example 8.2 Simple linking

This example searches for links to a specified databank in the results of an index search:

```
[swissprot-des:kinase] > pdb
```

The result will be a list of all the entries in the PDB databank that have links to the results of the index search:

```
[swissprot-des:kinase]
```

See section 8.2 “Searching in Indices” (p. 157), for an explanation of index searches.

Example 8.3 Multiple linking 1

It is possible to combine several linking queries. For example:

```
[swissprot-id:acha_human] > prosite > swissprot
```

The search first retrieves the SWISS-PROT entry “acha_human”. This entry will have links to PROSITE entries that document the protein family of which “acha_human” is a member. (In this case it is the family of neuronal acetylcholine receptors.) These items in the PROSITE databank are retrieved. The next link retrieves SWISS-PROT entries that are linked to the PROSITE entries, i.e., belong to the neuronal

acetylcholine receptor family. In effect, the entry “acha_human” is being amplified to retrieve all the entries in SWISS-PROT which document members of the protein family or families to which it belongs.

Example 8.4 Multiple linking 2

A similar technique to that used in example 8.3 Multiple linking 1, can be used to find related information in another databank to which the initial entry is not linked.

```
[swissprot-id:gshr_caeel] > prodom > pdb
```

The query retrieves a probable glutathione reductase (whose **ID** is “gshr_caeel”) from SWISS-PROT, searches for entries in ProDom which document related proteins, and then looks for links to PDB in these entries. The result is a set of PDB protein structures that are homologous to the SWISS-PROT entry “gshr_caeel”.

Example 8.5 Complex linking

This example can be thought of as being composed of two parts:

```
(q = [{swissprot swissnew}-des:kinase])!(q<swissnew)
```

The first part of the query searches the `description` fields of the SWISS-PROT and SWISSNEW databanks, looking for “kinase”. The second part of the search excludes any entries that have links to the SWISSNEW databank. Using this technique, it is possible to retrieve the entries for “kinase”, but exclude any that are replaced by (more up to date) entries in SWISSNEW.

The distinction being made here is that entries in SWISSNEW will not be linked to themselves (or other entries in SWISSNEW), so all the SWISSNEW entries will be kept. However, any entry in SWISS-PROT that has been replaced by an entry (to which that entry is linked) in SWISSNEW will be picked up and rejected. In this way, out-of-date entries are excluded.

Note: The first part of the query also defines a group, α , which contains the SWISS-PROT and SWISSNEW databanks. This is used in the second part of the query rather than listing the databanks explicitly (see section 8.2.6 “Searching Multiple Databanks”, p. 164).

Example 8.6 Searching multiple databanks and screening for overlaps

Many protein or DNA databanks overlap to a great extent, which creates a lot of redundancy; however, the annotation of equivalent entries in different databanks can be quite varied. This can be useful for string searching because the probability of finding a certain enzyme name is greater if you can search all sequence databanks. After the search, links can be used to remove any overlaps.

See Example 8.5 Complex linking, for how this might be done.

8.4 Entries and Subentries

Sets originating from the same databank may have different set types. Consider the two queries:

```
[swissprot-keywords:transmembrane]
[swissprot-ftkey:transmem]
```

The first query retrieves all SWISS-PROT entries that include “transmembrane” in the `keywords` index. The `ftkey` index, however, has a special type, allowing it to find features of a given type within entries. Thus the second query retrieves a set of subentries within SWISS-PROT entries. Each subentry is a transmembrane feature.

Note: The second query above will retrieve many more entries than the first because most transmembrane proteins have more than one membrane spanning segment. If you requested the sequences for the entries in the second set, you would get the transmembrane segments rather than the parent entry’s sequence.

8.4.1 Links with Sets Containing Subentries

Simple Links

It is not possible to combine sets of entries with sets of subentries using the logical operators; however, link operators may be used between sets of entries and sets of subentries. For example:

```
[swissprot-org:human] > [swissprot-ftkey:transmem]
```

returns a set of transmembrane segment subentries found in human proteins, whereas

```
[swissprot-org:human] < [swissprot-ftkey:transmem]
```

returns all human proteins that have transmembrane segments.

Parent Links

Sometimes it is necessary to do an explicit conversion from subentries to entries. This can be done using the operand, `parent`. This method looks for links from the subentries to their respective parent entries and retrieves a set containing parent entries. For example:

```
[swissprot-ftkey:transmem] > parent
```

retrieves the parent entries (the entries to which the subentries belong) for the set of subentries from SWISS-PROT that have transmembrane sequence features. Logical operators can then be used to combine the set of parent entries with another set of entries. Furthermore, as with other SRS query language commands, it is possible to combine the link to the parent entries and the subsequent operations into the same command, see section “More Complex Links using the Parent Operand” (p. 175).

More Complex Links using the Parent Operand

The command:

```
[swissprot-ftkey:transmem] > parent |  
[swissprot-key:transmembrane]
```

returns all entries that have the keyword, “transmembrane”, or that have “transmem” sequence features. The index search `[swissprot-ftkey:transmem]` results in a set of subentries whereas the index search `[swissprot-key:transmembrane]`

returns a set of entries. These sets cannot be combined directly using a logical operator. Instead, an extra step must be added, which finds the set of parent entries to which the subentries belong. The resultant set of parent entries can then be combined with the set from the second index search using a logical operator.

This type of search may be necessary to ensure all entries with a certain property are retrieved because the annotation is often incomplete.

8.5 Storing Intermediate Results in Sets

If a query becomes very complicated, it may be convenient to store an intermediate result, into a set with a name which can be used later in the query. This is particularly useful to save typing out the query several times, but does not actually save a file. This example is an index query in EMBL that is then linked to both SWISS-PROT and SWISSNEW:

```
[embl-org:escherichia coli] > SWISSPROT |  
[embl-org:escherichia coli] > SWISSNEW
```

The EMBL index query is specified twice, however, it is possible to store the result of the index query in a set, e.g., `temp`, which saves duplicate typing of the index query. The assignment operation must be within parentheses, as shown:

```
(temp=[embl-org:escherichia coli])>prosite |  
temp > SWISSNEW
```


INDEX

- (index search) 157

Symbols

(range search) 157

* (wildcard) 59

: (string search) 157

? (wildcard) 59

A

accessing

analysis tools 138

administration

projects 47

analysis tools 137

using 138

what are they 138

application invocation page 25, 29, 30

application select page 21

applications

batch 142

invocation page 25, 29, 30

parameters 23

select page 21

applications. See analysis tools

apply

views 111

B

batch

applications 142

applications

batch queued 24

not batch queued 27

browse

index 86

C

characters

special

databank group names 164

filenames 35, 51, 53, 131

passwords 44

usernames 44

command line 149

copy

project information 49

creating

views 123

D

databank group names

spaces and special

characters 164

databanks

search multiple 164

dates

in indices, search for 57

query language, search for 163

delete 52

dialog

file download 33

save as 34

download

file download dialog 33

download options page 32

E

entries 173

find all entries, linking 98

linking 174

entry page

linking from 103

examples

expression linking 107

getz 150

linking 171, 172, 173

regular expressions 161

expression

expression linking 104

examples 107

procedure 104

- expression query 79
 - using 80
- expressions
 - regular expressions 59
 - examples 161
 - query language 159
- extended query form 65
- F**
- field information page 88
- file download dialog 33
- filenames
 - spaces and special characters 35, 51, 53, 131
- forms
 - extended query 65
 - query 62
 - standard query 8, 62
- G**
- getz 149, 150
 - arguments 151
 - examples 150
- H**
- how to
 - expression linking 104
 - linking 95
- I**
- index
 - browse 86
 - search
 - 157
 - query language 157
 - syntax 157
- information
 - field information page 88
- interactive
 - tools 141
- intermediate results
 - storing 176
- J**
- job status
 - page 145
- L**
- language
 - query language 155
- link page 11
- linking 91, 92
 - entries and subentries 174
 - examples 171, 172, 173
 - expression linking 104
 - examples 107
 - procedure 104
 - find all entries options 98
 - from entry page 103
 - from query manager page 99
 - from query result page 101
 - link page, finding 99
 - operators 105
 - query language 169
 - parent 175
 - procedure 95
- list view 124, 125, 128
- logical operators
 - query language 169
- M**
- manage
 - manage your query results
 - page 31
- managing
 - project manager page 47
 - projects
 - permanent 46
 - temporary 42
- multiple
 - databanks, searching 164

N

- non-secure permanent projects 43
 - starting 44
- numbers
 - in indices, search for 57
- numerical ranges
 - combining 163
 - searching, query language 161

O

- open 52
- operands 167
- operators
 - linking 105, 169
 - logical 169
 - query language 168
 - regular expressions 160
 - using 171
- order
 - sorting results 82
 - overview 10

P

- page
 - application invocation 25, 29, 30
 - application select 21
 - download options 32
 - entry 103
 - extended query form 65
 - field information 88
 - job status 145
 - link 11
 - manage your query results 31
 - query manager 17, 20, 99
 - query result 10, 101
 - standard query form 62
 - start 2
 - view manager 1 14
 - view manager 2 15
- pages
 - query result 5

parameters

- applications 23
- parent
 - links 175
- passwords
 - spaces 44
 - special characters 44
- permanent projects 38
 - managing 46
 - non-secure 43
 - secure 43
 - starting 44

project

- starting 2
- projects
 - administration 47
 - permanent 38
 - managing 46
 - non-secure 43
 - secure 43
 - starting 44
 - project manager page 47
 - starting 39
 - switching 48
 - temporary 38, 41
 - managing 42
 - starting 41

Q

- query 55
 - expression query 79
 - manage your query results
 - page 31
 - sorting results 82
 - overview 10
 - using expression queries 80
- query forms 62
 - extended 65
 - standard 8, 62
- query language 155

- operands 167
 - operators 168
 - linking 169
 - logical 169
 - regular expressions 159
 - search
 - combining 165
 - dates 163
 - index 157
 - multiple databanks 164
 - numerical ranges 161
 - syntax 165
 - what is it 156
 - wildcards 159
 - query manager page 17, 20
 - linking from 99
 - query result page 5, 10
 - linking from 101
 - quick search 60
 - using 61
- R**
- range search
 - # 157
 - ranges
 - combining 163
 - numbers and dates 57
 - regular expressions 59
 - examples 161
 - query language 159
 - rename 53
 - results
 - sorting 82
 - overview 10
- S**
- save 33, 50
 - save as dialog 34
 - search
 - dates, query language 163
 - index, query language 157
 - numerical ranges, query language 161
 - query language
 - dates 163
 - index 157
 - numerical ranges 161
 - quick search 60
 - using 61
 - strings
 - query language 158
 - terms 56
 - dates 57
 - multiple word 57
 - numbers 57
 - regular expressions 59
 - single word 56
 - wildcards 59
 - secure permanent projects 43
 - starting 44
 - select
 - application select page 21
 - sets 176
 - single word
 - search terms 56
 - sort
 - results 82
 - overview 10
 - spaces
 - databank group names 164
 - filenames 35, 51, 53, 131
 - passwords 44
 - usernames 44
 - special characters
 - databank group names 164
 - filenames 35, 51, 53, 131
 - passwords 44
 - usernames 44
 - standard query form 8, 62
 - start

- permanent projects 44
- project 39
- start page 39
- temporary projects 41
- start page 2
- string
 - search, query language 158
- string search
 - : 157
- subentries 173
 - linking 174
- switch
 - projects 48
- syntax
 - index search 157

T

- table view 124, 125, 128
- temporary projects 38, 41
 - managing 42
 - starting 41
- tools
 - interactive 28, 141
- tools. See analysis tools

U

- UNIX 150
- upload 52
- usernames
 - spaces 44
 - special characters 44
- using
 - analysis tools 138
 - getz 151

- operators 171
- views 110

V

- view
 - list view 124, 125, 128
 - table view 124, 125, 128
- view manager page 1 14
- view manager page 2 15
- views 109
 - applying 30, 111
 - creating 123
 - using 110

W

- what are
 - analysis tools (tools) 138
- what is
 - linking 92
 - query language 156
 - view 110
- white spaces
 - databank group names 164
 - filenames 35, 51, 53, 131
 - passwords 44
 - usernames 44
- wildcards 59
 - * 59
 - ? 59
 - query language 159
- word
 - multiple word search terms 57
 - single word search terms 56

