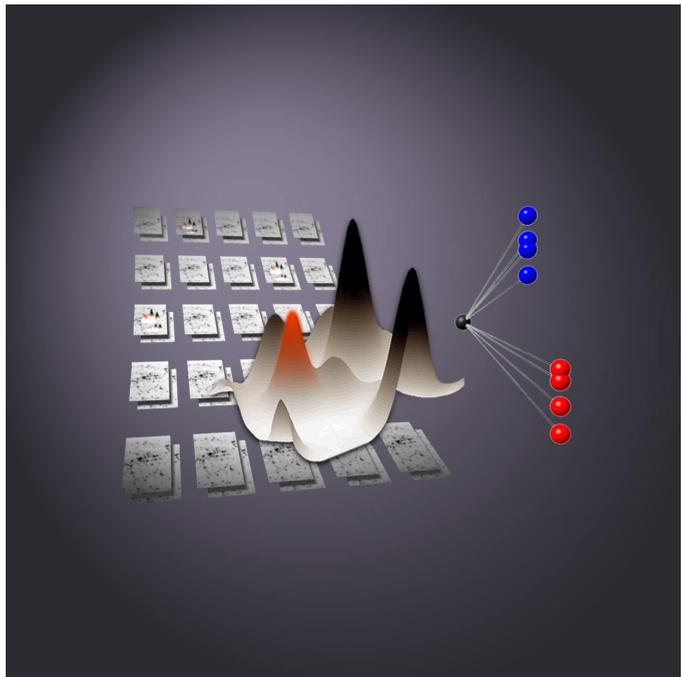


DeCyder Differential Analysis Software, Version 5.0

User Manual



Terms and Conditions of Sale

Unless otherwise agreed, all goods and services are sold subject to the terms and conditions of sale of the company within the Amersham Biosciences group which supplies them. A copy of these terms and conditions is available on request.

Trademarks

Amersham and Amersham Biosciences are trademarks of Amersham plc.

Cy, CyDye, DeCyder, Drop design, Ettan, Immobiline, IPGphor, PlusOne, Sierra and Typhoon are trademarks of Amersham Biosciences Limited.

HASP is a trademark of Aladdin Knowledge Systems Ltd.

Microsoft, Windows, Word, Excel, Internet Explorer are trademarks of Microsoft Corporation.

OpenGL is a trademark of Silicon Graphics, Inc.

Pentium is a trademark of Intel Corporation or its subsidiaries.

Spotfire is a trademark of Spotfire Inc.

SYPRO is a trademark of Molecular Probes Inc.

Patents and Licences

CyDye: 2-D Fluorescence Difference Gel Electrophoresis (2-D DIGE) technology is covered by US patent numbers US6,043,025, US6,048,982, US6,127,134, and US6,426,190 and foreign equivalents and exclusively licensed from Carnegie Mellon University.

CyDye: this product or portions thereof is manufactured under licence from Carnegie Mellon University under US patent number US5,268,486 and other patents pending.

The purchase of CyDye fluors includes a limited license to use the CyDye fluors for internal research and development, but not for any commercial purposes. A license to use the CyDye fluors for commercial purposes is subject to a separate license agreement with Amersham Biosciences.

Amersham Biosciences has patent applications pending relating to its DeCyder software technology, European patent application number EP1,234,280.

Office Addresses

Amersham Biosciences AB

Björkgatan 30
SE-751 84 Uppsala
Sweden

Amersham Biosciences UK Limited

Pollards Wood, Nightingales Lane
Chalfont St Giles
Buckinghamshire
HP8 4SP
UK

Amersham Biosciences Corp.

800 Centennial Avenue
P.O. Box 1327 Piscataway N.J. 08855-1327
USA

Amersham Biosciences Europe GmbH

Munzinger Strasse 9
D-79111 Freiburg
Germany

Amersham Biosciences KK

Sanken Building
3-25-1 Hyakunincho, Shinjuku-ku
Tokyo
Japan

©Amersham Biosciences AB 2003

-All rights reserved

Contents

1 Introduction to DeCyder Differential Analysis Software

1.1	Introduction	11
1.2	The DeCyder Differential Analysis Software User Manual	11
1.3	Measuring differential protein abundance using Ettan DIGE system	12
1.4	Experimental Design using an Internal Standard	14
1.5	Integration of DeCyder Differential Analysis Software with Ettan DIGE system experimental design	17
1.6	Steps involved in Image Analysis using DeCyder Differential Analysis Software	19
1.7	Structure of DeCyder Differential Analysis Software	19

2 Computer requirements and installation

2.1	Computer requirements	23
2.2	Installation	23
2.2.1	<i>Installation with a previous version present</i>	24
2.2.2	<i>De Novo Installation</i>	25

3 DIA (Differential In-gel Analysis) Module

3.1	Overview	27
3.2	Creating and opening workspaces	29
3.3	Spot detection and quantitation	31
3.3.1	<i>Detection</i>	31
3.3.2	<i>Quantitation</i>	33
3.4	Viewing spot data	34
3.4.1	<i>Image View</i>	34
3.4.2	<i>3-D View</i>	36
3.4.3	<i>Table View</i>	38
3.4.4	<i>Histogram View</i>	40
3.5	Data analysis	42
3.5.1	<i>Spot exclusion</i>	43
3.5.2	<i>Spot confirmation</i>	46
3.5.3	<i>Protein of Interest</i>	48
3.5.4	<i>PTM assignment</i>	48
3.6	Customizing display colors	48
3.7	Saving, exporting and printing	49

4 BVA (Biological Variation Analysis) Module

4.1 Overview	51
4.1.1 <i>Functionality</i>	51
4.1.2 <i>Structure</i>	52
4.2 Protein quantitation	53
4.3 Creating and opening workspaces	53
4.3.1 <i>Creating workspaces</i>	53
4.3.2 <i>Opening workspaces</i>	54
4.4 Viewing spot data	54
4.4.1 <i>Image View</i>	54
4.4.2 <i>3-D View</i>	56
4.4.3 <i>Table View</i>	57
4.5 Defining spot map attributes	58
4.5.1 <i>Function assignment</i>	58
4.5.2 <i>Group assignment</i>	59
4.5.3 <i>Comment assignment</i>	61
4.5.4 <i>Spot Map Table Mode</i>	61
4.5.5 <i>Spot Map Table properties</i>	62
4.6 Inter-gel matching	63
4.6.1 <i>Landmarking</i>	63
4.6.2 <i>Matching</i>	65
4.6.3 <i>Match confirmation</i>	65
4.6.4 <i>Match Quality Metric</i>	67
4.6.5 <i>Spot Merging</i>	68
4.6.6 <i>Match Table</i>	70
4.6.7 <i>Match Table properties</i>	70
4.7 Graphical representation	71
4.8 Protein statistics	73
4.8.1 <i>Opening the protein statistics dialog box</i>	73
4.8.2 <i>Defining groups</i>	74
4.8.3 <i>Overview of Statistical tests</i>	75
4.8.4 <i>Independent and paired analyses</i>	76
4.8.5 <i>Student's T-test & Average ratio</i>	78
4.8.6 <i>ANOVA</i>	81
4.8.7 <i>One-Way ANOVA</i>	83
4.8.8 <i>Two-Way ANOVA</i>	85
4.8.9 <i>Further statistical analyses</i>	90
4.9 Protein Filter	90
4.10 Molecular weight (Mw) and isoelectric point (pI)	91
4.10.1 <i>Entering Mw and pI of known proteins</i>	91
4.11 User-defined protein labelling	94
4.11.1 <i>Confirmation</i>	94
4.11.2 <i>Name</i>	94
4.11.3 <i>Comment</i>	94
4.11.4 <i>Protein of interest</i>	94
4.11.5 <i>PTM</i>	94
4.12 Database linking	94
4.12.1 <i>Adding databases</i>	95

4.13 Viewing protein data	98
4.13.1 Protein Table	98
4.13.2 Appearance Table.....	101
4.13.3 Spot annotation.....	102
4.13.4 Customizing display colors	103
4.13.5 Saving, exporting and printing.....	104

5 Spot picking

5.1 Overview	107
5.2 Spot detection on the preparative gel	107
5.3 Identification of reference markers	108
5.4 Identifying Proteins of Interest	109
5.4.1 Identifying proteins for picking using the DIA module	110
5.4.2 Identifying proteins of interest using the BVA module	112
5.5 Assigning Spots for Picking	114
5.6 Editing Pick Locations	116
5.7 Generating a Pick List	119

6 Batch processor

6.1 Overview	121
6.2 Batch list creation	121
6.2.1 Setting up the DIA Batch List	122
6.2.2 Setting up the BVA Batch List	125
6.3 Editing the batch list	126
6.3.1 DIA batch list.....	126
6.3.2 BVA batch list.....	126
6.4 Saving Data	127
6.5 Running the batch processor	128

7 XML Toolbox

7.1 Overview	129
7.2 Opening the XML Toolbox module	129
7.3 Extracting data	130
7.4 Tag definitions	132
7.4.1 BVA parameters.....	132
7.4.2 DIA parameters.....	135

8 LWS Integration

8.1 Overview	139
8.2 Ettan LWS with samples generated within Ettan DIGE system ...	140
8.3 Enter a protein tube generated within Ettan DIGE system, into Ettan LWS	142
8.4 Enter a gel, prepared for spot picking within Ettan DIGE system, into Ettan LWS	144

8.5 XML toolbox - LWS Pick List	146
8.5.1 Preparations before XML conversion	146
8.5.2 Using the LWS Pick List tool	147
8.5.3 After XML conversion.....	148
8.6 XML Toolbox - Protein Identification	148
8.6.1 Preparations in Ettan LWS before XML conversion.....	149
8.6.2 Using the Protein Identification tool.....	150
8.7 Importing protein data templates into the BVA module	151
8.8 Database queries in DeCyder BVA	151

9 Tutorials Introduction

9.1 DeCyder Differential Analysis Software structure	155
9.2 Scope of tutorials	157

10 Tutorial I - Using DIA module for preliminary investigation of protein changes

10.1 Objective	159
10.2 Overview	159
10.3 Experimental design	160
10.4 Analysis of control-control gel	160
10.4.1 Selecting gel images.....	160
10.4.2 Spot detection	162
10.4.3 Assigning an area of interest.....	165
10.4.4 Gel artifact removal.....	166
10.4.5 Ascertaining the 2 S.D. (standard deviation) threshold value.....	168
10.5 Analysis of control-treated gel	168
10.5.1 Selecting gel images.....	168
10.5.2 Spot detection	168
10.5.3 Spot filtering.....	168
10.5.4 Setting a threshold.....	169
10.6 Assigning protein of interest	170
10.6.1 Selecting spots for picking.....	170
10.6.2 Spot confirmation	171
10.6.3 Exporting the Pick List and physically excising spots from the gel.....	172

11 Tutorial II - Employing an internal standard to Analyze Protein Changes

11.1 Objectives	173
11.2 Overview	173
11.3 Experimental design	174
11.4 Spot detection and quantitation	175
11.4.1 Selecting gel images.....	175

11.4.2	<i>Spot detection and quantitation</i>	176
11.4.3	<i>Viewing the DIA workspace</i>	177
11.4.4	<i>Exporting spot map data</i>	179
11.5	Creating the BVA workspace.....	180
11.5.1	<i>Selecting gels</i>	180
11.5.2	<i>Function assignment</i>	182
11.5.3	<i>Experimental group assignment</i>	183
11.5.4	<i>Landmarking</i>	184
11.6	Assessing matching.....	186
11.7	Post-matching landmarking.....	189
11.8	Statistical analysis.....	190
11.9	Assigning spots as proteins of interest.....	193
11.9.1	<i>Selecting proteins of interest</i>	193
11.9.2	<i>Spot confirmation in Protein Table</i>	195
11.9.3	<i>Exporting the Pick List and physically excising spots from the gel</i>	195

12 Tutorial III - Processing the Preparative Gel and Generating a Pick List

12.1	Objective.....	197
12.2	Overview.....	198
12.3	Experimental design.....	198
12.3.1	<i>Identifying protein spots on SYPRO Ruby stained gel</i>	199
12.3.2	<i>Spot detection</i>	200
12.3.3	<i>Assigning an area of interest</i>	203
12.3.4	<i>Gel artifact removal</i>	206
12.4	Matching to analytical gels.....	208
12.4.1	<i>Preparing the workspace</i>	208
12.4.2	<i>Editing picking reference markers</i>	211
12.4.3	<i>Matching and landmarking</i>	212
12.4.4	<i>Exporting the Pick List</i>	216

13 Tutorial IV - Fully automated identification of differentially expressed proteins

13.1	Objective.....	217
13.2	Overview.....	217
13.3	Experimental design.....	218
13.4	Protein spot filtering (optional).....	219
13.4.1	<i>Selecting gel images</i>	220
13.4.2	<i>Spot detection</i>	221
13.4.3	<i>Gel artifact removal</i>	224
13.5	Processing multiple images.....	226
13.5.1	<i>Setting up the batch processor</i>	226
13.5.2	<i>Assignment of spot map attributes</i>	229
13.5.3	<i>Protein statistics</i>	231

13.5.4	<i>Protein Filter</i>	232
13.5.5	<i>Saving the Pick List</i>	233
13.5.6	<i>Saving the batch workspace</i>	234
13.5.7	<i>Running the batch processor</i>	235

Appendices

Appendix A:	Recommended workflow for analysis of multiple gels...	237
Appendix B:	Understanding the digital image.....	243
Appendix C:	Spot processing algorithms	251
Appendix D:	Experimental design and set up examples	257
Appendix E:	DeCyder Differential Analysis Software keyboard shortcuts	267
Appendix F:	Related products and documentation.....	271
Appendix G:	Glossary.....	275

1 Introduction to DeCyder Differential Analysis Software

1.1 Introduction

Two-dimensional electrophoresis (2-D electrophoresis) is a leading tool in proteomics research today, capable of visualizing many components of complex proteomes in a single gel. Ettan™ DIGE (Difference Gel Electrophoresis) system is a method for pre-labelling protein samples prior to 2-D electrophoresis. The system is based upon the specific properties of CyDye™ DIGE Fluor dyes which enable multiplexing of separate protein mixtures on the same 2-D gel.

DeCyder™ Differential Analysis Software is an automated image analysis software suite which enables detection, quantitation, matching and analysis of Ettan DIGE system gels. The software was developed as part of Ettan DIGE system, to exploit the multiplexing capabilities of the CyDye DIGE Fluor dyes. Multiplexing, the co-migration of more than one sample per gel, enables the inclusion of an internal standard. The internal standard is used to derive statistical data within and between gels. This experimental design using the internal standard, effectively eliminates gel-to-gel variation, allowing detection of small differences in protein levels to be achieved. Using DeCyder Differential Analysis Software, system variability is minimized enabling expression differences identified by 2-D DIGE to be confidently assigned to induced biological change.

1.2 The DeCyder Differential Analysis Software User Manual

This user manual is broadly divided into two main parts, the reference manual (Chapters 1-8) and the tutorials (Chapters 9-13).

It is recommended that new users first work through the tutorials, in order to gain a rapid understanding of the software's capabilities. The tutorials are step by step guides that take the user through the main applications of the software by employing real data. A CD containing the necessary files is required to perform each of the tutorials. The tutorials are designed to be worked through without prior knowledge of the reference component of the manual.

The reference manual provides a detailed technical account encompassing all aspects of the built-in functionality of DeCyder Differential Analysis Software, which can be used as a source of further information for experienced users.

1.3 Measuring differential protein abundance using Ettan DIGE system

To compare protein abundance in different samples, conventional 2-D methods require the separation of each sample on an individual gel. This “one-sample-per-gel” approach exposes the data to a high level of system variation, i.e. the variation that arises from differences in protein uptake into the first dimension strip, second dimension gel running, etc. This high level of system variation can outweigh the often subtle, induced biological changes that the experiment is intended to detect, for example, differences that are caused by a disease state, drug treatment or life-cycle stage.

To compound this problem, it is also necessary to separate the induced biological changes within an experiment from the inherent biological variation, i.e. the differences between two individual animals, cultures, plants or flies, that are present irrespective of the applied experimental test conditions. To achieve this, multiple sample replicates must be incorporated within each experimental design. This requires the separation and analysis of a large number of samples and can be a slow process if each sample is separated on a different gel.

Ettan DIGE system has been developed to address these problems. The system includes CyDye DIGE Fluor CyTM2, Cy3 and Cy5 minimal dyes, which are mass and charge-matched, spectrally resolvable fluorescent dyes. Additional CyDye DIGE Fluor Cy3 and Cy5 saturation dyes have been developed specifically to be used where only small amounts of sample are available. Please refer to section F.1 for ordering details of the Scarce Sample Labelling kit. A protein sample labelled with any of the CyDye DIGE Fluor dyes, will migrate to the same position on a 2-D gel. This permits the multiplexing of two or three samples within the same 2-D gel, allowing the inclusion of an internal standard (see section 1.4).

Gels are scanned using the TyphoonTM Variable Mode Imager, generating overlaid, multi-channel images for each gel. Images can then be analyzed using the DeCyder Differential Analysis Software, which contains novel algorithms for co-detection of multiplexed gel images and has been specifically developed for use with Ettan DIGE system (see section 1.4).

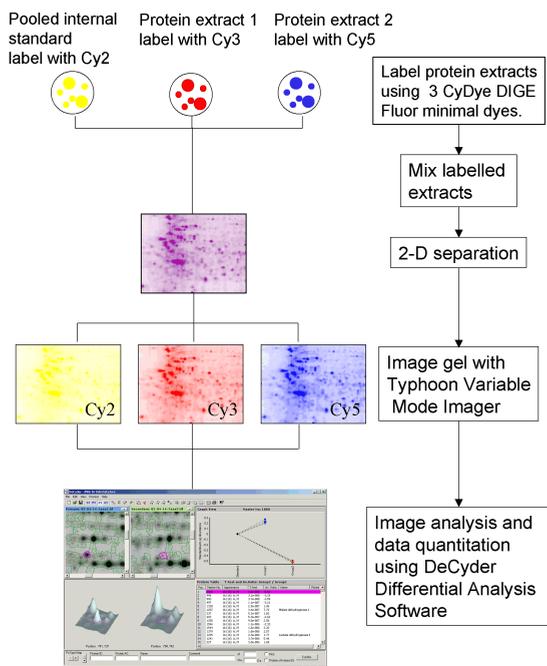


Fig 1-1. Scheme showing the workflow for Ettan DIGE system

The benefits offered by Ettan DIGE system are:

- Accurate quantitation and statistical analysis of protein abundance changes
- High sensitivity and wide dynamic range (5 orders of magnitude)
- Minimization of system (gel-to-gel) variation
- Easier matching between gels, with increased confidence
- Fewer gels required per experiment
- Faster analysis due to fully automated gel-processing workflow

1.4 Experimental Design using an Internal Standard

Ettan DIGE system provides the ability to multiplex samples, enabling the use of an internal standard within each 2-D gel. Ideally, the internal standard should consist of a pool taken from all of the samples within the experiment. The internal standard is labelled with one of the CyDye DIGE Fluor minimal dyes (usually Cy2 if using CyDye DIGE Fluor minimal dyes or Cy3 if using CyDye DIGE Fluor saturation dyes) and run on each gel in the experiment. This creates an image that is the average of all experimental samples, with all proteins in the experiment represented. The presence of the internal standard in every gel provides an intrinsic link between samples. Ettan DIGE system is currently the only 2-D gel electrophoresis protein difference analysis technique to utilize the internal standard approach.

There are several benefits of using an internal standard in 2-D experiments. Firstly, each protein spot in a sample can be compared to its representative within the internal standard on the same gel, to generate a ratio of relative protein levels. Quantitative comparisons of samples between gels are made based on the relative change of sample to its in-gel internal standard. This process effectively removes the system (gel-to-gel) variation enabling accurate quantitation of induced biological change between samples (Fig. 1-2). The need to run gel replicates is also overcome, reducing the number of gels required per experiment.

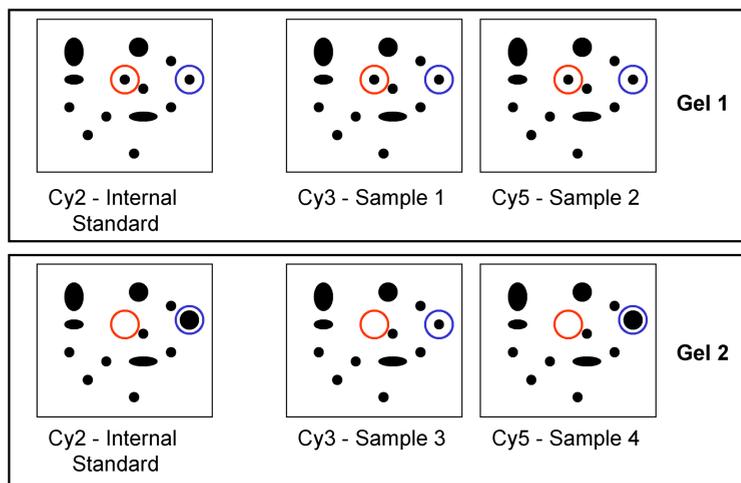


Fig 1-2. Analysis of samples 1-4 in the absence of an internal standard would suggest that the protein spot circled in red on gel 1 is absent in samples 3 and 4. However, reference to the internal standard, which is an identical pool of all samples run on both gels, shows that this protein has not entered gel 2. This indicates that the observed absence of this protein spot in samples 3 and 4 is due to system variation (e.g. gel distortions, differences in first dimension focusing etc.) and not to sample differences. Similarly, analysis without the internal standard of the protein spot circled in blue suggests that this protein is present in a greater abundance in sample 4. Reference to the internal standard indicates that system variation has resulted in this protein having an increased volume in gel 2 relative to gel 1. Hence the abundance of this protein is unchanged in samples 1,2 and 4 and decreased in sample 3.

A further benefit of using an internal standard is that matching between gels is more straightforward. The internal standard image is common between all gels in an experiment, therefore matching can be performed between internal standard images which have the similar spot patterns. Conventional 2-D electrophoresis requires matching between different samples on different gels, which introduces differences in spot patterns from sample-to-sample and gel-to-gel variation. Matching between internal standards allows matching between identical samples, so variations in spot patterns are due only to electrophoretic differences.

The internal standard approach can be applied to two-color or three-color experiments by including the internal standard plus one sample or two samples respectively on each gel.

To maximize the quality of the data obtained when working with Ettan DIGE system, the correct experimental design should be implemented. Table 1-1 shows an example of the experimental design used for a simple three-color experiment using control and treated groups, each containing four individuals

<i>Gel</i>	<i>Cy2</i>	<i>Cy3</i>	<i>Cy5</i>
1	Pooled internal standard	Control A	Treated B
2	Pooled internal standard	Control B	Treated C
3	Pooled internal standard	Treated D	Control C
4	Pooled internal standard	Treated A	Control D

Table 1-1. Using an internal standard for accurate measurement of protein abundance changes between samples in an experiment (for an experiment using >2 samples).

Each gel contains a pooled internal standard labelled with CyDye DIGE Fluor Cy2 minimal dye. Four biological replicates (A-D) have been included for control and treated samples that have each been labelled with CyDye DIGE Fluor Cy3 or Cy5 minimal dyes. A greater degree of statistical confidence can be assigned to the experimental results by increasing the number of biological replicates employed. Half of the control group are labelled using the CyDye DIGE Fluor Cy3 minimal dye, half using the CyDye DIGE Fluor Cy5 minimal dye and similarly for the labelling of the treated group, thereby confirming to best experimental practices.

For more detailed information about Ettan DIGE system and using an internal standard, please refer to Ettan DIGE system User Manual (code no. 18-1173-17).

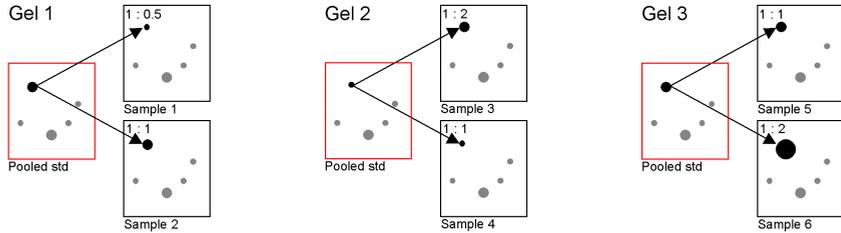
1.5 Integration of DeCyder Differential Analysis Software with Ettan DIGE system experimental design

DeCyder Differential Analysis Software was developed specifically for the 2-D DIGE methodology and therefore all the advantages of this approach are utilized in the software.

- DeCyder Differential Analysis Software in conjunction with Ettan DIGE system allows the analysis of experimental designs with various degrees of complexity. A simple control – treated experiment through to a multi-condition experiment addressing factors such as dose and time can be performed in a single analysis.
- The relationship between any number of samples can be accurately quantified and statistically analyzed in DeCyder Differential Analysis Software by employing the internal standard (see Fig. 1-2). This approach results in unparalleled accuracy allowing experimental conclusions to be drawn with high confidence. No other 2-D electrophoresis technique available is capable of resolving multiple samples in this manner and hence Ettan DIGE system is unique in exploiting use of an internal standard on every gel.
- The novel co-detection algorithm exploits the identical spot patterns generated when multiple samples are resolved on the same gel. The algorithm generates identical spot detection patterns on all images derived from the same gel. Hence all spots on the same gel are effectively matched with the identical spot boundaries.
- Spot quantitation is performed automatically by normalizing spot volumes against the internal standard (Fig. 1-3). The co-detection algorithm ensures that the internal standard and the quantified analytical spot have an identical spot boundary. This results in a highly accurate and robust protein quantitation.
- DeCyder Differential Analysis Software utilizes experimental design incorporating an internal standard and performs gel to gel matching on the standard samples.

A. Differential In-gel Analysis (DIA)

In-gel spot matching and determination of ratios, standard : sample



B. Biological Variation Analysis (BVA)

Quantitative comparison of protein abundance across multiple gels

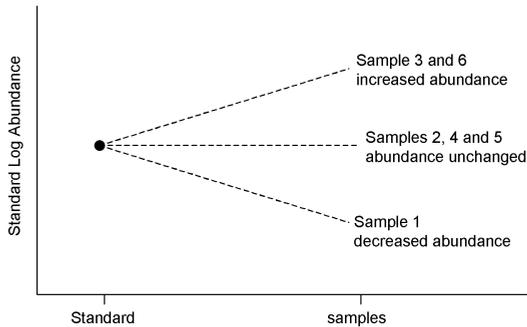


Fig 1-3. A. DeCyder Differential Analysis Software utilizes an internal standard to a) aid spot matching between samples within the same gel and b) generate a ratio of protein abundance between the proteins of the internal standard and each sample within the same gel. Because the internal standard is the same sample run within each gel, this effectively normalises all the data. These functions are performed within the Differential In-gel analysis (DIA) module of the software.

B. The Biological Variation (BVA) module is used to provide a quantitative comparison of protein abundance between all samples within the experiment.

1.6 Steps involved in Image Analysis using DeCyder Differential Analysis Software

Image analysis performed using DeCyder Differential Analysis software is performed using a number of complex algorithms, some of which are patent pending. They have been designed specifically for use with multiplexed 2-D images. These steps can be broken into the following processes:

- Spot detection
- Background subtraction
- In-gel normalization
- Gel artifact removal
- Gel to gel matching
- Statistical analysis

The complex algorithms associated with these steps form part of the in-built functionality of the DeCyder Differential Analysis Software. The various stages of gel processing are performed by different modules within the software suite.

1.7 Structure of DeCyder Differential Analysis Software

The software consists of four modules (see Fig. 1-4):

DIA (Differential In-gel Analysis) – Protein spot detection and quantitation on a set of images, from the same gel.

BVA (Biological Variation Analysis) – Matches multiple images from different gels to provide statistical data on differential protein expression levels between multiple groups.

Batch Processor – Fully automated image detection and matching of multiple gels without user interaction.

XML Toolbox – Extracts user specific data facilitating automatic report generation.

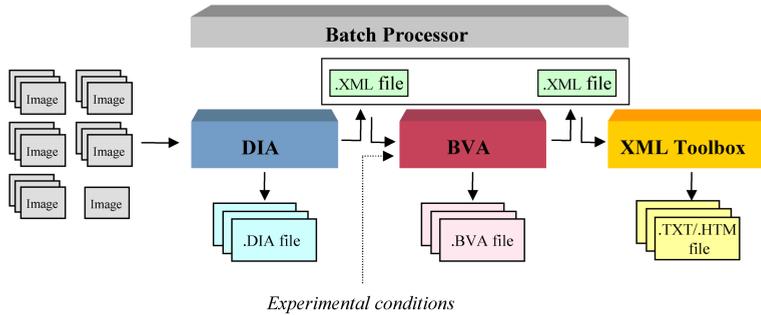


Fig 1-4. Structure of DeCyder Differential Analysis Software

DIA module

A set of images (each saved as either 16-bit TIFF or customized *.GEL* file) from a single gel are loaded then processed simultaneously. The DIA algorithms perform spot detection on a combined image derived from all loaded images. The protein spot quantitation is calculated and expressed as a volume ratio to the internal standard. The results can be saved as a *.DIA* file, which can be re-opened in the DIA module. This pick list is created from data generated in a single DIA module analysis (i.e. an experiment based on a single gel). Alternatively, the generated results can be saved in an XML format, or opened in the XML Toolbox module in order to extract specific data, or exported to the BVA module for multi-gel analyses (Fig. 1-5).

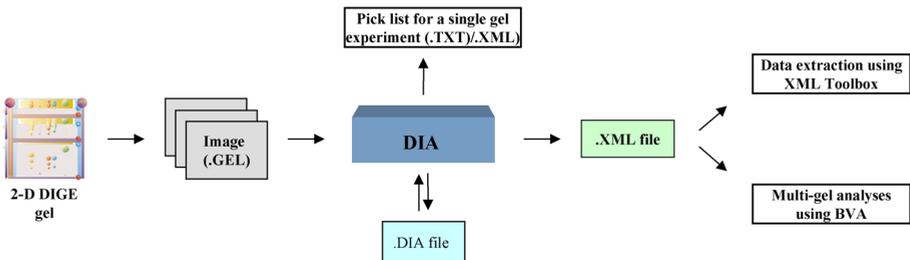


Fig 1-5. Schematic representation of DIA module workflow

BVA module

The BVA module utilizes the XML files exported from the DIA module together with the original gel images to match protein spots on different gels. Data is then subjected to statistical analyses to accurately assess protein changes occurring between control and test groups within the experiment. The results can be saved as .BVA file, which can be re-opened in the BVA module. Furthermore, data can be exported in XML format for data extraction using the XML Toolbox.

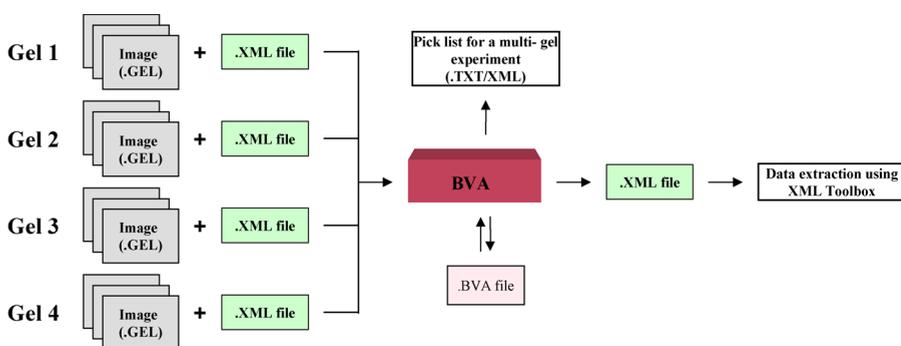


Fig 1-6. Schematic representation of BVA module workflow

Batch Processor

The Batch Processor integrates both the DIA and BVA modules enabling fully automated processing of multiple gels without user intervention. The Batch Processor can be configured to analyze several gels in the DIA module exclusively. Alternatively multiple gels can be processed through both modules to produce a final .BVA file and subsequent pick list.

XML Toolbox

The XML Toolbox enables the extraction of user specific data from the XML files generated in either the DIA or BVA modules. This data can be saved in either text or html format enabling users to access data in DeCyder Differential Analysis Software workspaces for other applications.

The XML Toolbox also provides an interface for linking Ettan DIGE system data with Ettan Laboratory Workflow System.

2 Computer requirements and installation

2.1 Computer requirements

- Operating System: Windows™ XP.
- Minimum Processor: Pentium™ 4 processor, 1.5 GHz.
- 1 Gbyte RAM.
- Video card capable of 32 bit color.
- The video card driver needs to support Open GL™ (v 1.2 or later) - ensure the latest compatible driver is installed.
- The color resolution of the PC should be set to 32 bit color.
- The screen resolution should be set to 1024 x 768 pixels, landscape (with 24 bits Z-buffer / preferably 32 bits).
- The virtual memory should be set so that the total amount of available memory, including physical RAM, is greater than 2.5 GB.
- Internet Explorer™ version 5.5 or higher must be installed to run the XML Toolbox module.

Note: Avoid running other programs at the same time as the various DeCyder Differential Analysis Software modules.

2.2 Installation

DeCyder Differential Analysis Software is protected by a HASP™ key that should be attached to either the USB or parallel port on the computer when DeCyder Differential Analysis Software is in use. The HASP device driver software is installed during installation of DeCyder Differential Analysis Software. The color resolution must be set to at least 32 bit color and the screen resolution should be set to 1024 x 768 pixels. These settings are accessed on the **Settings** tab in the **Display** icon on the Control Panel. Local administrative privileges must first be obtained for installation of DeCyder Differential Analysis Software.

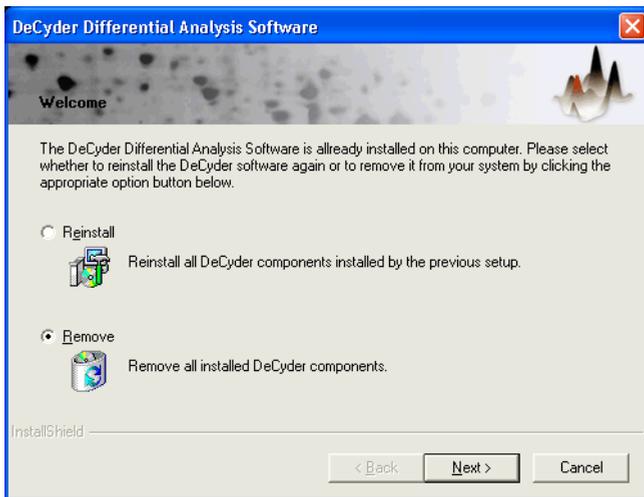


Insert the DeCyder Differential Analysis Software CD-ROM and select the appropriate disc drive in Windows explorer. From the files on the disc double click the **Setup** icon.

2.2.1 Installation with a previous version present

If a previous version of DeCyder Differential Analysis Software was installed select **Reinstall all DeCyder components installed by the previous setup** from the dialog box to remove the old version and automatically install the new version. Click **Next**.

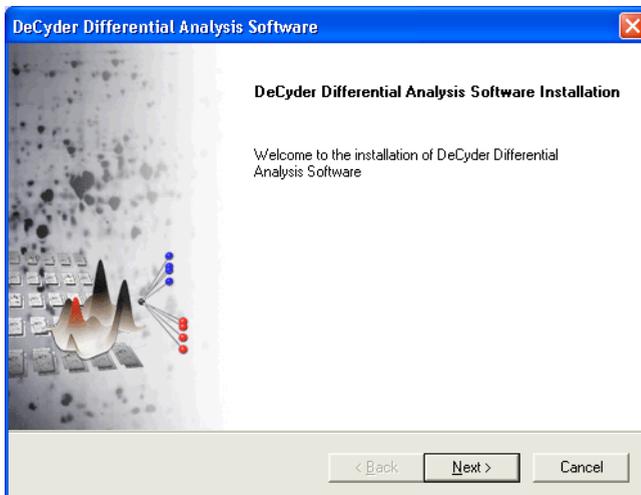
Note: If DeCyder Differential Analysis Software version 4 or earlier is already installed, the installation process will continue as a de novo installation (see section 2.2.2) automatically removing the previous version of the software.



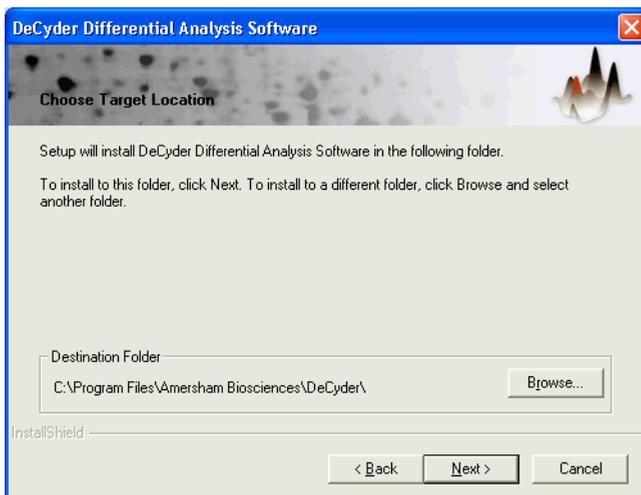
DeCyder Differential Analysis Software can be un-installed automatically by selecting **Remove all installed DeCyder components**.

2.2.2 De Novo Installation

If DeCyder Differential Analysis Software has not been previously installed on the system a different dialog box will appear when the **Setup** icon is selected. Click **Next**, to install DeCyder Differential Analysis Software.



In the subsequent dialog box, a default file path for the installation is given, if this is not suitable browse for an alternative path. Click **Next**, to install the software to the selected path.



The HASP key and DeCyder Differential Analysis Software will now be installed. When setup is complete click **Finish**.

Note: If DeCyder Differential Analysis Software is supplied pre-installed on a computer, no further software should be installed.

3 DIA (Differential In-gel Analysis) Module

3.1 Overview

DIA Processes

DeCyder Differential Analysis Software DIA processes images from a single gel, performing spot detection and quantitation.

The DIA module algorithms detect spots on a cumulative image derived from merging up to three individual images from an in-gel linked image set. This co-detection ensures that all spots are represented in all images processed. The DIA module algorithms then quantitate spot protein abundance for each image and express these values as a ratio thereby indicating changes in expression levels by direct comparison of corresponding spots. This ratio parameter can be used, in small scale experiments, to directly evaluate changes between two labelled protein samples. Alternatively, the ratio can be used for protein spot quantitation of a sample against an internal standard to allow accurate inter-gel protein spot comparisons (see Chapter 1). Once spot maps incorporating an internal standard have been analyzed in the DIA module, the spot data can be used in DeCyder Differential Analysis Software BVA for accurate quantitative inter-gel studies. Generally when multiple gel analyses are performed, only the spot detection and quantitation algorithms are utilized in the DIA module. Data is then transferred to the BVA for inter-gel analysis.

Data can be saved and re-opened in a .DIA file format from within the DIA module. The DIA module can also export for both Ettan Spot Picker or Ettan Spot Handling Workstation. Data can also be exported in an XML format for either multi-gel analysis in DeCyder Differential Analysis Software BVA, querying in DeCyder Differential Analysis Software XML toolbox or copying and pasting into applications such as Microsoft™ WORD™ and EXCEL™.

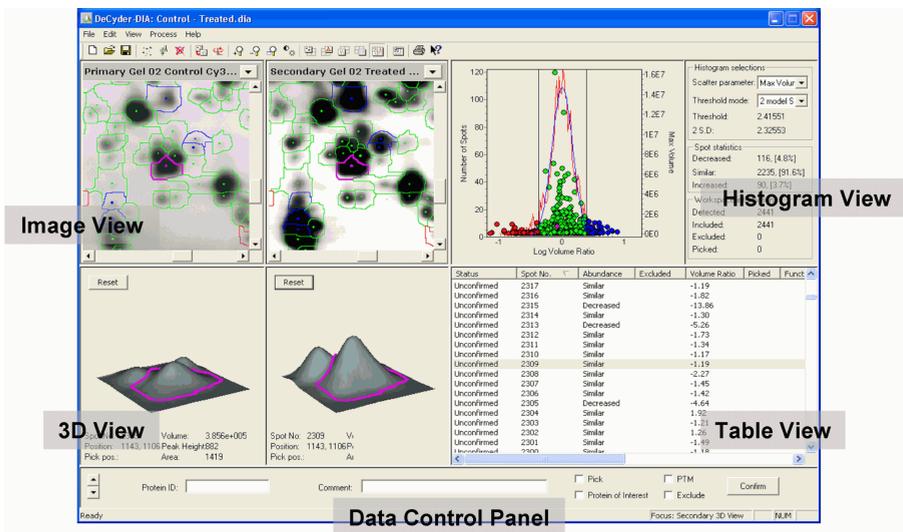
DIA Graphical User Interface

DeCyder Differential Analysis Software DIA graphical user interface is divided into four equally sized inter-linked views.

- **Image View** - primary and secondary gel images
- **3-D View** - a three dimensional representation of the gel localized on the spot
- **Histogram View** - graphical representation of data associated with the spots displayed in the image view
- **Table View** - tabulated data associated with selected spots displayed in the image view

All four views are linked, therefore selecting a spot in, for example the Image View, will display the spot in the Histogram View (in magenta), the 3-D View and the Table View. The role of the four views will be discussed in detail later in this section.

Below the four views the **Data Control Panel** is found. This panel contains tools and user definable functions associated with the specific data displayed in each mode.

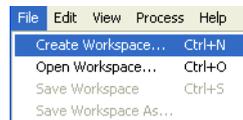


3.2 Creating and opening workspaces



Creating Workspaces

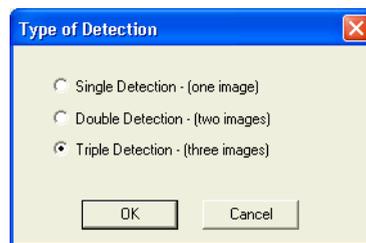
From the Desktop or **Start** button menu select the **DeCyder Differential Analysis Software - DIA** icon to open the software module. Create a new workspace by selecting **File:Create Workspace** in order to load the images for analyzing.



There are three distinct algorithms employed for image spot detection, which process different numbers of images simultaneously.

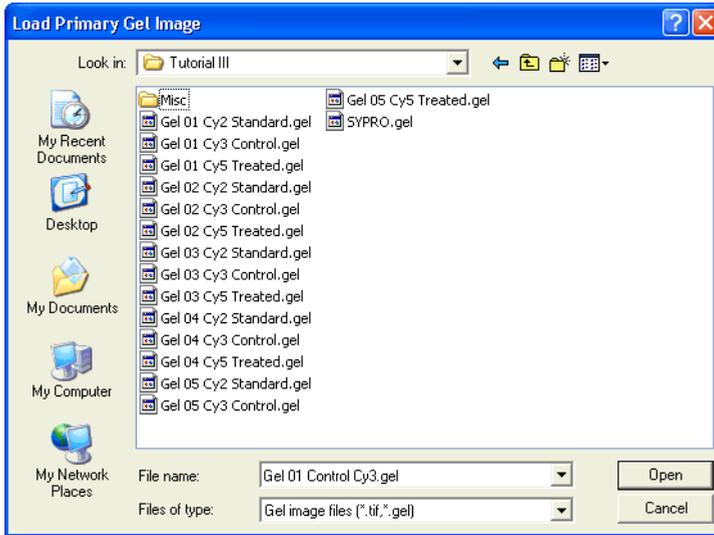
- Single detection - one image
- Double detection - two images
- Triple detection - three images

Select the detection algorithm appropriate for the number of images present on the gel being analyzed, then click **OK**.



In the resulting window click on the image file that is to become the primary image, and click **Open**.

***Note:** DeCyder Differential Analysis Software has been validated for file formats generated by Amersham Biosciences imaging devices recommended for Ettan DIGE system (see Appendix B).*



Repeat for the secondary and tertiary images if required.

Note: In an experiment in which an internal standard sample is being used, the standard sample image is designated as the primary image and the analytical samples are designated as the secondary and tertiary images.

Once the image(s) have been loaded spot detection and quantitation can be performed.

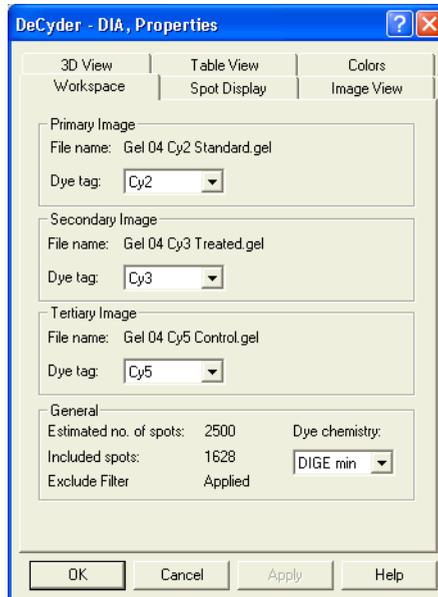
Opening Workspaces

To open previously created and saved workspaces select **File:Open** and browse to locate the **DIA** file. When the file is located select the file and click **Open** or double click on the **DIA** file.

Workspace Properties



Properties associated with the workspace are defined in the Workspace Properties window. Selecting **View:Properties...** (or selecting the **Properties** icon) then selecting the **Workspace** tab, displays the workspace properties window.



The dye tag used and the dye chemistry used can be selected from the pull down menus. The dye tag label is automatically selected if the information is included in the image file name. The remaining information is derived from the data processing performed.

3.3 Spot detection and quantitation

3.3.1 Detection

There are three distinct spot detection algorithms, which are able to simultaneously process a different number of images derived from a single gel.

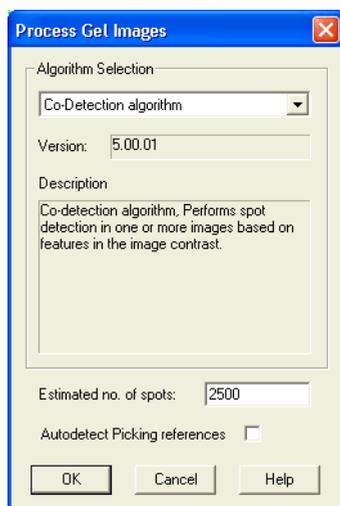
- Single detection - one image
- Double detection - two images
- Triple detection - three images

Single detection is performed on images of post-stained gels used for picking, a case where there is a single image associated with the gel.

Double and triple detection algorithms are designed to take advantage of the inherent co-migration benefits of the CyDye DIGE Fluor dyes. A set of co-run images (two images in double detection and three images in triple detection) is merged together thereby incorporating all spot features in a single image. Spot detection and spot boundary definition is then performed using pixel data from all the individual raw images and the merged image. The resultant spot map is overlaid back onto the original image files. Since the spot boundaries are identical for all images, the spots are effectively already matched. This process results in highly accurate volume ratio calculations.

When creating workspaces the user selects the applicable algorithm and is subsequently prompted to load the appropriate number of image files. Once the images are loaded the pre-selected spot detection algorithm can be applied.

Select **Process:Process Gel Images** to open the Process Gel Images window.



An estimation of the number of spots present on the images must be entered, before clicking **OK**. For example, a mammalian lysate run on an 18 cm pH 4-7 Immobiline™ DryStrip and a large format gel, such as the Ettan DALT electrophoresis unit (20 cm x 26 cm), a value of 2500 for Estimated Number of Spots should be satisfactory. If all the spots have not been identified the spot detection process can be repeated with a higher number of estimated spots. It is recommended that this value be overestimated to compensate for the detection of non-proteinaceous spots on the image, e.g., dust particles which are subsequently excluded from the analysis.

3.3.2 Quantitation

Numerical data for individual spots are calculated (e.g. volume, area, peak height and slope).

Spot volumes (sum of pixel intensities within the spot boundary) are always expressed with background subtracted. Background is subtracted on a spot specific basis, by excluding the lowest 10th percentile pixel value on the spot boundary, from all other pixel values within the spot boundary. The spot volume is the summation of these corrected values.

Spot ratios are calculated (volume of secondary image spot/volume of primary image spot). This ratio indicates the change in spot volume between the two images.

These ratio values are normalized, so that the modal peak of volume ratios is zero (since the majority of proteins are not up or down regulated). This ratio parameter is referred to as the volume ratio. In all DeCyder Differential Analysis Software DIA tables the volume ratio is expressed in the range of 1 to ∞ for increases in spot volumes and -1 to $-\infty$ for decreases in spot volumes. Values between -1 and 1 are not represented, hence a two-fold increase and decrease is represented by 2 and -2 , respectively (not 2 and 0.5).

When using either double or triple detection algorithms, the images displayed in the primary and secondary views can be selected using the drop down in the Image View title bar. Spot quantitation is automatically recalculated upon alteration of the primary and secondary image. When using single detection the volume ratio value is left blank, since there is no secondary image.

***Note:** The volume ratio is displayed in the Table View in this manner but all data analyses are based on the log of the normalized ratio measurement (see Appendix C for further details).*

3.4 Viewing spot data

3.4.1 Image View

Image View simultaneously displays both primary and secondary gel images.

Five toolbar functions are associated with the Image View. All these functions can also be accessed through menu pull down options, indicated in parentheses.



(**View:Image View**). Expands the Image View to fit the workspace.



(**View:Zoom:Zoom In**). Zooms in to selected region of the image.



(**View:Zoom:Zoom Out**). Zooms out of selected region of the image.



(**View:Zoom:Fit to window**). Fits the entire gel image to Image View window.



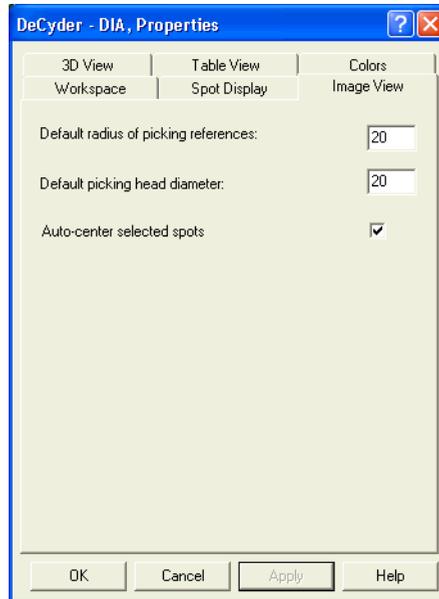
(**View:Contrast Brightness**). Changes the brightness and contrast of the Image View. Raising the positions of the slide bar increases either the contrast or brightness. Selecting the **Apply to all Images** check box results in linking the controls to all images. Altering the contrast and brightness does not change the raw pixel data contained within the image, hence subsequent analyses are not effected

Zooming in and out of the image can also be performed using the mouse by dragging the mouse over a square area of the image to be zoomed. Dragging the mouse from top left to bottom right results in zooming into the image, whilst dragging the mouse from bottom right to top left result in resizing the image to fit.

Properties



Properties associated with Image View are defined in the Image View properties window. Selecting **View:Properties...** or selecting the **Properties** icon then selecting the **Image View** tab displays the Image View properties window.



Changing the default radius of picking references allows the user to define the reference marker for Ettan Spot Picker.

The default picking head diameter is used to define the size of the automated picking head. For example, Ettan Spot Picker manual recommends a 2 mm picking head diameter. Therefore with a 100 micron resolution image this translates to a 20 pixel picking head diameter (the default value).

Selecting the **Auto-center selected spots** check box results the image view automatically shifting so that the selected spot is in the center of the view.

3.4.2 3-D View

The 3-D View function provides a three dimensional representation of the primary and secondary images localized on the selected spot, representing the raw image without filtering. The representation is plotted along the X-Y axes in the plane of the gel. The Z axis scale is normalized between the two images based on the volume ratios, facilitating direct visual comparison between the two 3-D spot images.

Three toolbar functions are associated with the 3-D View. All these functions can also be accessed through menu pull down options.



(View:3-D View). Expands the 3-D View to fit the workspace.



(View:Area in 3-D). Displays the area selected in the Image View in the 3-D View.



(View:Rotate 3-D). Rotates the 3-D View, re-clicking the icon stops rotation.

The 3-D View can be rotated manually by dragging the mouse over the 3-D image.

Adjustment of the 3-D View can be performed using the mouse. Holding down both mouse buttons and dragging the mouse upwards and downwards, results in zooming in and out the 3-D View, respectively.

The position of the spot image within the 3-D View can be adjusted by holding down the right mouse button then dragging the mouse over the 3-D View.

Data associated with each of the highlighted spots in the 3-D View are displayed beneath the corresponding spot.

The orientation of the scaling of the 3-D View can be reset to the initial viewing settings by selecting the **Reset** button above in the 3-D View.

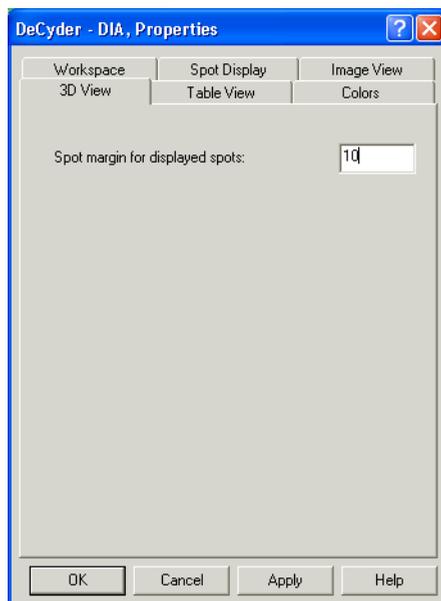
Spot number	Spot reference number (unique to a co-detected spot on a set of images from the same gel).
Position	Gel X,Y co-ordinates of the co-detected spot
Volume	Spot pixel volume
Peak Height	Largest pixel value within the spot boundary (expressed with background subtraction).
Area	Number of pixels within the spot boundary
Pick Position	Gel X,Y co-ordinates of the spot pick location

Properties

Properties associated with 3-D View are defined in the 3-D View properties window

Selecting **View:Properties...** or selecting the **Properties** icon then selecting the **3-D View** tab, displays the 3-D View properties window.

The size of area displayed in the 3-D View can be altered by entering a positive integer between 3 and 80.



3.4.3 Table View

Displays data associated with selected co-detected spots, in a tabulated format. The data within the table can be sorted into ascending or descending order by clicking on the column headers of the table.

Status	Spot No.	Abundance	Excluded	Volume Ratio	Picked	Max Slope	Area	Max Peak...
--------	----------	-----------	----------	--------------	--------	-----------	------	-------------

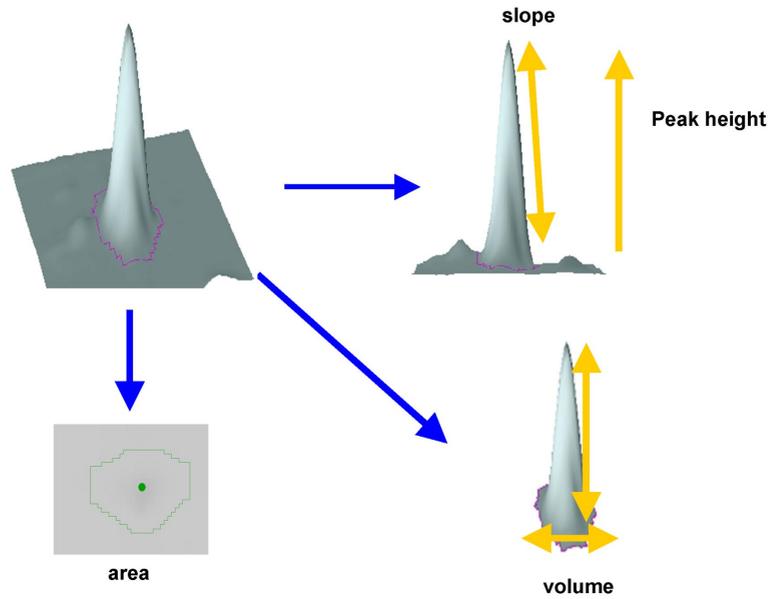


Clicking on the **Table View** icon (or selecting **View:Table View**) expands the Table View to fit the workspace.

The following information on co-detected spots is contained within the Table View. The spots displayed in the table can be adjusted using the **Table View Properties** Tab (see section 3.5.1).

Status	Indicates whether a spot has been confirmed by the user (see section 3.5.1).
Spot No.	Spot reference number (unique to a spot pair on a set of images).
Abundance	“Decreased”, “Similar” or “Increased” depending on thresholds set in DIA Histogram View
Excluded	Spot assigned by user or Exclude Filter for removal from analysis set. An excluded spot is never completely removed from the workspace and can be recovered by the user. Excluded spots are not imported to BVA.
Volume Ratio	Normalized Volume Ratio between co-detected spots in the primary and secondary images.
Picked	“Pick” designation indicates that a spot has been selected for picking (see section 5.4.1)
Max Slope	Largest gradient associated with co-detected spots.
Area	Number of pixels within the spot boundary
Max Peak Height	Largest pixel value associated with co-detected spots.
Max Volume	Volume of the largest of the co-detected spots.
Protein ID	User defined protein identification (manually entered in the Protein ID text box at the bottom of the screen).
Comment	User defined comment (manually entered in the Comment text box at the bottom of the screen).
Function	Indicates whether a spot has been selected as a protein of interest denoted by the letter “I” in the column
PTM	Indicates whether a spot has been selected as a protein that carries a post translational modification denoted by the letters “PTM” in the column

The following diagram illustrates some of the parameters in the Table View associated with a spot.

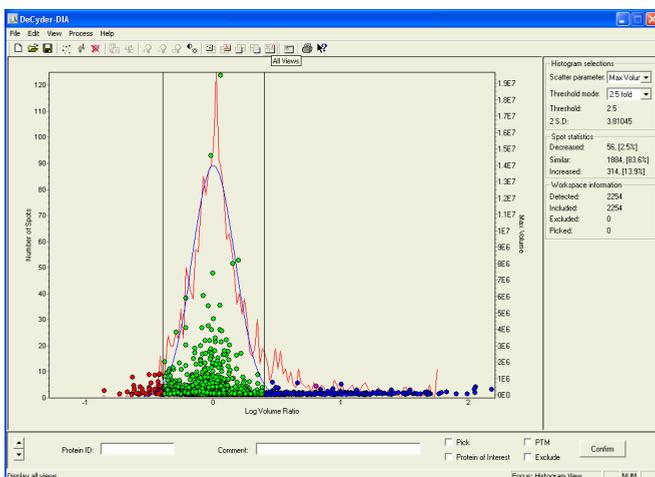


3.4.4 Histogram View

The histogram view is only visible when more than one image is being analyzed (i.e. analyses employing either double or triple detection algorithms).



Clicking on the **Histogram View** icon (or selecting **View:Histogram View**) expands the Histogram View to fit the workspace.



Histogram

The histogram displays data associated with all detected spots in the primary and secondary images. Spot data is plotted against log volume ratio on the X-axis, using two Y-axes.

- **Left Y-axis:** displays the spot frequency. The red curve represents the frequency distribution of the log volume ratios. The curve in blue represents a normalized model frequency fitted to the spot ratios so that the modal peak is zero (see Appendix C).
- **Right Y-axis:** represent the scatter parameter selected in the histogram selection box (right of the histogram). A plotted single data point on the histogram represents an individual protein spot.

The histogram is automatically recalculated when the primary and secondary images are changed using the pull-down menu in the **Image View** title bar.

Histogram Selection

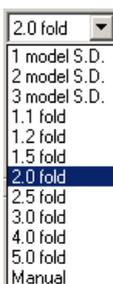
The Histogram Selections toolbar contains two drop down menus and two information boxes: Scatter Parameter and Threshold mode.

Scatter Parameter

The pull down menu can be used to plot either maximum peak height, area, maximum slope, maximum volume (displayed on the right Y-axis of the histogram) against log volume ratios.

Threshold mode

Thresholds are user-defined values that enable highlighting of spots that differ between the primary and secondary image. Consequently, the threshold functionality is predominantly used when performing small-scale experiments to determine expression differences in two samples run on a single gel.



A variety of values can be selected in the threshold pull-down menu.

- **Model S.D.** (standard deviation): number of standard deviations based on the normalized model curve displayed in the histogram view.
- **Fold:** magnitude of volume ratios.
- **Manual:** user-defined magnitude of volume ratios (entered in the **threshold** text box).

The threshold values are represented by vertical black lines in the histogram. Spot boundaries in the Image View and data points in the histogram are automatically color coded to denote the spots that are below, within or above the threshold values. The abundance column in Table View is also updated automatically.

Threshold

Indicates the volume ratio selected in the threshold mode pull down menu (e.g. the volume ratio represented by 2 model S.D.).

2 S.D.

Indicates the volume ratio for 2 S.D. based on the raw data. In a normally distributed data set 95% of data points fall within this value.

Spot statistics

Displays information on the spot population illustrated in the histogram view based on the user-defined threshold settings.

- **Decreased:** the number of spots classified as decreased in their abundance in the primary image, compared to the secondary image.

-
- **Similar:** the number of spots classified as not differentially expressed in the primary image, compared to the secondary image.
 - **Increased:** the number of spots classified as increased in their abundance in the primary image, compared to the secondary image.

The Table View, Image View and Histogram View can be adjusted to selectively display any of these subgroups by selecting the **Spot Display** and **Table View** tabs in the properties dialog box then selecting/de-selecting the appropriate check boxes.

The number of spots designated as excluded and included (designated either manually or by the Exclude Filter) are displayed, see next section for more details.

Workspace information

Displays frequency data on the various categories of spots contained within the entire workspace.

3.5 Data analysis

The DIA module does not require the user to perform any form of spot boundary editing. After spot detection the user only has two post detection tasks to perform.

- 1 Creation and execution of an exclusion filter to remove non-proteinaceous spots.
- 2 Confirmation of detected spots. This is used for small-scale experiments to confirm expression differences in two samples run on a single gel.

This approach was employed to remove user-to-user variation from software analysis. In-house studies have shown that different users generate statistically different results if spot editing is permitted.

The spot detection process (DIA) in DeCyder Differential Analysis Software has been demonstrated to be highly accurate (in-house studies have demonstrated 98% accuracy). The accuracy achieved without editing is taken as an acceptable level and is more than outweighed by the removal of user-to-user variation and higher throughput from not editing.

3.5.1 Spot exclusion

Assigning an area of interest

The automated spot detection algorithm in DeCyder Differential Analysis Software DIA may detect artifacts due to gel heterogeneity at the edges of the images. These spots can be removed by setting an area of interest to define the gel only. Those spots outside of the area of interest will be automatically removed when the Exclude Filter is applied. The area of interest function is only operational when the Exclude Filter is executed. If an area of interest was set before detection all the spots on the image will still be detected (even those outside the area of interest).

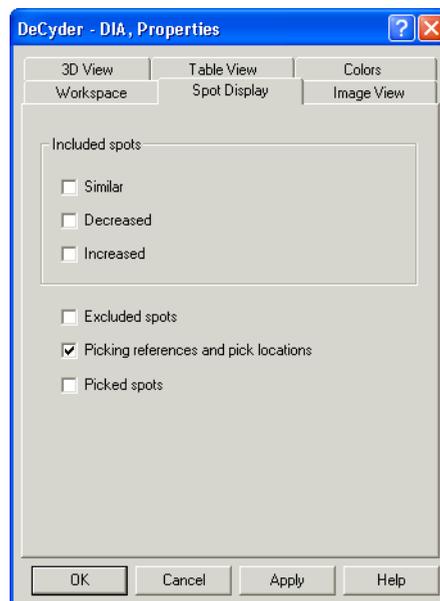


To set an area of interest click on the **Fit to window** icon on the tool bar to fit the gel images to the Image View.



Click on the **Image View** icon on the tool bar to have a full screen view of the gel images.

Click on the **Properties** icon and select the **Spot Display** tab in the window that now appears. De-select **Similar**, **Increased** and **Decreased** and Click **OK**, to remove the spot boundaries displayed in the Image View (de-selecting the spot boundaries on view is optional but helps the user to see the gel more clearly).



Select **Edit:Define Area of Interest** and using the rectangular target pointer drag the mouse to draw a rectangle on either of the gels, taking care to

exclude edge artifacts. The area of interest is automatically defined on the other image(s). Click on the **Properties** icon and select the **Spot Display** tab and reselect the **Increased**, **Decreased** and **Similar** options and click **OK**.

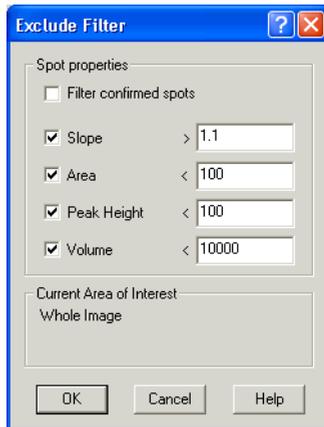
*Note: To remove the area of interest select **Edit:Remove area of interest**.*

Creating an Exclude Filter

The Exclude Filter removes specific spots (based on their physical characteristics) in order to eliminate non-proteinaceous spots from further analyses.

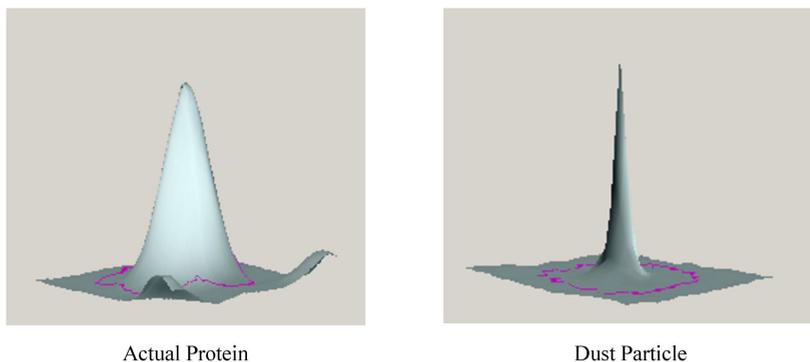


Select **Process:Exclude Filter** (or select the icon from the toolbar) to open the Exclude Filter dialog box.



The 3-D View generally reveals if a detected spot is a gel artifact rather than a protein spot. Dust particles normally have very steep sides and a pointed peak when compared to the smoother curve of a protein spot. Therefore dust particles have high slope values and low area values.

The figure below illustrates a protein spot and a dust particle spot.



Irregularities in the gel surface may also be detected as a spot, however such artifacts have low peak height and volume values.

Artifact spots can therefore be excluded on the basis of slope, area, peak height and volume.

The appropriate filter parameter values can be efficiently obtained by ordering the table according to each parameter in the filter. To order the table according to Max Slope click once on the **Max slope** header on the table so that the largest maximum slope is at the top of the column (if the smallest maximum slope is at the top, click again).

Max Slope	Area	Max Peak...	Max Volume
-----------	------	-------------	------------

Click on the first spot in the Table View and observe the spot in the 3-D View. The 3-D View clearly shows if the detected spot is a dust particle rather than a protein spot. Working down the table leads to spots with smaller and smaller Max Slope values. After a certain value the spots change from gel artifacts to real protein spots. The max slope value of the artifact immediately before the real protein spot in the table is the value entered into the slope parameter of the filter. The same procedure can be applied with the Area, Peak Height and Volume. The only difference with finding these parameters is that the table has to be ordered so that the smallest value is present at the top of the table before scrolling down. If triple detection has been employed, it may be necessary to toggle between the three images, to ascertain more accurately the exclude filter parameters.

After the filter parameters have been established select **Process:Exclude Filter** to display the Exclude Filter window and enter the values determined, click **OK**. The Exclude Filter will automatically remove

detected spots on the basis of these values. In addition those spots outside of the area of interest will also be removed.

***Note:** These excluded spots can be returned to the experiment by de-selecting the check boxes in the Exclude Filter window and then re-running the filter. Alternatively, the values in the exclude filter can be edited, followed by re-running the filter. Excluding spots only removes them from the analysis, not from the workspace.*

Manual spot exclusion

Non-proteinaceous spots that are not removed by the exclusion filter can be manually removed from the data set by highlighting the spot then selecting the **Exclude** check box at the bottom of the screen.

The screenshot shows a software interface with two input fields: 'Protein ID:' and 'Comment:'. To the right, there are four checkboxes: 'Pick', 'Protein of Interest', 'Exclude', and 'Confirm'. The 'Exclude' checkbox is circled in red. There is also a 'Confirm' button to the right of the checkboxes.

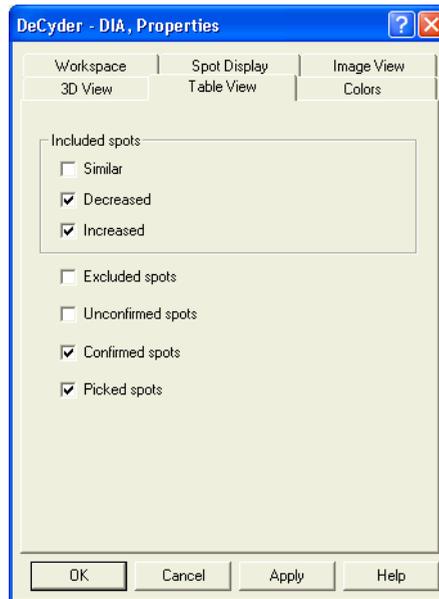
Select **Process:Re-Normalize** to perform the normalization process again with the manually excluded spots removed from the calculation.

***Note:** Re-normalization is automatically performed when the Exclude Filter is used.*

3.5.2 Spot confirmation

Remaining, non-excluded spots can be manually confirmed, for the purposes of visually verifying each spot. Three options are available during spot confirmation:

- 1 Only confirm spots that are increased or decreased in their abundance. This method is relatively rapid since the increased and decreased spots are automatically identified by setting the threshold mode (see section 3.4.4). There is often very little benefit in investigating spots that do not change in expression levels. Select **View:Properties**, and select the **Table View** tab. Ensure in this window that the **Decreased** and **Increased** options as well as **Picked Spots** are ticked, and that the **Excluded** box is not ticked. Click **OK**. It is useful to have the **Confirmed** option checked.



- 2 Only confirm spots that have been assigned with a pick status. This is the recommended option if the protein filter has been applied prior to spot confirmation.
Select only **Picked spots** in the **Table View** tab. The procedure is then identical to option 1.
- 3 All spots (Decreased, Increased and Similar) are manually verified (this takes approximately 1.5 hours for 1000 spots and is not recommended).
Select the **Similar**, **Decreased** and **Increased** options on the **Table View** tab. Click **OK**. The procedure is then identical to option 1.

With the above options, a convenient way to start is to sort the Table View data based on a spot characteristic (e.g. Max Volume). Click once on the column header entitled **Max Volume**. Scroll to the top of the table to see whether the spot with the largest volume has been sorted to the top of the table (if the spot with the lowest volume is at the top of the table click once more on the max volume header). The spots with the largest Max volume (and hence at the top of the sorted table) are almost exclusively protein spots, which can be quickly confirmed. Scrolling down the table to spots with lower Max volumes, non-proteinaceous spots start to appear which can be manually excluded.

Spot confirmation pertains to a Spot Number, therefore when using triple detection a confirmed spot number will possess the confirm status irrespective of which spot maps are displayed in the image view.

3.5.3 Protein of Interest

Protein spots can be assigned as a **Protein of Interest**, thereby highlighting these spots for further investigation. Protein of Interest status can be assigned manually by selecting the **Protein of Interest** check box when highlighting the desired protein spot.



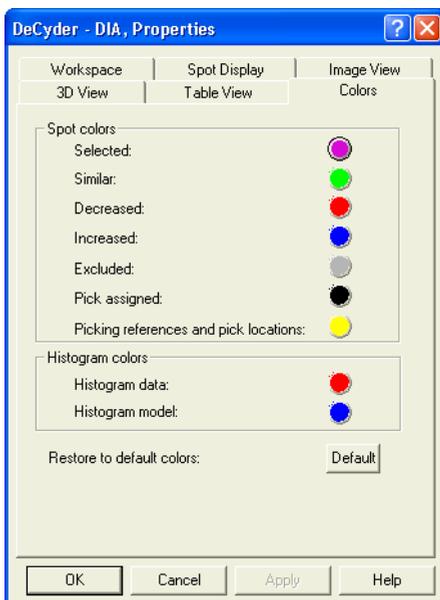
Alternatively, the Protein filter can be used to automatically assign Proteins of Interest using spot property as a selection criteria.

3.5.4 PTM assignment

Protein spots possessing a post-translational modification (PTM) identified by methods outlined in Ettan DIGE system application notes (e.g. application note 18-1170-83 AA, obtainable on the Amersham web site) can be denoted by checking the PTM check box.

3.6 Customizing display colors

The colors used in the various views can be customized by selecting the **Colors** tab in the Properties dialog box.



Spot colors: The colors of spot boundaries displayed in the Image Views and the spot colors in the histogram view can be altered by clicking on the colored circles and selecting a new color.

Histogram colors: The colors of the line graphs in the histogram view can be altered by clicking on the colored circles and selecting a new color.

Clicking **Default** restores the original settings.

3.7 Saving, exporting and printing

Save As

Workspaces are saved as DIA files (.dia file extension), which contain all the data associated with all the loaded spot maps.

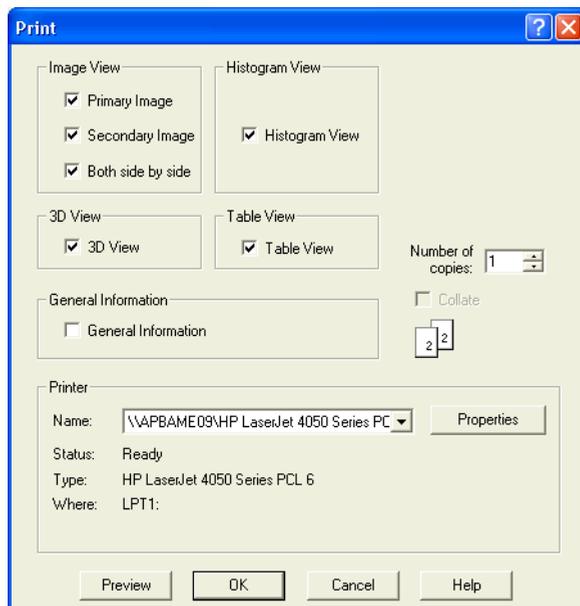
Select **File:Save As**, and browse to locate the folder to save the workspace in. Enter a file name then click **Save**.

Save

Save workspace updates the saved version of the file to include any changes that have been made to the workspace.

Printing

Select **File:Print** to open the **Print** dialog box.



Multiple check boxes can be selected for printing the required workspace views.

DeCyder Differential Analysis Software software supports the following printers:

- HP LaserJet 4000N
- HP color LaserJet 5M
- HP 2500C
- HP DeskJet 895 Cxi
- HP DeskJet 950C
- HP Laser Jet 4100 DTN

Note: Other manufacturers and models of printers have been demonstrated to be compatible with DeCyder Differential Analysis Software.

Using the clipboard

Any aspect of the different views within the workspace can be copied to the clipboard for pasting into another application.

Click on the part of the workspace chosen for pasting then select **Edit:Copy** (the part of the workspace selected is described in this pull down menu). The clipboard can then be pasted into other applications.

Exporting data

Each of the export functions in DIA is located under the **File** menu.

1 Export Result Table

Exports all the data in the Table View in a text format that can be opened in a spreadsheet.

2 Export pick list

Exports the proteins which have been assigned for picking by the user in the form of a .txt or .xml file that can be recognized by Ettan Spot Picker or Ettan Spot Handling Workstation.

3 Export Spot Maps

Exports all the data from the DIA workspace in an XML format that contains all the data generated in the DIA workspace. This can be opened by either the DeCyder Differential Analysis Software BVA or the DeCyder Differential Analysis Software XML toolbox.

4 BVA (Biological Variation Analysis) Module

4.1 Overview

4.1.1 Functionality

DeCyder Differential Analysis Software BVA processes multiple Ettan DIGE system gel images, performing gel to gel matching of spots, allowing quantitative comparisons of protein expression across multiple gels.

BVA processes gel images that have undergone spot detection in DIA. The BVA module utilizes the XML files generated in DIA (Section 3.7) together with the original scanned image files. The images are then matched to a single master image, identifying common protein spots across the gels. Various experimental designs can be assigned in BVA, facilitating the statistical analysis tools to highlight proteins that demonstrate significant protein changes under different experimental conditions.

Additional in-built functionality allows post-matching activities to be performed:

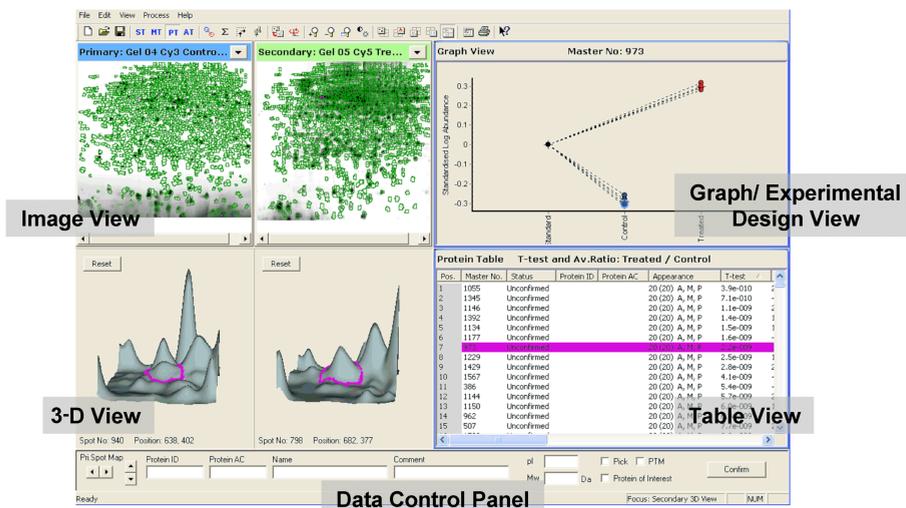
- Molecular weight calculation
- Isoelectric point calculation
- Database linkage
- Statistical analyses
- Spot pick list generation

Data can be saved then re-opened in a BVA file format. Pick lists can be exported as either a text file or XML file for use in Ettan Spot Picker or Ettan Spot Handling Workstation, respectively. Data can also be exported in an XML format for querying in DeCyder Differential Analysis Software XML toolbox or copying and pasting into applications such as Microsoft™ Word and Excel™.

4.1.2 Graphical user interface

The BVA graphical user interface is similar to DIA. It is divided into four equally sized inter-linked views. Selecting a spot in, for example, the Image View will display the same spot in the Graph View, 3-D View and the Table View.

- Image View: displays two gel images.
- 3-D View: a three dimensional representation of the images localized on the spot selected.
- Graph View, graphical representation of data.
- Table View, tabulated data.



There is also a data control panel present on the bottom of the workspace, which incorporates specific functionality features. The contents of the control panel and the Table View are dependent on the mode selected.

4.1.2 Structure

The BVA is composed of four different modes, which display data manipulated through various tables and associated controls. Each mode provides different functionality associated with specific processes in BVA analysis.

ST Spot Map Table

The Spot Map table is used to set up images for spot matching and statistical analysis. The table lists data related to the Spot Maps imported from DeCyder Differential Analysis Software DIA module.

HT Match Table

The Match table is used for the processes associated with inter-gel matching. The table lists all data associated with the matching algorithm.

Protein Table

PT

The Protein table is used to display and process data associated with the protein spots identified across the gels. Each row of the table corresponds to one protein spot, which may be present in several Spot Maps.

Appearance Table

AT

The Appearance table is used to display data associated with a selected protein across the gels.

4.2 Protein quantitation

When using DeCyder Differential Analysis Software BVA for inter-gel analysis it is highly recommended that an internal standard is used in the experimental design (see Section 1.4). As discussed in the previous section, the DIA quantifies the spots as a function of the internal standard. This value is used in the BVA for analyses facilitating direct comparison of spot maps, and is referred to as the standardized abundance. The \log_{10} of this value (referred to as standardized log abundance) is used to aid scaling in graphical representations and is employed in all statistical analyses.

If an internal standard is not used in the experimental approach DeCyder Differential Analysis Software can express spot volumes of specific proteins across gels as a ratio relative to the lowest volume spot of that protein. This parameter is referred to as the abundance in BVA. This latter approach is not recommended, since inter-gel quantitative comparisons are not as accurate as the quantitation using an internal standard. It is not possible to perform the BVA statistical functionality with this non-internal standardized data.

As with the standardized abundance, \log_{10} of this value is often used to aid scaling in graphical representations (referred to as log abundance).

4.3 Creating and opening workspaces



From the Desktop or **Start** button menu select the DeCyder Differential Analysis Software – **BVA** icon to open the module.

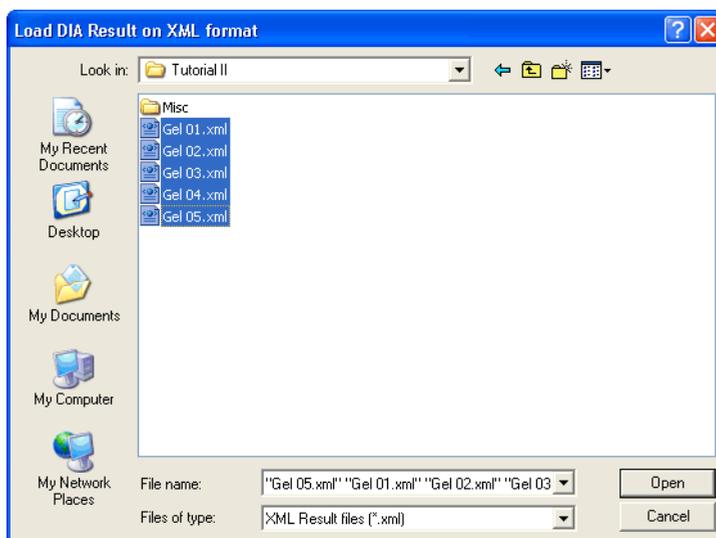
4.3.1 Creating workspaces

DeCyder Differential Analysis Software BVA requires XML files generated in the DIA (section 3.7) together with the original scanned image files, to create a new BVA workspace.



To load XML files in BVA select **File:Create Workspace** in the DeCyder Differential Analysis Software BVA window or select the **Create workspace** icon.

Locate the XML files and select the files for opening by holding down the Control key and clicking on the required files. Click **Open**.



*Note: The XML files contain the necessary information to open the associated image files automatically, providing the image and XML files are in the same relative file path when processed in the DIA module. If this is not the case, a message box will appear giving the option to locate the image file(s). Click **Yes** and browse to locate and select the corresponding image files. The newly created workspace should now be saved (see section 4.13.5).*

4.3.2 Opening workspaces



To open previously created and saved workspaces select **File:Open** and browse to locate the required BVA file. When the file is located select the file and click **Open**.

4.4 Viewing spot data

The general toolbar functions associated with the Image, 3-D and Table View are identical to those in the DIA and are described in sections 4.4.1, 4.4.2 and 4.4.3, respectively.

4.4.1 Image View

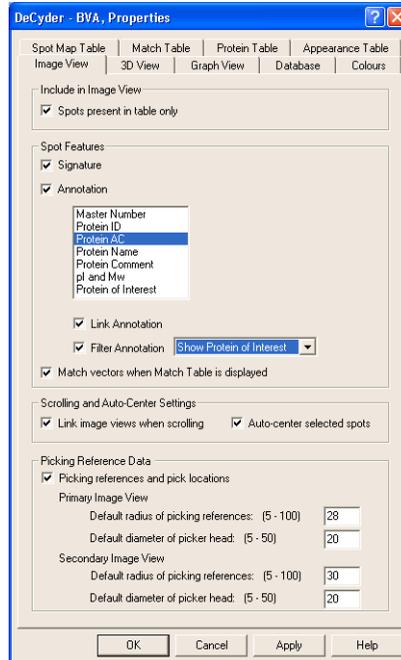
Simultaneously displays two gel images. The gel image displayed can be selected using the pull down menu in the title bar of each gel image window.

Properties

Properties associated with Image View are defined in the Image View Properties window



Selecting **View:Properties...** or selecting the **Properties** icon and then selecting the **Image View** tab displays the Image View Properties dialog box.



Include in Image View: Select check box to display only the spots displayed in the Table View.

Signature: Select to display the spot boundaries.

Annotation: Select to display user determined annotation as text boxes on the gel Image Views. The annotation displayed is selected from the list, only one category can be selected.

Selecting the **Link Annotation** check box results in spots derived from different isoforms of the same protein (defined by the user by having the same Protein ID, AC, name or comment) being grouped and labelled with a single label. The **Filter Annotation** box can be selected to only display annotations for pick status spots or spots selected as proteins of interest.

Match vectors when Match Table is displayed: Select to display the match vectors (see section 4.6.2).

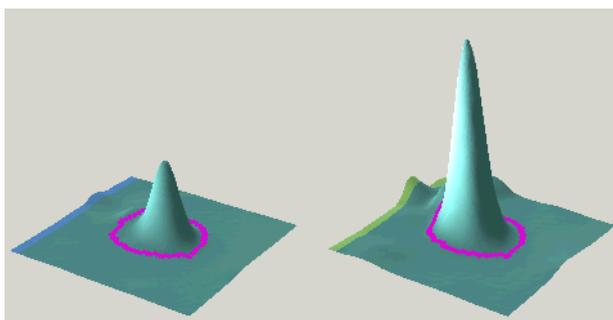
Link Image Views when scrolling: Select to link scrolling in both images.

Auto-center selected spots: Select to position selected spots in the center of the Image View.

Picking References: The radius of the picking reference markers is entered in the dialog boxes.

4.4.2 3-D View

A three dimensional representation of the two images selected in the Image View, localized on the selected spot.



This view is not visible in the Spot Map mode. With the exception of the data displayed below each of the 3-D images, the BVA and DIA 3-D Views are identical (see section 3.4.2).

A colored banner across one of the edges of the 3-D View images corresponds to the color of the title bar in the corresponding image view. The banner reveals the orientation of the 3-D View image, denoting the edge of the view closest to the top of the gel.

Pick locations, when applicable, are displayed in the 3-D View (see section 5.6).

Properties

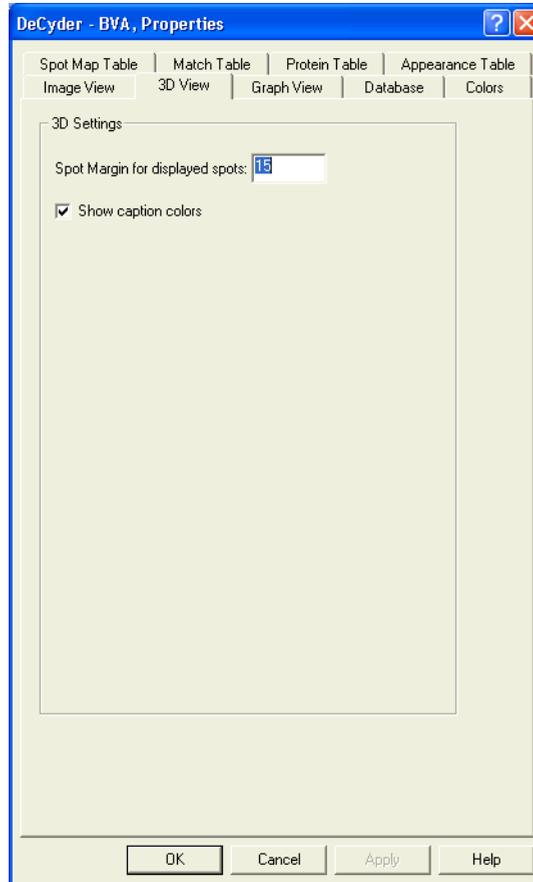
Properties associated with the 3-D View are defined in the 3-D View Properties window.

Selecting **View:Properties...** or selecting the **Properties** icon and then selecting the **3-D View** tab displays the Image View Properties dialog box.

The size of area displayed in the 3-D View can be altered by entering a

positive integer between 3 and 80.

The **Show caption colors** check box allows the colored banner indicating the upper edge of the 3-D View (as seen in the Image View) to be removed.



4.4.3 Table View

The contents of the Table View is dependent on BVA mode selected. Similarly the editable properties associated with the Table View are dependent upon the mode selected and are discussed in detail in later sections. The data within the table can be sorted into ascending or descending order by clicking on the column headers of the table.

No.	Status	Image	Type	Label	No.of Spots	Matched	Function	Group	Group Description	Condition1	Condition2	Sample ID
-----	--------	-------	------	-------	-------------	---------	----------	-------	-------------------	------------	------------	-----------

4.5 Defining spot map attributes

All spot map attributes are assigned in the Spot Map Table mode. Select **View:Spot Map Table** or select the toolbar icon to work in the Spot Map Table mode (if a new BVA workspace has been created the interface will automatically open to the Spot Map Table). This mode consists of an Image, Table View and Experimental Design View.

4.5.1 Function assignment

Spot Maps can be assigned with up to four functions that indicate the role of the Spot Map in the BVA processing. The Spot Map functions are:

Analysis (A) – Spot data is included in quantitative analysis. Those that should not be assigned as analysis images may include preparative gels for spot picking and images to be used as a template.

Master (M) –One spot map can be assigned as the Master Spot Map in each workspace. All other Spot Maps are matched to the Master. When creating a new workspace, the Spot Map with the largest number of spots is automatically set to be the Master Spot Map.

Template (T) - One spot map can be assigned as the Template spot map in each workspace. Protein ID, Protein AC, Name, Comment, pI, Mw and Protein of Interest assigned to the spots of the Template Spot Map will be cross annotated to matching spots in all the other images.

Pick (P) – This function should be assigned to a spot map if, in DIA, pick assignments have been made and the user wishes to transfer the pick status to corresponding spots in the BVA workspace. (see section 5.5).

To assign the spot map functions select **View:Table View** or click on the **Magnify Table View** icon to display the Table View only. The table shows the images that have been loaded into the workspace and the number of spots that have been detected on them.

Spot Map Table												
No.	Status	Image	Gel No.	Type	Label	No. of Spots	Matched	Function	Group	Group Description	Sample ID	Comment
1	Matched	Gel 01 Cy2 Standard.g1		DIGE Min	Cy2	2214	1865	A		Standard		
3	Master	Gel 01 Cy2 Standard.g2		DIGE Min	Cy2	2564	2564	M, A, P		Standard		
5	Matched	Gel 02 Cy2 Standard.g3		DIGE Min	Cy2	2204	1635	A		Standard		
7	Matched	Gel 02 Cy2 Standard.g4		DIGE Min	Cy2	2529	1843	A		Standard		
9	Matched	Gel 03 Cy2 Standard.g5		DIGE Min	Cy2	2246	1679	A		Standard		
11	Matched	Gel 03 Cy2 Standard.g6		DIGE Min	Cy2	2455	1855	A		Standard		
13	Matched	Gel 04 Cy2 Standard.g7		DIGE Min	Cy2	2422	1654	A		Standard		
15	Matched	Gel 04 Cy2 Standard.g8		DIGE Min	Cy2	2323	1690	A		Standard		
17	Matched	Gel 05 Cy2 Standard.g9		DIGE Min	Cy2	2339	1674	A		Standard		
19	Matched	Gel 05 Cy2 Standard.g10		DIGE Min	Cy2	2138	1696	A		Standard		
4	Matched	Gel 01 Cy5 Treated.g12		DIGE Min	Cy5	2564	2564	A	Treated	BA treated	1	
8	Matched	Gel 02 Cy5 Treated.g11		DIGE Min	Cy5	2529	1843	A	Treated	BA treated	2	
12	Matched	Gel 03 Cy5 Treated.g16		DIGE Min	Cy5	2455	1855	A	Treated	BA treated	3	
16	Matched	Gel 04 Cy5 Treated.g18		DIGE Min	Cy5	2323	1690	A	Treated	BA treated	4	
20	Matched	Gel 05 Cy5 Treated.g10		DIGE Min	Cy5	2138	1696	A	Treated	BA treated	5	
2	Matched	Gel 01 Cy3 Control.gel 1		DIGE Min	Cy3	2214	1865	A	Control	No treatment	1	
6	Matched	Gel 02 Cy3 Control.gel 3		DIGE Min	Cy3	2204	1635	A	Control	No treatment	2	
10	Matched	Gel 03 Cy3 Control.gel 5		DIGE Min	Cy3	2246	1679	A	Control	No treatment	3	
14	Matched	Gel 04 Cy3 Control.gel 7		DIGE Min	Cy3	2422	1654	A	Control	No treatment	4	
18	Matched	Gel 05 Cy3 Control.gel 9		DIGE Min	Cy3	2339	1674	A	Control	No treatment	5	

Note: A set of images from the same gel will have the same number of spots since the DIA detection algorithm is designed to detect the same number of spots on images from the same gel.

In the **Function** column all the images are, by default, assigned **Analysis (A)** and one of the analysis images is assigned as **Master (M)**. The analysis function designation indicates that each of the images will be included in the statistical analysis.

The Master function can be assigned to a different image, by selecting the image to be assigned as master using the left mouse click, then selecting the Master check box at the bottom of the workspace in the area entitled Spot Map Functions.

The Master function can be assigned to a different image, by selecting the image to be assigned as master using the left mouse click, then selecting the **Master** check box at the bottom of the workspace in the area entitled **Spot Map Functions**.



All other function assignments are similarly performed.

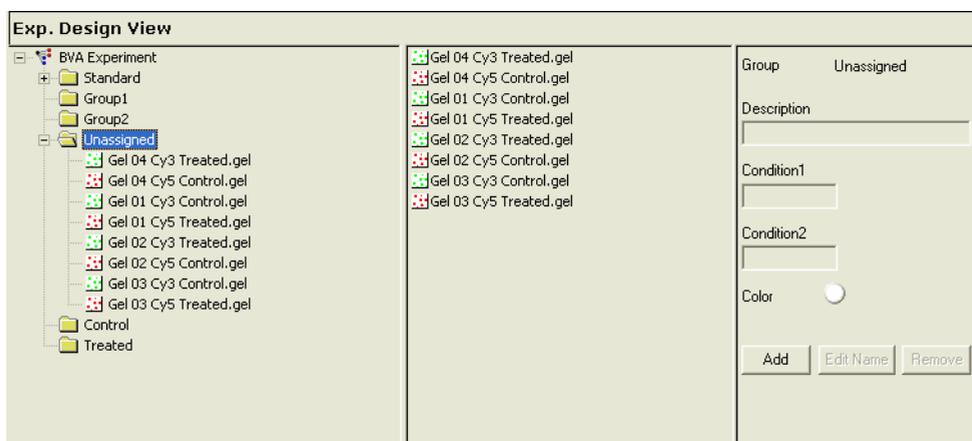
4.5.2 Group assignment

A group is a collection of spot maps that for the purposes of the experiment cannot be broken down into further sub-groups. Groups can include analysis sets such as control groups, treated groups, time points and temperature points. This spot map assignment is necessary to facilitate the inter-group statistical analyses in DeCyder Differential Analysis Software.

The internal standard (non-experimental group) spot maps are assigned

as standard if the word **standard**, **pooled** or **std** is present in the image file name. By default non-standard spot maps are assigned as “Unassigned”.

The experimental group assignment is displayed in Experimental Design View when within the Spot Map Table mode.



Spot Maps are located in the appropriate folder on the left side of the Experimental Design view. The Spot Maps can be viewed by selecting the desired group folder.

Spot Maps can be assigned to the predefined “Group1” and “Group2” by selecting the Unassigned folder, then dragging and dropping Spot Maps from the center panel of the Experimental Design view to the desired folders.

The names of predefined “Group1” and “Group2” experimental groups can be renamed by selecting the **Edit Name** button on the right of the Experimental Design view.

New custom named groups can be generated. To create new groups select **Add**, then type in the desired name of the group in the Group box. A new group folder is subsequently displayed.

The assigned groups are also displayed in the Spot Map Table in the Group column.

The description text box adjacent to the Group assignment can be used to enter user defined information regarding the description of the experimental group. This description is linked to the experimental group.

4.5.3 Comment assignment

The comment text box on the right of the data control panel can be used to enter a user defined comment that is specific to the Spot Map in the primary view.



4.5.4 Spot Map Table Mode

The Spot Map Table Mode displays data associated with the spot maps. This data is derived from the XML file generated in the DIA module.

No.	Spot map number in table
Status	Indicates whether the image is Matched, Pending or Processing.
Image	Gel image file name.
Type	Type of labelling chemistry used.
Label	Dye used to label the protein (e.g., CyDye DIGE Fluor Cy2, Cy3 and Cy5 minimal dyes).
No. Of Spots	Number of protein spots identified in the DIA.
Matched	Number of protein spots that have been matched to the protein spots on the master.
Function	Spot Map function (e.g. Master, Template, Analysis and Pick).
Group	Spot Map group (e.g. Control, Treated and Standard).
Group Description	User defined description of group.
Condition 1	User defined condition assigned numerically (example Time point 1,2,3...).
Condition 2	User defined condition assigned numerically (example Dose point 1,2,3...).
Sample ID	User specified identification of sample
Comment	User defined comment on Spot Map.

4.5.5 Spot Map Table properties

Properties associated with Spot Map Table View are defined in the Spot Map Table Properties window

Selecting **View:Properties...** or selecting the **Properties** icon then selecting the **Spot Map Table** tab displays the Spot Map Table Properties dialog box.

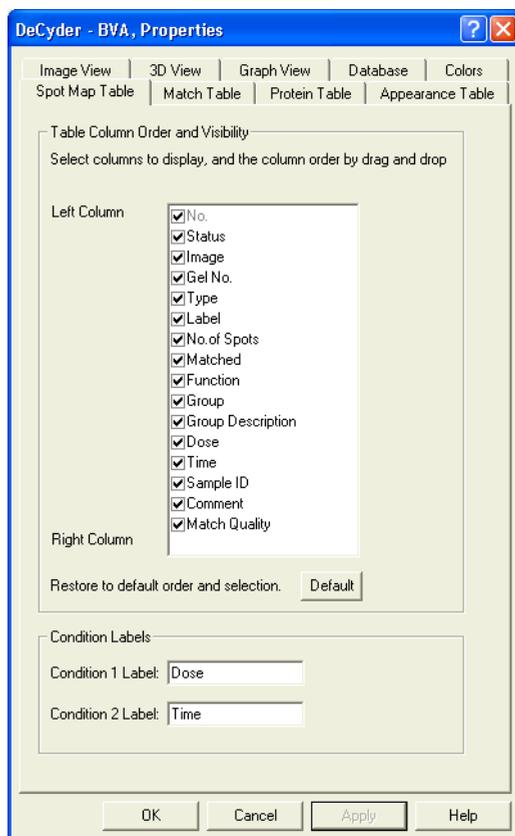


Table Column Order and Visibility: The column titles selected will be displayed in the Table View. Dragging the column titles in the list defines the order of the columns in the Table View. Clicking **Default** restores the original settings.

Condition Labels: Labels for condition 1 and condition 2 (e.g. time, dose) can be entered (see section 4.8.8).

4.6 Inter-gel matching

All spot maps are matched to a master image sequentially. The spot-matching algorithm is a pattern recognition algorithm that matches one single spot in one gel to a single spot in another gel based on its neighboring spots.

4.6.1 Landmarking

Landmarking allows the user to manually define matched protein spots in order to improve the accuracy of the gel-to-gel matching process. In some cases images from different gels may vary sufficiently to require landmarks to be set.

Landmarking is performed in the Match Table mode. Select **View:Match Table** or select the **toolbar** icon to work in the Match Table mode. This mode consists of an Image, 3-D and Table View.

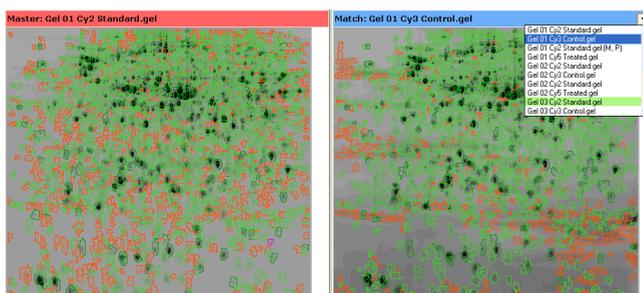
To use a spot as a landmark the spot must be present on both the match image and the master.



Click on the **Magnify Image View** icon to display the gel images only. The image on the left represents the master image (in the Match Table mode the left gel image is always the master image) while the image on the right represents one of the images to be matched to the master.



To commence setting landmarks click on the button entitled **Landmark Mode**, in the data control panel at the bottom of the workspace (this button will now become yellow or to a color the user has assigned as the landmark color).



In the drop down menu in the Image View title bar select the image for landmarking.

If all the spots are unmatched the spot boundaries are red to indicate this. If the spot boundaries are green in an unmatched set of spots then the images are from the same DIA file (thus landmarks are not necessary).

If the spots on the images are not clearly visible it may be necessary to alter the contrast/brightness settings of the images.



Click on the **Contrast/Brightness** icon and alter the position of the bars to alter the contrast and brightness of the images until only the most intense spots are visible. To set a landmark click on a clearly defined spot on the master image (the spot boundary should become magenta - color set as default). Now select the spot on the match image which corresponds to the master image spot, so that this spot also becomes magenta. A few seconds later a vector line, showing the displacement of the spot, should appear showing that the spots have been matched and the landmark has been set.

Note: Landmarking can be aided by viewing the pattern around the selected spot in the 3-D View to confirm that the two selected spots do correspond to each other. The number of spots being displayed around the selected spot in 3-D can be altered by opening the Properties window and selecting the 3-D View tab. Altering the numerical value in the Spot Margin For Displayed Spot option adjusts the parameter size which may consequently change the number of spots displayed around the selected spot.

It is recommended that landmarks be evenly distributed across the image as this aids the matching process. Multiple gels can be landmarked in an identical manner. Images from the same DIA file as a previously matched gel are automatically landmarked. It is therefore only necessary to set landmarks in one image from a set of images derived from a single DIA analysis. Ideally the landmarked image should be the standard spot map. The landmarks which were set on previous spot maps will be colored yellow on the Master image, thus the same landmarks can be set on the next match image by finding the spots which correspond to the yellow master image spots. It is not necessary to set the same landmarks on each gel, but as landmarks chosen by a user tend to be the clearest spots, it is useful to have these highlighted so that the landmarking process is speeded up.

When landmarking is completed deselect the **Landmark** button.

Note: It is usually only necessary to set landmarks on those images that differ significantly from the Master image. In addition landmarks may have to be set after matching if there are a high number of wrongly matched spots, in which case matching will have to be repeated. Often landmarks do not need to be set at all.

4.6.2 Matching

The matching process is divided into two steps (Level 1 and Level 2).

- 1 The algorithm first matches large spots spread evenly over the whole gel. The algorithm matches by comparing positions and sizes of the neighboring spots. The matched spots are then used as landmarks, equivalent to manual matches. The algorithm then starts from these landmarks and matches neighboring spots. In this way the matching propagates out from the landmarks by matching a spot and then matching its neighbors.
- 2 Matching is performed by transforming the detected spot center from one gel to the other gel. A spot center transformation is considered a match if the transformed co-ordinates are close enough to a spot center on the other gel. After each matching step is a control that removes obvious mismatches.



To perform matching select **Process:Match...** or click the **Match** icon.

Select one of the different options to match all the images or just a specific set of images.



Match all: matches all spot maps in the workspace (including those already matched).

Match Primary: matches spot map selected as primary in the Image View.

Match pending and landmarked: matches spot maps that have not been matched and those that have been landmarked after a previous matching process.

Click **Match** to commence matching.

4.6.3 Match confirmation

The matches in the Table View are of two types, Auto Level 1 and Auto Level 2. By default the two types of matches are displayed differently in the Image View. Auto level 1 and Auto level 2 spots are represented by

a darker and lighter spot boundary highlights, respectively. The boundaries of the unmatched spots are colored red. The vector lines in the Image View indicate the positional difference for the same protein spot on different gels.

The matching can be checked within the Match Table mode to ascertain the accuracy of the matching process. This can be done in several ways:

- 1 Confirm a few matches from different areas on the gel by clicking around the image then moving onto the next image (recommended option).
- 2 Order the spots in the Match Table by match quality values then confirm the spots with the highest match quality values (see section 4.6.4).
- 3 Confirm randomly selected auto level 1 (e.g. 5) and auto level 2 (i.e. 10) matches by scrolling down the table.
- 4 Confirm all matches on every gel (time consuming and not recommended).

Note: Areas of incorrect matching can be identified rapidly by examining the match vectors displayed in the Image View. The vectors should be orientated in a similar direction across the gel. If an area of the gel does not follow this pattern (e.g. the vectors are perpendicular or cross), it is highly likely that mismatches are present in that area and should be looked at closely.

Deciding whether a match is accurate can be aided by viewing the selected spot and the surrounding cluster in the 3-D View as well as looking at the matched spots in the Image View. If the match is correct, confirm the match, by clicking the **Confirm match** button in the data control panel.

If the match is incorrect, click on the **Break match** button, which now becomes **Add Match**. Click the correct spot on the match image.

If the spot on the match image has also been wrongly matched break this match as before. Select the corresponding spots on the master image and the match image and click on the button entitled **Add Match**.





If a matched spot on the primary image is not present on the master for any reason it can be added to the master by clicking the **Add to Master** button.

A comment can be appended to each match in the comment text box.

4.6.4 Match Quality Metric

Match quality values are typically between 1 and 15, expressed as real numbers approximated to two decimal points. This value represents morphological, but not positional similarities in matched spots. The match quality is calculated for only the internal standard spots that are matched across a set of gel images, as we know that these proteins should be the same.

The match quality value is calculated from the surface profile of each internal standard on each gel matched to the master spot maps. The surface profiles of all standard spots in a set of matched spots are combined to form an “average” surface profile. Subsequently, all the individual internal standard spot surfaces in the match are compared to this “normal” surface (including the spot derived from the master spot map). The quality score is the calculated difference between the individual matched spots and the “normal” spot.

The spot surfaces that deviate a great deal from the “normal” surface receive the highest score, whereas the spot surfaces that have close resemblance to the “normal” surface will be given the lowest score. Therefore, values approaching zero indicate good spot similarity and hence accurate matching. As the quality score increases the degree of similarity decreases revealing incorrect matches. The match quality value in the Match Table can therefore be used to rapidly identify incorrect matches when reviewing inter-gel matching.

In addition to quality score for each matched spot being displayed in the match table, each set of matched spots (a protein) are also assigned a global match quality score (displayed in the Protein table). This protein match quality score represents the spot within the matched group that has the highest quality score. This protein match quality value enables the identification of poorly matched spots, thereby highlighting matched spots in the protein set that may contain outliers. All the match quality values for all matched spots in a protein can be viewed simultaneously in the Appearance Table facilitating the rapid identification of the outlying data point.

There is also a quality score value designated for each spot map, which is displayed in the Spot Map Table. This value refers to the average quality score for all matched standard spots on each gel. Hence, gels that are poorly matched possess high match quality values. High spot map match quality values are often due to poorly run gels that cannot be matched efficiently without land marking (see section 4.6.1).

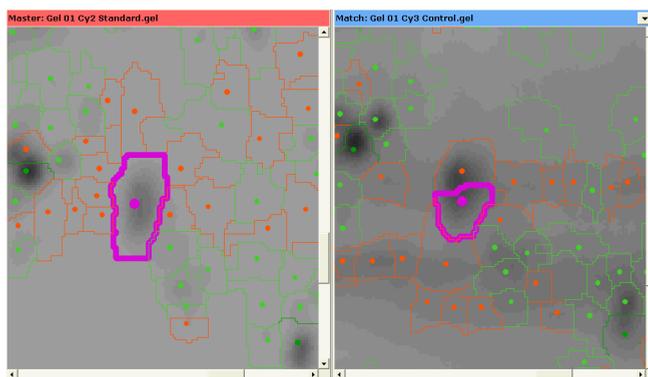
Note: The match quality score is based exclusively on the surface profile of the spots on the internal standard spot maps. Therefore, the quality score cannot be calculated for experimental designs that do not contain an internal standard.

Furthermore, each co-detected spot for multiple images on the same gel all possess the same match quality value, which is associated with the profile of the internal standard spot only.

4.6.5 Spot Merging

Very occasionally, a protein peak will be dissected into several spots despite appearing as a single discrete spot. This is normally due to subtle gradient changes on the spot surface that are not always apparent to the naked eye. These “split” spots can be merged to form a single spot boundary within the Match Table mode.

Select the matched spot in the image view that requires to be merged to surrounding spot(s).

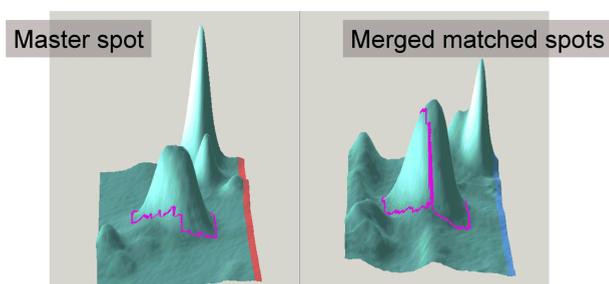


Selecting **Edit:Merge Spots...** results in the surrounding unmatched spots being exclusively displayed. Click on the spot in the image view that requires merging to the spot selected originally.

***Note:** Two spots can be merged if at most, one of them is matched to another spot. If one master spot is matched to several spots, it's possible to merge it with another spot, but only if that other spot is not matched.*



Both the image view and the 3-D View display the new spot boundary showing all the component spots. The volume of the merged spot and subsequent volume ratios (when using an internal standard) are recalculated automatically.



The merging process can be reversed by selecting **Edit:Split Previously Merged**. Splitting merged spots results in all the associated spots being unmatched.

4.6.6 Match Table

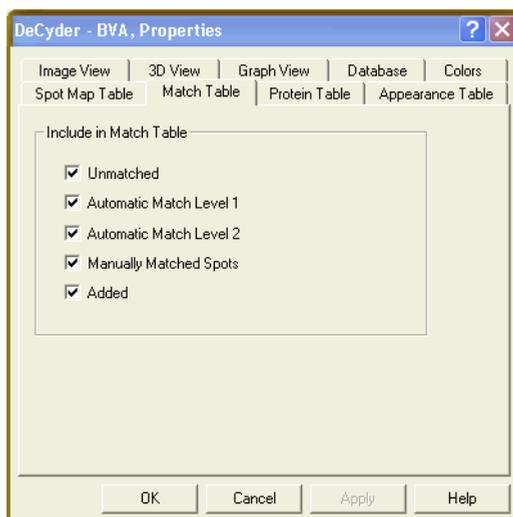
The Match Table displays spot specific information.

Pos.	Match Table row number
Master No.	Protein spot number on Master Spot Map
Status	Match confirmation status (confirmed or unconfirmed)
Master	Protein spot co-ordinates on Master Spot Map
Match	Protein spot co-ordinates on Match Spot Map
Type	Types of match: Unmatched: Unmatched spots Auto Level1: Spots matched after first level of matching Auto Level2: Spots matched after second level of matching Manual: Spots matched by the user Added: Spots that have been added to the master
Comment	User defined comment regarding matched protein spots.

4.6.7 Match Table properties

Properties associated with Match Table are defined in the **Match Table Properties** dialog box.

Selecting **View:Properties...** or selecting the **Properties** icon then selecting the **Match Table** tab displays the Match Table Properties dialog box.



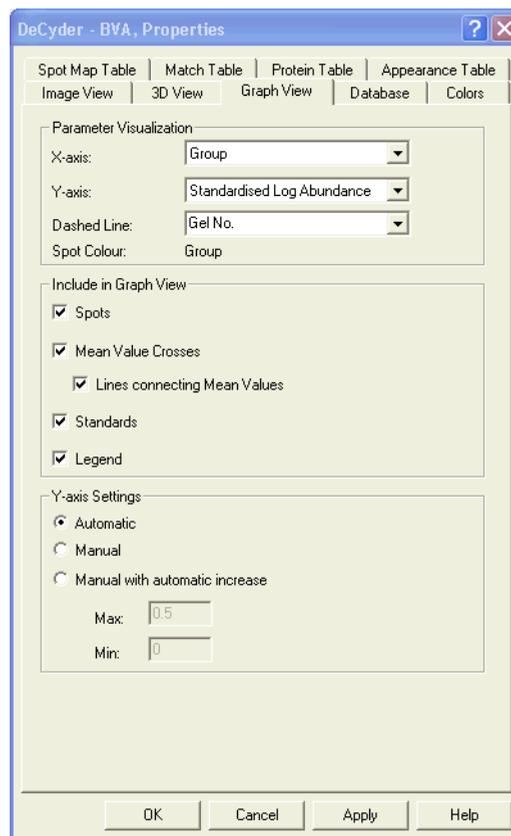
Include in Match Table: Selects the categories of spots that are to be displayed in the **Match Table**.

4.7 Graphical representation

The graphical view in DeCyder Differential Analysis Software BVA allows the user to view scatter plot representations of data points associated with individual proteins. For example the manner in which the expression of a protein changes with time or drug dosage can be graphically viewed. The graphical views can only be seen in the Protein Table mode and Appearance Table mode. Examples of the graphical representations are illustrated in the next section.

Properties associated with graph view are defined in the Graph View dialog box.

Selecting **View:Properties...** or selecting the **Properties** icon then selecting the **Graph View** tab displays the **Graph View Properties** dialog box.



Parameter visualization

Allows the user to define how data is displayed on the graph by designating the parameters displayed on the X and Y-axis of the graph.

The **X-axis** options allow the user to display data according to group, conditions or fluor type. The sequence of the groups on the X-axis are determined by the order of group folders in the Experimental Design view within the Spot Map Table mode. This can be changed by dragging and dropping the folder into the desired order within the Experimental Design view.

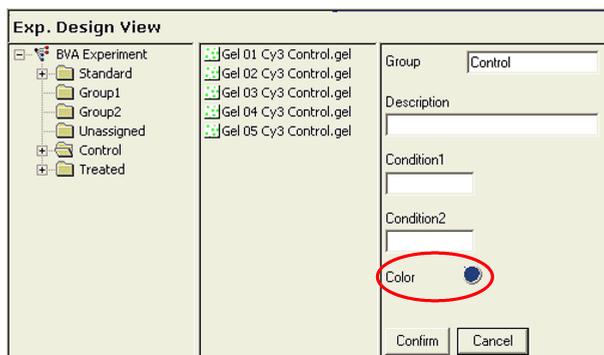
The **Y-axis** options allow the user to display either the abundance, log of the abundance, standardized abundance or log standardized abundance. However the statistical functions within DeCyder Differential Analysis Software only utilize log standardized abundance. Therefore the graphical representation of this parametric value only reflect the data points used in the statistical analyses.

Dashed line options allow the user to display dashed lines on the scatter plot linking data points by the specified association. Samples either derived from the same gel or possessing the same sample ID can therefore be easily identified in the graph view.

Include in graph view

The options available determine the data points plotted on the graph.

By selecting the **Spot** check box all proteins spot data points associated with the user defined X-axis parameter are shown. Colored circles represent these data points. The colors representing the groups can be assigned in the Experimental Design view.



Selecting the **Mean Value Crosses** check box results in the mean value of the data points associated with the user defined X-axis parameter being displayed (represented by a cross). Lines connecting these mean values can be displayed by selecting the subsequent check box.

If either Groups or conditions are assigned to the X-axis, the Y-axis values corresponding to the internal standard can also be displayed by selection of the **Standard** check box.

Note: *The DeCyder Differential Analysis Software standardization process defines all protein abundance values associated with the internal standard to be 1 (or zero if the log values are displayed). For further details see Appendix C.*

The **legend** check box allows the user to determine whether legends defining the color coding of groups will be displayed.

Y axis settings

The Y-axis setting options enable the user to change the Y-axis scaling:

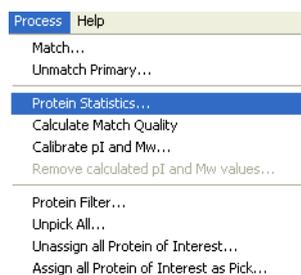
- **Automatic** allows optimization of the Y-axis scaling.
- **Manual** allows the user to define the maximum and minimum values on the Y-axis.
- **Manual with Automatic increase** allows the user to define the maximum and minimum values on the Y-axis but adjusts the Y-axis values if the data points fall outside the manual defined values.

4.8 Protein statistics

4.8.1 Opening the protein statistics dialog box

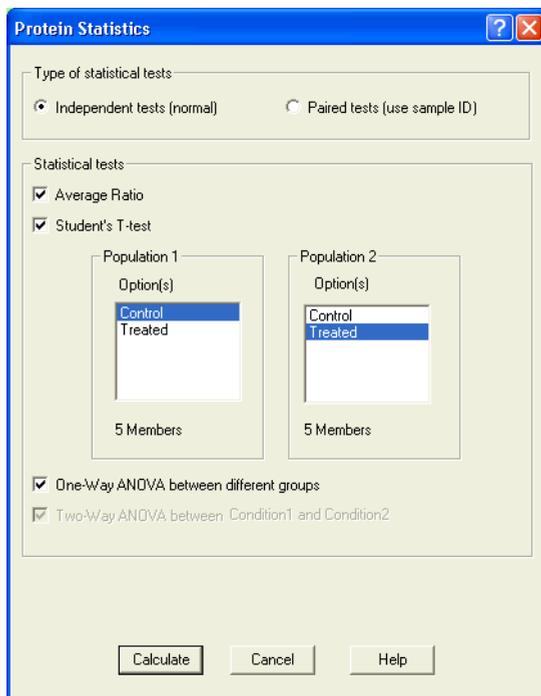
DeCyder Differential Analysis Software BVA possesses several statistical analysis methods that can be employed to ascertain whether changes in expression of specific proteins are significant between samples from different experimental groups.

To open the statistical analysis dialog select menu **Process:Protein Statistics**.



Alternatively, the statistical analysis dialog can be opened using the toolbar icon.

The dialog box contains various options allowing statistical analysis of protein data from spot maps loaded in the DeCyder Differential Analysis Software BVA.



4.8.2 Defining groups

A group in BVA is a collection of spot maps that for the purposes of the experiment cannot be broken down into further sub groups. For example, 3 replica gels from one sample are considered one group. Alternatively, 3 gels of 3 different samples treated with exactly the same experimental conditions can also be considered as a group. There is no limit on the number of groups that can be assigned in DeCyder Differential Analysis Software BVA.

The different experimental groups must first be identified in DeCyder Differential Analysis Software BVA in order to undertake statistical analysis of protein expression changes (see section 4.5.2).

4.8.3 Overview of Statistical tests

The BVA module has four “principle groups comparison” or “experimental condition comparison” methods that can be applied to analyze protein spot data.

- Average ratio between two groups or two populations of groups.
- Student’s T-test: Statistical analysis between two groups or two populations of groups.
- One-Way ANOVA (ANalysis Of VAriance): Statistical analysis between all groups.
- Two-Way ANOVA: Statistical analysis between the two conditions in an experimental design where there are two independent factors (e.g. time-dose study). This analysis allows the internal and mutual effects of the two factors to be quantified.

The log standardized abundance is the only variable subjected to the above statistical analyses within DeCyder Differential Analysis Software BVA. The standardized abundance is derived from the normalized spot volume standardized against the intra-gel standard. The log of the standardized abundance values are used in order that the data points approach a normal distribution around zero, thereby fulfilling the requirements of the subsequent statistical tests.

Consequently, the statistical analysis functionality is not valid unless the experimental design includes an internal standard on every gel.

General requirements

All spot maps included in the analysis need to be co-run with an internal standard.

Only Spot Maps assigned as “Analysis” in the Spot Map Table are included in the statistical analysis.

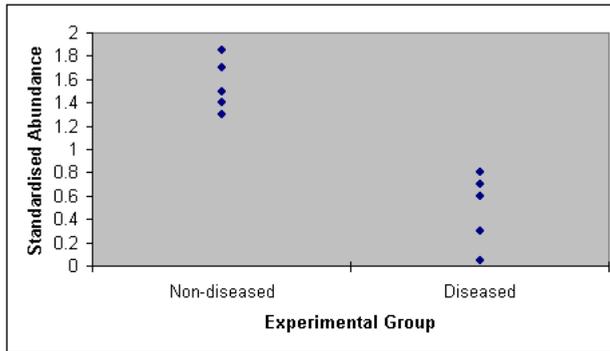
Automatic recalculation

Protein Statistics are recalculated automatically if any data that affects the statistics are changed, e.g. if a match is broken or if a spot map is assigned to a new group.

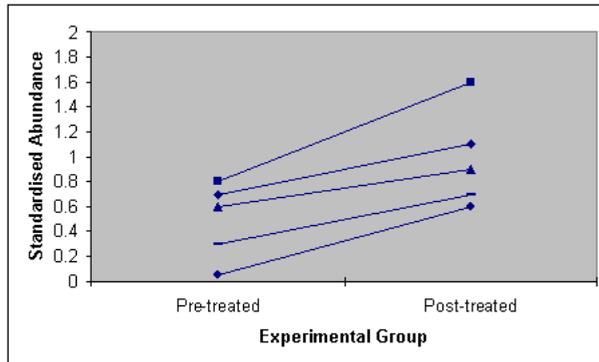
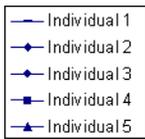
4.8.4 Independent and paired analyses

The Student's T-test, One-Way ANOVA and Two-Way ANOVA can be either independent (normal) or paired (individual). Both types are used to test the hypothesis that a variable, in this case protein abundance, differs between groups or experimental conditions.

The paired test is specifically used when each data point in one group corresponds to a matching data point in the other group(s). A typical example would be the same group of patients before and after a treatment. The unpaired tests are more general techniques that can be used to test whether standardized protein abundance differs between groups, and does not require that the groups be paired in any way, or even of equal sizes. The following graphs are examples illustrating independent and paired tests.



Independent test. Abundance of a specific protein in five diseased individuals compared to a group of five non-diseased individuals. Data points are independent since there are different individuals in each group.



Paired test. Abundance of a specific protein in the above five diseased individuals before and after drug treatment. Data points are paired because the same individuals are present in pre- and post-treated groups. Although there is no significant difference between the means of standardized abundance in the two groups, a paired T-test reveals that the two groups are significantly different, since each individual exhibits an increase in protein abundance after drug treatment. In this example, there is a single result (i.e. one image) for each individual in each group. Experiments employing replicate images from the five individual scan also be applied to a paired test. In this instance the means of the replicas are calculated then subjected to pairing (ensuring the pairs are independent).

Note: The above graphs are conceptual (not generated in DeCyder Differential Analysis Software).

In order to perform paired testing, the pairing of spot maps must be pre-assigned in the Spot Table view. This can be done by inserting a numerical identifier in the text box labelled **Sample ID** at the bottom of the Spot Table view. For example, spot maps of protein samples from individual 2 in each of the groups can be assigned. In this way the dependency can be associated with each spot map.

4.8.5 Student's T-test & Average ratio

Average ratio

Log standardized protein abundance is the only variable analyzed in the Student's T-test within DeCyder Differential Analysis Software BVA module. The degree of difference in the standardized abundance between 2 protein spot groups is expressed as the average ratio. The average ratio value indicates the standardized volume ratio between the two groups or populations. Values are displayed in the range of $-\infty$ to -1 for decreases in expression and $+1$ to $+\infty$ for increases in expression. Values between -1 and 1 are not represented, hence a two-fold increase and decrease is represented by 2 and -2 , respectively (not 2 and 0.5).

Note: The average ratio parameter is displayed in this manner but the statistical analyses are based on the log of the true ratio measurement.

Student's T-test

Student's T-test, often known simply as the T-test, is one of the most commonly used of all statistical tests. The Student's T-test is used to test the hypothesis that a variable differs between two groups. The Student's T-test performed is an equal variance two-tailed test, therefore, direction of change (i.e. increases and decreases) in the standardized abundance parameter are considered.

The T-test requires a minimum of two data points in each of the two groups. Greater statistical validity can be achieved using larger number of replicates. It is recommended that the largest possible number of biological replicates are performed for optimal validity. However, if biological replicates are not possible, gel replicates can be used for the purposes of the statistical analysis. In these instances it is recommended that gels be run, at least, in triplicate, hence each group has a minimum of three data points.

Null hypothesis

The Student's T-test null hypothesis is that there is no change in the protein abundance between experimental groups (i.e. the average ratio between two groups is 1). Therefore the T-test p value (seen in the T-test column of the protein table) represents the probability of obtaining the observed data if the two groups have the same protein abundance. For example, if the T-test p value between two groups is 0.01 , then the probability of obtaining the observed difference in protein abundance by stochastic variation alone is 1 in 100 .

Protein abundance differences are generally assumed to be statistically significant when $p \leq 0.05$. However a critical value of 0.05 would mean that 5% of the data points would be expected by stochastic events alone

(e.g. 50 spots out of 1000 tested). It is therefore advisable to review the critical value applied to the data. Due to the small amount of experimental variation within Ettan DIGE system and subsequent DeCyder Differential Analysis Software analysis, a critical value of 0.01 is often applied.

Assumptions and limitations

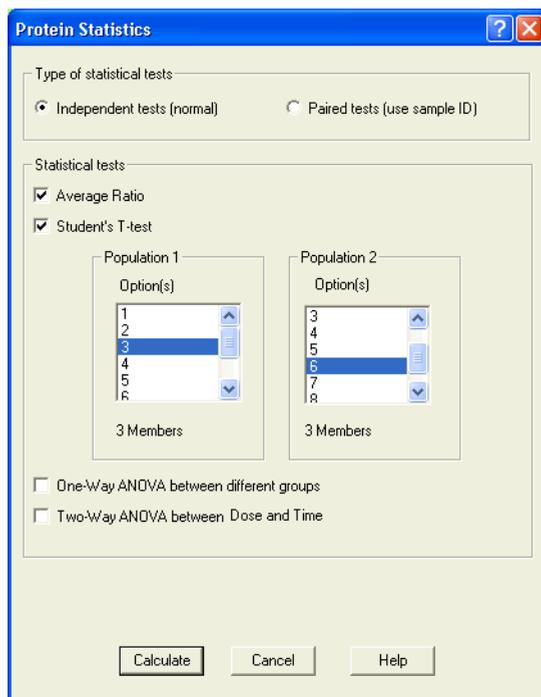
The main assumptions of the Student's T-test are the following:

- Log standardized abundance values within a group are normally or approximately normally distributed. When the assumption of normality is violated, the T-test tends not to perform well when the sample size is small, or the significance level is small (e.g. $p < 0.01$). However, two-tailed tests are surprisingly robust even for skewed data.
- The parametric variances of the two groups are equal. However, the test tends to be robust when the sample sizes are approximately equal. If the sample sizes are not equal, then the worst case occurs when the smaller sample comes from the population with the larger variance.

Performing analysis

The user can perform T-test calculations and average ratio calculations between any two groups (or populations of groups).

- 1 In the **Protein Statistics** dialog box, select the option button to indicate whether the T-test is independent or paired. Pairing of spot maps has to be pre-assigned in the Spot Map Table mode (see section 4.8.4) for paired testing.
- 2 Select the two groups in the in the population 1 and population 2 lists (population of groups can be selected by ctrl + left mouse clicking the selected groups). Select the **Average ratio** and the **Student's T-test** check boxes. Click **Calculate**.



The T-test p values and average ratios will be displayed in the protein table.

- 3 The spots in the Protein Table can then be sorted by clicking on the **T-test** column header. This will order the table so that the protein spots exhibiting the most significant changes are listed at the top of the table. Since the T-test is two-tailed the statistical significance of both increases and decreases in protein abundance quantified are ordered together. A positive average ratio value indicates an increase from population 2 to population 1 (the order is stipulated when defining groups in the protein statistics dialog box). Conversely a decrease in abundance is denoted by a negative average ratio.

Protein Table T-test and Av.Ratio: 6 / 3												
Pos.	Master No.	Status	Protein ID	Protein AC	Name	Appearance	T-test	Av. Ratio	1-ANOVA	2-ANOVA-Dose	2-ANOVA-Time	2-ANOVA-Int
16	230	Unconfirmed				52 (60) A, M, P	0.0007	-1.35				
17	1071	Unconfirmed				44 (60) A, M	0.0010	1.93				
18	652	Unconfirmed				38 (60) A, M	0.0011	-1.37				
19	1031	Unconfirmed				44 (60) A, M	0.0015	1.97				
20	214	Unconfirmed				50 (60) A, M	0.0019	-1.45				
21	801	Unconfirmed				52 (60) A, M	0.0019	-1.36				
22	207	Unconfirmed				50 (60) A, M	0.0021	1.42				
23	638	Unconfirmed				38 (60) A, M, P	0.0021	-1.10				
24	430	Unconfirmed				46 (60) A, M, P	0.0022	1.18				
25	525	Unconfirmed				50 (60) A, M	0.0022	-1.06				
26	237	Unconfirmed				42 (60) A, M, P	0.0025	-1.62				
27	728	Unconfirmed				60 (60) A, M, P	0.0026	1.10				
28	1078	Unconfirmed				38 (60) A, M	0.0026	1.32				
29	202	Unconfirmed				50 (60) A, M, P	0.0027	1.88				
30	592	Unconfirmed				54 (60) A, M, P	0.0028	1.10				
31	765	Unconfirmed				36 (60) A, M	0.0028	-1.22				
32	1089	Unconfirmed				26 (60) A, M	0.0028	1.24				
33	894	Unconfirmed				54 (60) A, M	0.0035	1.08				
34	514	Unconfirmed				48 (60) A, M, P	0.0040	1.09				
35	597	Unconfirmed				58 (60) A, M, P	0.0040	1.09				
36	576	Unconfirmed				58 (60) A, M, P	0.0046	-1.13				
37	715	Unconfirmed				38 (60) A, M	0.0046	1.29				
38	730	Unconfirmed				38 (60) A, M	0.0046	1.48				
39	391	Unconfirmed				38 (60) A, M	0.0048	-1.28				
40	541	Unconfirmed				48 (60) A, M	0.0052	-1.14				
41	223	Unconfirmed				40 (60) A, M	0.0054	-1.46				
42	624	Unconfirmed				40 (60) A, M, P	0.0055	1.16				
43	115	Unconfirmed				30 (60) A, M	0.0057	1.84				
44	1020	Unconfirmed				46 (60) A, M	0.0057	-1.14				
45	203	Unconfirmed				56 (60) A, M, P	0.0059	1.76				
46	625	Unconfirmed				60 (60) A, M, P	0.0060	1.07				
47	860	Unconfirmed				50 (60) A, M, P	0.0060	-1.15				
48	734	Unconfirmed				60 (60) A, M, P	0.0061	-1.12				
49	341	Unconfirmed				58 (60) A, M	0.0062	-1.07				
50	1035	Unconfirmed				36 (60) A, M	0.0062	2.94				
51	581	Unconfirmed				54 (60) A, M, P	0.0066	-1.13				

Specific Requirements

If there are at least two members in each group, T-tests and average ratios will be calculated. If there is only one member in each group only the “average” ratio will be calculated.

4.8.6 ANOVA

Overview

Analysis of variance (ANOVA) is a family of methods that can be used to analyze the results from both simple and complex experiments. It is one of the most important statistical tests available for biologists and at its lowest level it is essentially an extension of the logic of Student’s T-tests to those situations where the concurrent comparison of the means of three or more samples is required. Thus, when comparing two means, the ANOVA will give the same results as the T-test for independent samples (if comparing two different groups or observations). There is no restriction on the number of groups that can be analyzed. It is equally valid for testing differences between 2 groups or among 20.

There are two types of ANOVA analyses in DeCyder Differential Analysis Software BVA module, One-Way and Two-Way. The One-Way ANOVA evaluates differences in all assigned groups, whereas the Two-Way ANOVA can also evaluate the statistical significance of effects

from independent factors (such as time, temperature and dose). These are entered as conditions 1 and 2.

The ANOVA tests require a minimum of two replicate data points. Greater statistical validity can be achieved using larger number of replicates. It is recommended that the largest possible number of biological replicates are performed for optimal validity. However, if biological replicates are not possible, gel replicates can be used for the purposes of the statistical analysis. In these instances it is recommended that gels be run, at least, in triplicate, hence each group has at least three data points.

As with the Student's T-test the ANOVA tests can be either independent or paired. Paired ANOVA tests require dependencies to be applied (see section 4.8.4).

Null hypothesis

The ANOVA test null hypothesis is that there is no change in the protein abundance between any of the experimental groups analyzed. The ANOVA test compares the variance between groups with the variance within groups. The ratio of the between-groups variance to the within-groups variance is known as the F ratio, and can be used to generate a p value. If the F ratio is large, then it shows that the variation between groups is large compared with the variation within groups, and thus that the groups may be different. Therefore the p value (displayed in the ANOVA column of the protein table) is a measure of the probability of obtaining the observed data if the groups have the same protein abundance. For example, if the ANOVA p value between two groups is 0.01, then the probability of obtaining the observed difference in protein abundance by stochastic variation alone is 1 in a 100. Protein abundance differences are generally assumed to be statistically significant when $p \leq 0.05$. However a critical value of 0.05 would mean that 5% of the data points would be expected by stochastic events alone (e.g. 50 spots out of 1000 tested will be false positives). It is therefore advisable to review the critical value applied to the data. Due to the small amount of experimental variation within Ettan DIGE system a critical value of 0.01 is often applied.

Assumptions and limitations

The main assumptions of the ANOVA tests are the following:

- The populations from which the samples were obtained are normally or approximately normally distributed.
- The variances of the populations must be equal.

There is a certain amount of leeway in both these assumptions, particularly the first: small departures from a normal distribution are

unlikely to be of any consequence. However, if the variance among the groups differs greatly, the ANOVA p value may have poor validity.

4.8.7 One-Way ANOVA

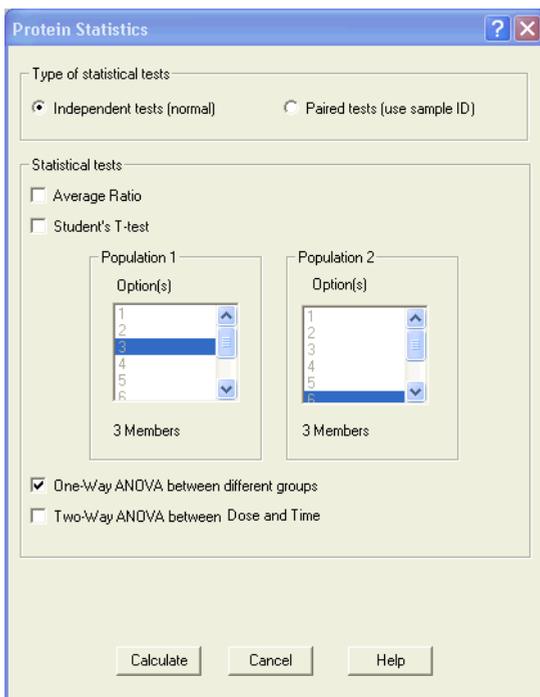
Overview

The simplest form of ANOVA is known as one-way ANOVA, and in DeCyder Differential Analysis Software is used to test for differences in standardized abundance. For example, a clinical trial might consist of three groups of patients according to whether they were given drug A, drug B, or placebo. ANOVA could be used to test the null hypothesis that there is no difference in standardized protein abundance among the groups.

The test will not indicate which groups are different from which other groups, just that there is an overall difference. In the above example, a significant difference does not necessarily mean that drug A is different from drug B, or that either drug is different from placebo. All a One-Way ANOVA test indicates is that, in some way, the patients are affected by one of the drugs.

Performing analysis

- 1 In the **Protein Statistics** dialog box, select the option button to indicate whether the test is independent or paired. A paired ANOVA test is called a repeated measures ANOVA (RMANOVA). Pairing of spot maps have to be pre-assigned in the spot map Table View (see section 4.8.4) for paired testing.
- 2 Select the **One-Way ANOVA** check box (all assigned groups are included in the One-Way ANOVA test). **Click Calculate**.
- 3 Select the **Properties** icon on the tool bar and ensure that **1-ANOVA** check box is selected in the **Spot Map Table** tab. If the One-Way ANOVA analysis is paired, select the **Paired tests** option button. A repeated measures (paired) One-way ANOVA will then be automatically calculated. Click **Calculate** to perform the analysis.



The One-Way ANOVA p values will then be displayed in the protein table. As with the T-test analysis, spots can be ordered by significance by clicking on the 1-ANOVA column heading in the protein table.

Specific requirements

All groups that have at least two members will be included in the calculation. To analyze a subset of groups, those that are to be excluded have to be de-selected as analysis in the Spot Map function tab (see section 4.5.2).

4.8.8 Two-Way ANOVA

Overview

Two-Way ANOVA calculates the significance of the difference between groups with the same condition 2 and different condition 1 value (Two –Way ANOVA Condition 1) and the other way around (Two –Way ANOVA Condition 2). The Two-Way ANOVA analysis also calculates a significance value of the mutual effect of the two factors (Two –Way ANOVA Interaction).

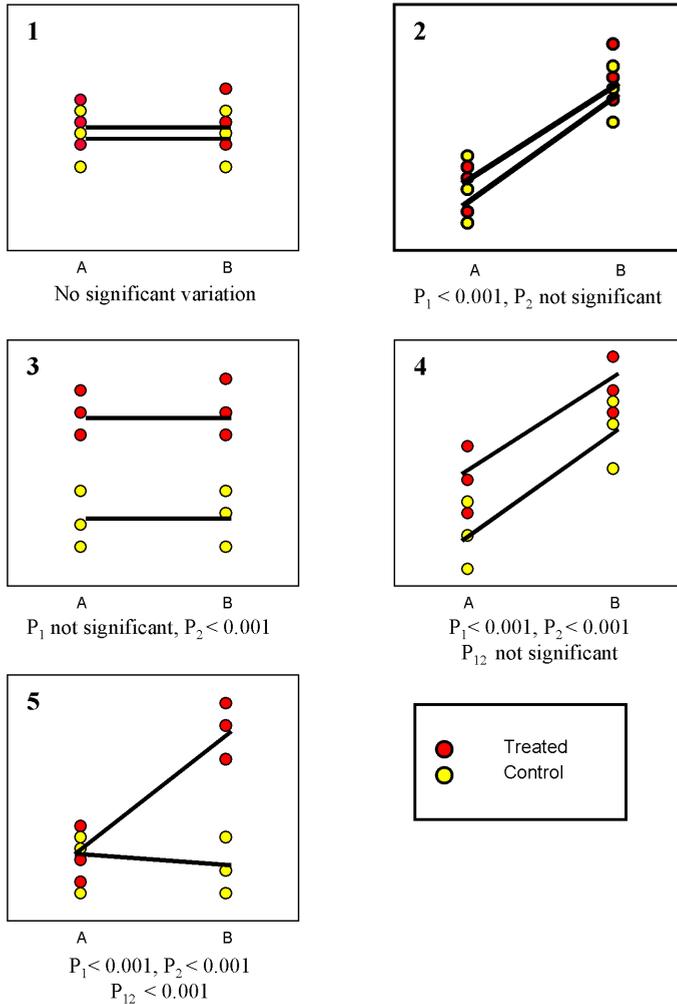
There are therefore three sets of hypothesis with the Two-Way ANOVA. The null hypotheses for each of the sets are given below.

The population means of the first condition are equal. This is like the One-Way ANOVA for condition 1 exclusively.

The population means of the second condition are equal. This is like the One-Way ANOVA for condition 2 exclusively.

There is no interaction between the two factors. A significant ANOVA Interaction value indicates that the two factors affect each other due to synergy or interference.

Examples of such effects are illustrated on next page:



Graphical examples of Two-Way ANOVA analyses.

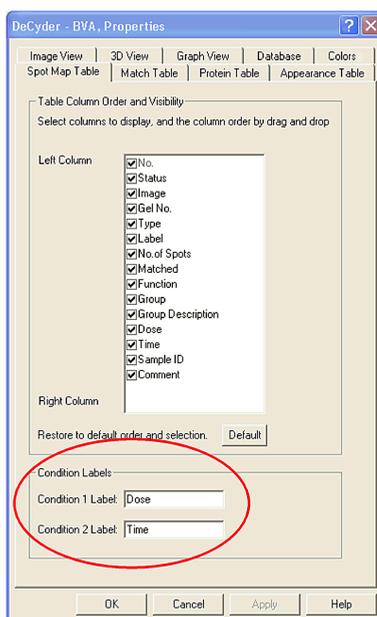
Graphs illustrate changes in protein abundance (y-axis) for a two-condition experiment. Condition 1 (x-axis) represents two temperatures A and B; condition 2 (red and yellow circles) represents drug treated and control samples. Each condition is in triplicate, hence there are four experimental groups with 3 samples in each group. Conditions 1 and 2 are used to link groups together based on one common factor, i.e. group 1 and 2 may have the same condition 1 value (both temperature 1) but different condition 2 value (drug treated or control). Groups 3 and 4 will have the other condition 1 value (both temperature 2) with different condition 2 values (drug treated or control).

- Experiment 1, there is no variation in abundance from temperature A to B or from treated and non-treated samples.
- Experiment 2, temperature B results in increased protein expression whereas drug treatment has no effect.
- Experiment 3, drug treatment results in increased protein expression whereas temperature changes have no effect.
- Experiment 4, both drug treatment and increasing temperature result in increased protein expression i.e. the two conditions are independent and P_{12} is not significant.
- Experiment 5, drug treatment only in conjunction with change in temperature results in an increased protein abundance i.e. the two conditions are synergistic hence $P_{12} < 0.001$.

P_1 , P_2 and P_{12} are labelled as **2-ANOVA-“condition 1”**, **2-ANOVA-“condition 2”** and **2-ANOVA-interact** in the Protein Table.

Defining conditions

In order to perform Two-Way ANOVA analysis the experimental conditions for each gel image must first be assigned. Prior to this the condition labels can be defined by selecting the **properties** icon and selecting the **Spot Map Table** tab, the condition label can then be entered into the appropriate text boxes.

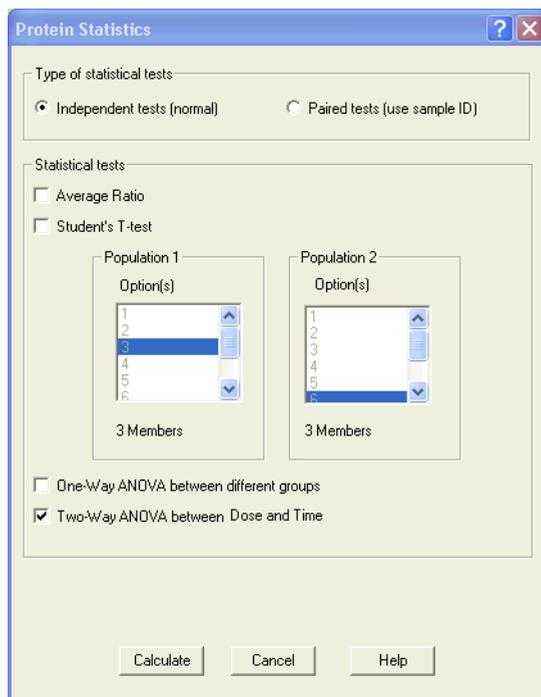


The experimental conditions of each gel image can then be entered at the bottom of the spot map view. Conditions can be entered as integers to represent the exact conditions (e.g. dose 1,2,3...n can represent a series of 5 mg increments in a drug treatment). Up to two conditions can be designated for each spot map.

Note: It can, therefore, be seen that when assigning a two condition experimental design that the Condition labels must be defined for a Two-Way ANOVA test. When assigning a single condition experiment the data is assigned into Groups for One-Way ANOVA testing.

Performing analysis

- 1 In the **Protein Statistics** dialog box, select the option button to indicate whether the test is independent or paired. Pairing of spot maps has to be pre-assigned in the Spot Map Table View (see section 4.8.4) for paired testing.
- 2 Select the Two-Way ANOVA check box. Click **Calculate**.
- 3 Select the **Properties** icon on the tool bar and ensure that **2-ANOVA-“condition 1”**, **2-ANOVA-“condition 2”** and **2-ANOVA-interact** check boxes are selected in the spot map tab. If the Two-Way ANOVA analysis is paired, select the **Paired tests** option button. A repeated measures (paired) Two-way ANOVA will then be automatically calculated. Click **Calculate** to perform the analysis.



The Two-Way ANOVA p values will then be displayed in the protein table. As with the other analyses, spots can be ordered by significance by clicking on the appropriate 2-ANOVA column headings in the protein table.

Specific requirements

At least two experimental groups with at least two members in each are required. The experimental design should also be set up so that the following criteria are applied:

- Condition 1: The experimental groups should have both condition values filled in, so that there are at least two groups with different condition 1 values and the same condition 2 value.
- Condition 2: The experimental groups should have both condition values filled in, so that there are at least two groups with different condition 2 values and the same condition 1 value.

Additional requirement for paired Two-Way ANOVA:

- At least two different sample IDs in each group (spot maps with the same condition 1 and condition 2 value).

For further information on experimental design and set up examples, see Appendix D, Experimental design and set up examples.

4.8.9 Further statistical analyses

The data generated within DeCyder Differential Analysis Software can be exported as an XML file and then extracted using the XML toolbox, enabling post processing of several parameters. The user can therefore apply their own statistical analysis algorithms to normalized and pre-normalized data to suit their own specific requirements.

4.9 Protein Filter



The protein filter is a tool that allows the selection of proteins based on various user-defined criteria. To open the protein filter select **Process: Protein Filter** or click the protein filter icon. The filter allows proteins to either be given a **Protein of Interest** or **Pick Status** by selection of the appropriate check box.

Filter action

Assign Proteins of Interest Assign Pick status

The general filter settings allows the selection of all detected proteins by selecting the **Select all** check box. This panel within the filter also permits the restriction of the filtering to only a subset population of proteins either those confirmed or those present on a specified number of spot maps.

General filter settings

Select all

Restrict to confirmed proteins

Restrict to proteins present in spot maps

Proteins can also be selected on the statistical parameters described in section 4.8. The desired statistical analyses check boxes are selected and the desired values can then be entered in the appropriate text boxes.

Select proteins with

<input checked="" type="checkbox"/> Student's T-test value	<=	<input type="text" value="0.05"/>	
<input checked="" type="checkbox"/> Average Ratio	>=	<input type="text" value="1.5"/>	
or			
<input checked="" type="checkbox"/> Average Ratio	<=	<input type="text" value="-1.5"/>	
<hr/>			
<input type="checkbox"/> Average Ratio	>=	<input type="text" value="-2"/>	and <= <input type="text" value="2"/>
<hr/>			
<input checked="" type="checkbox"/> One-way ANOVA value	>=	<input type="text" value="0"/>	and <= <input type="text" value="0.05"/>
<input checked="" type="checkbox"/> Two-way ANOVA - Time value	>=	<input type="text" value="0"/>	and <= <input type="text" value="0.05"/>
<input checked="" type="checkbox"/> Two-way ANOVA - Dose value	>=	<input type="text" value="0"/>	and <= <input type="text" value="0.05"/>
<input checked="" type="checkbox"/> Two-way ANOVA - Interaction value	>=	<input type="text" value="0"/>	and <= <input type="text" value="0.05"/>

Proteins can also be selected on the basis of their physical properties, such as **volume**, **X-co-ordinate** and **Y-co-ordinate**. These physical properties pertain to the spots on the **Master spot map** or the **Pick spot map** (if present) depending on which is selected from the pull down menu.

Properties for proteins in spot map

<input type="checkbox"/> Volume	>=	<input type="text" value="1.00e+005"/>	and <=	<input type="text" value="1.00e+008"/>
<input type="checkbox"/> X co-ordinate	>=	<input type="text" value="0"/>	and <=	<input type="text" value="1000"/>
<input type="checkbox"/> Y co-ordinate	>=	<input type="text" value="0"/>	and <=	<input type="text" value="1000"/>

Note: The use of the protein filter in selecting spots for picking is further discussed in section 5.

4.10 Molecular weight (Mw) and isoelectric point (pI)

DeCyder Differential Analysis Software BVA provides the user with the functionality to calculate and display the molecular weight (Mw) and isoelectric point (pI) of the proteins on all the images being analyzed. The Mw and pI of known proteins can be entered into the Protein Table, these values (known as the base list) are then used to calculate the Mw and pI for all the other proteins on the images.

4.10.1 Entering Mw and pI of known proteins

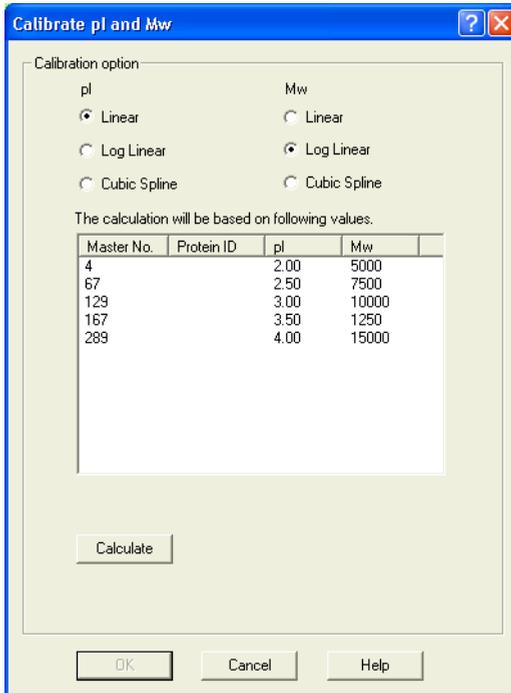
To enter the Mw and pI of the known proteins in your experiment select the protein in the **Protein Table (PT)** and enter the Mw and pI into the **Mw** and **pI** text boxes, in the data control panel. The Mw and pI entered for a protein can be edited by clicking on the protein and changing the values in the **Mw** and **pI** boxes.



Calculating Mw and pI

To calculate the Molecular weight (Mw) and isoelectric point (pI) of all the proteins on the images select **Process:Calibrate pI and Mw** or click the **Calibrate pI and Mw** icon.

The subsequent dialog box, displays the pI and Mw values that have been manually entered into the protein table, using the data control panel in the Protein Table mode. These values are used as a base list to calculate the pI and Mw values for all the proteins.



Calibration of the pI and Mw values can be user determined by selecting **Linear**, **Log Linear** or **Cubic Spline**:

- **Linear:** Creates a straight line between all values in the base list.
- **Log Linear:** Similar to the linear function, except that the values in the base list values are first logged (Log_{10}). So the line is based on logarithmic values.
- **Cubic Spline:** Creates a line using the cubic spline function, which fits a curved line through all points in the base list.

The type of first dimension Isoelectric Focusing (IEF) strip influences pI calibration used. For example, a linear Isoelectric Focusing (IEF) strip

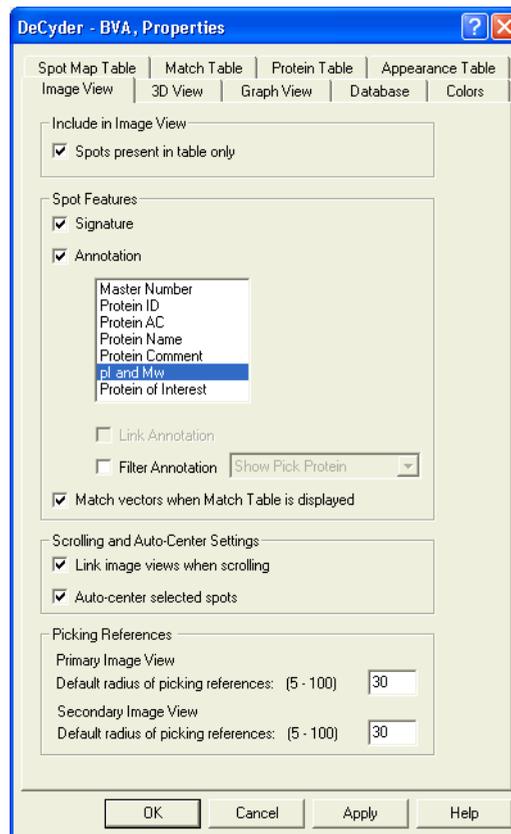
makes the Linear pI calibration method most suitable.

Mw calibration is influenced by the second dimension gel used. For example, Log Linear Mw calibration method is most suitable for non-gradient gels.

After choosing one of the three available options for calculating the pI and Mw click **OK** to perform the calculation. The pI and Mw of the proteins should now appear in the **Protein Table**.

Displaying the Mw and pI values on the images

To display the calculated pI and Mw values of the proteins on the gel images, click on the **Properties** icon to display the **Properties** window. By default the window should open on the **Image View** tab. Select the **Annotations** check box then highlight the **pI and Mw** option from the list of different annotations. Click **OK** to display these values on the images.



Use the **Zoom in** function to make the annotations more legible.

4.11 User-defined protein labelling

4.11.1 Confirmation

Protein confirmation allows the marking of proteins for user reference purposes (e.g. for visually checked proteins). No part of the analysis in DeCyder Differential Analysis Software BVA demands confirmation of proteins.

To confirm a protein, select the **Confirm** button in the data control panel.

If the selected protein is confirmed, it can be unconfirmed by selecting the **Unconfirm** button.

4.11.2 Name

User-defined names can be entered through the text box in the data control panel.

4.11.3 Comment

User-defined comments can be entered through the text box in the data control panel.

4.11.4 Protein of interest

The Protein of Interest check box is selected for proteins that may warrant further investigation by the user.

4.11.5 PTM

Protein spots possessing a post-translational modification (PTM) identified by methods outlined in Ettan DIGE application notes (e.g. application note 18-1170-83 AA, obtainable on the Amersham web site) can be denoted by checking the PTM check box.

4.12 Database linking

DeCyder Differential Analysis Software provides the user with functions that enable rapid access to both local and web based databases through Protein IDs and Accession Numbers (AC). This functionality is designed to aid further study of identified proteins using information from existing databases. A list of web based databases are provided as default in DeCyder Differential Analysis Software. The linking uses the default web browser to interact with the databases. The Protein ID or AC is substituted into the web address of the database. On condition that the protein ID entered in the protein table is in the format that is recognized by the search algorithm of a database the user

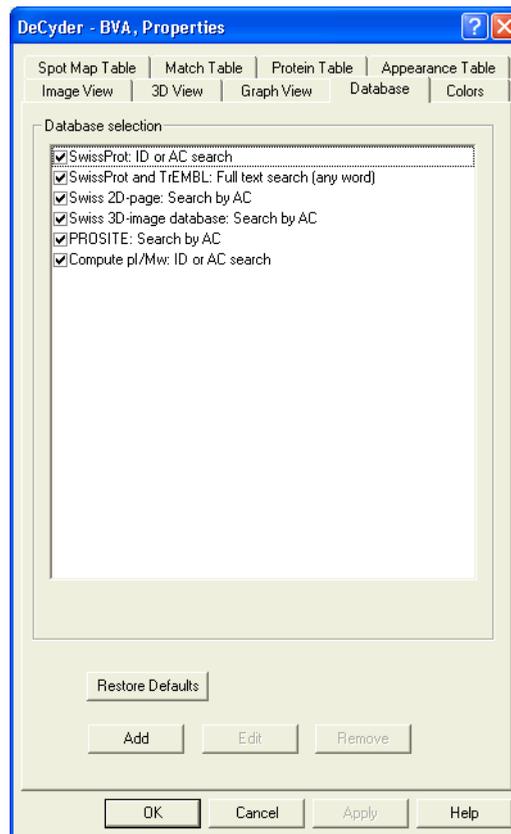
can directly access a selected protein from BVA.

In DeCyder Differential Analysis Software BVA module it is possible to link proteins to information in databases

Use the Database tab of the BVA module property pages to set up how Protein IDs and ACs are inserted in the web address of different databases.

4.12.1 Adding databases

To add a database to the default list of databases in BVA, click on the **Properties** icon (or select **View:Properties**) to open the Properties window then click on the **Database** tab.



To add a database to the list click on the **Add** button.

Add new web database label and address

Label

Address

A placeholder string has to be included in the address, and will be replaced by the protein's ID or AC.
E.g. http://www.expasy.ch/cgi-bin/get-sprot-entry?#BVA_PROTID#

#BVA_PROTID#	- Protein ID
#BVA_PROTAC#	- Protein AC
#BVA_PROTIDAC#	- Both Protein ID and AC

Test

Select a Protein ID or AC that will replace the placeholder string.

OK Cancel Help

The name of the database is entered in the **Label** text box. The web address (URL) in the format specified is entered in the address text box. The address should contain a placeholder string, which will be replaced by the protein's ID or Accession Number (AC) depending on the string is used and the information available.

There are three different placeholder strings:

- #BVA_PROTID# - The database will be available if there is information in the Protein ID column of the Protein Table.
- #BVA_PROTAC# - The database will be available if there is information in the Protein AC column of the Protein Table.
- #BVA_PROTIDAC# - The database will be available if there is information in either the Protein ID or the Protein AC column of the Protein Table.

Upon selecting a protein, only databases that have placeholder strings that fit the information entered for the protein will be available.

To verify that the address is entered correctly, select a protein in the **Test** pull down menu and click the **Test** button. If the address is correct, the browser should open and information on the test protein should be shown.

Example

Database - SwissProt: AC or ID search

Address (with substitution string) - http://www.expasy.ch/cgi-bin/get-sprot-entry?#BVA_PROTIDAC#

Protein AC - (Human myoglobin)- P02144

The web browser will open at the address:
<http://www.expasy.ch/cgi-bin/get-sprot-entry?P02144>

If a system for identifying proteins other than AC or ID is being used, text can be entered manually or pasted into the test place holder string box.

Accessing databases

Ensure that the required database is selected in the **Properties** window by clicking on the **Properties** icon and selecting the **Database** tab. Select the desired database(s) from the **Database Selection** list then click **OK**.

To access a database right mouse click the protein to be queried in the Protein Table, then select the database. If both Protein ID and Protein AC labels are available a choice of the label to be used is given. The web browser is then opened at the appropriate address.

***Note:** The linking is dependent on the place holder string being compatible with the selected database.*

Removing and editing databases

Removing and editing are performed in the **Database Properties** dialog box.

To remove a database, highlight the selected database and click the **Remove** button.

To edit a database, highlight the selected database and click the **Edit** button. The database can then be amended. Click **OK** to accept the changes.

4.13 Viewing protein data

Information relating to proteins (such as those described in the previous four sections) can be viewed through the Protein Table and the Appearance Table. The Protein Table displays several proteins simultaneously, allowing rapid comparison of data from different proteins. The appearance table displays a single protein showing data from individual gels where the protein occurs.

Data can also be displayed by annotating protein spots in the gel images view, which can be a useful tool for presentation of data.

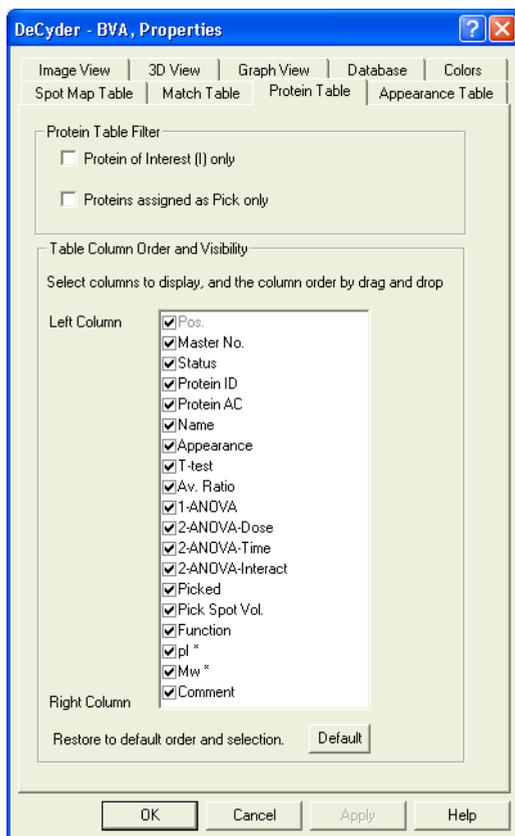
4.13.1 Protein Table

The table below describes the information contained in the Protein Table:

Pos.	Protein Table row number
Master No	Protein spot number on Master Spot Map
Status	Confirmation status of protein
Protein ID	Protein identification for database linking
Protein AC	Protein accession number for database linking
Appearance	Indicates the number of Spot Maps in which the selected spot appears, and the function these Spot Maps have Example: "20(20) A, M" means that the spot is present in 20 out of 20 Spot Maps. Furthermore, the Spot Maps are all assigned to be used in statistical analysis. The spot is also present on the master image.
T-test	p value calculated using the Student's T-test
Average ratio	Average ratio between the groups selected in the protein statistics dialog box
1-ANOVA	p value calculated using One-way ANOVA statistical test
2-ANOVA – "Condition 1"	p value calculated using Two-way ANOVA statistical test with respect to Condition 1
2-ANOVA – "Condition 2"	p value calculated using Two-way ANOVA statistical test with respect to Condition 2
2-ANOVA-Interact	p value calculated using Two-way ANOVA statistical test with respect to interactions between conditions 1 and 2
Picked	Indicates whether the protein spot has been assigned for picking
Table continued on next page	

Pick spot volume	Displays the volume of a spot designated for picking after spot detection and background correction. This value can be used as an indication of whether a spot can be identified by mass spectrometry
Function	Proteins assigned Protein of Interest are marked with an 'I' in this column.
pI	Isoelectric point of protein (can be user-defined or calculated)
Mw	Molecular weight of the protein (can be user-defined or calculated)
Name	User defined protein names
Comment	User defined notes on proteins
Match Quality	Morphological similarity metric describing deviation of internal standard spot to an "average" internal standard spot in protein set.
PTM	Indicates that the protein has a post translational modification.

The columns displayed in the protein table can be changed using the **Protein Table Properties** dialog box.



Protein Table Filter: Select check boxes to display only proteins of interest or pick assigned proteins in the Image View and the Table View. Both check boxes can be selected if there are spots selected as proteins of interest and pick status.

Table Column Order and Visibility: The selected column titles are displayed in the Table View. Clicking and dragging the column titles in the list sets the order of the columns in the Table View. Clicking **Default** restores the original settings.

4.13.2 Appearance Table

The table below describes the information contained in the Appearance Table:

No.	Spot map number
Image	Gel image file name
Type	Indicates the dye chemistry used
Label	Indicates the CyDye DIGE Fluor minimal dye label
Function	Spot map function (e.g. Master, Analysis, Pick and Template)
Abundance	Relative volume of all spots representing a particular protein in a BVA data set, used when no internal standard is present.
Std Abundance	Protein volume calculated relative to internal standard
Volume	Spot pixel volume (expressed background subtraction)
Peak Height	Largest pixel value within the spot boundary (expressed background subtraction)
Group	Assigned spot map group
Group Description	Description of spot map group
Condition 1	Condition 1 numerical value
Condition 2	Condition 2 numerical value
Sample ID	User defined sample identification
Comment	User defined comment
Match Quality	Morphological similarity metric describing deviation of internal standard spot to an "average" internal standard spot in protein set.

The columns displayed in the appearance table can be changed using the **Appearance Table Properties** dialog box (see next page).

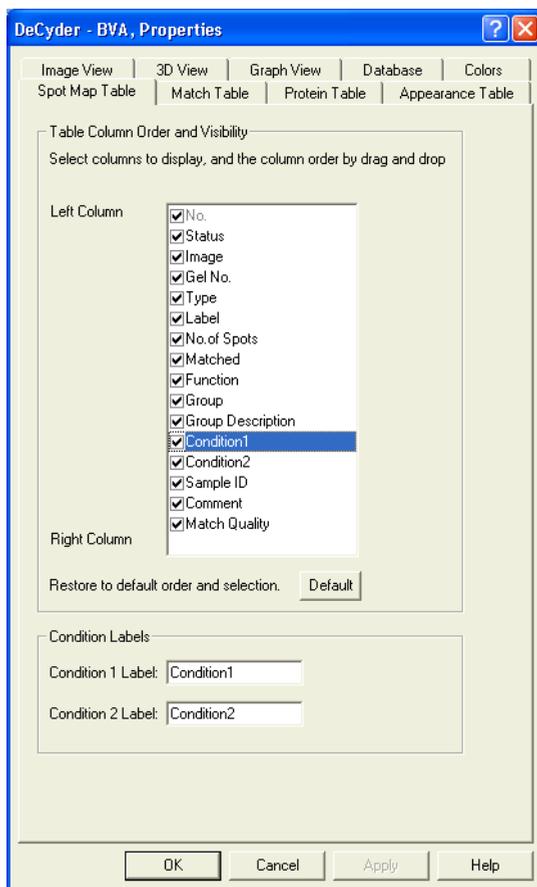


Table Column Order and Visibility: The selected column titles are displayed in the Table View. Clicking and dragging the column titles in the list sets the order of the columns in the Table View. Clicking **Default** restores the original settings.

4.13.3 Spot annotation

Displaying annotation

Various protein data elements can be displayed on the Image View. The annotation can be defined using the **Image View Properties** dialog box. See Section 4.4.1 for details.

Positioning the annotations

The position of the annotations in the Image View can be altered to ensure that the different annotations are clearly visible. This feature is useful when one annotation obscures another annotation. To alter the position of an annotation, select **Edit:Move Annotations** so that a tick appears next to this option. The annotation labels in the Image View can then be dragged to the new positions.

4.13.4 Customizing display colors

The colors used in the various views can be customized by selecting the **Colors** tab in the **Properties** dialog box.



Match Colors: The spot boundary colors of the different categories of spots displayed in the Match Table Image View can be changed by clicking on the colored circles then selecting a new color.

Spot Colors: Spot boundary colors of the different categories of spots

displayed in the Protein Table and Appearance Table Image Views can be changed by clicking on the colored circles then selecting a new color.

Annotation Colors: The colors of the gel image annotation text boxes and linking lines can be changed by clicking on the colored circles then selecting a new color.

Spot Map Colors: The colors that designate different categories of spot map in the Image Views can be changed by clicking in the colored circles then selecting a new color.

The original settings can be restored by clicking **Default**.

4.13.5 Saving, exporting and printing

Save As

Workspaces are saved as BVA files, which contain all the data associated with all the loaded spot maps. The BVA files do not contain the image files but do contain information on their file path.

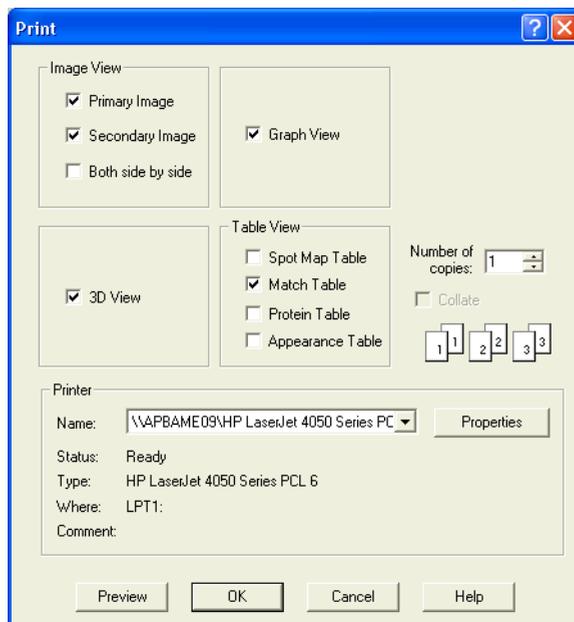
Select **File:Save As**, and browse to locate the folder to save the workspace. Enter a file name then click **Save**.

Save

Save workspace updates the saved version of the file to include any changes that have been made to the workspace.

Printing

Select **File:Print** to open the **Print** dialog box.



Multiple check boxes can be selected for printing the required workspace views.

DeCyder Differential Analysis Software supports the following printers:

- HP LaserJet 4000N
- HP color LaserJet 5M
- HP 2500C
- HP DeskJet 895 Cxi
- HP DeskJet 950C
- HP Laser Jet 4100 DTN

Using the clipboard

Any aspect of the different views within the workspace can be copied to the clipboard for pasting into another application.

Depending on the BVA mode different views will be available for copying. Click on the part of the workspace chosen for pasting then select **Edit:Copy** (the part of the workspace selected is described in this pull down menu). The clipboard can then be pasted into other applications.

Exporting data

Data associated with the entire workspace can be exported in an XML based format (DeCyder Differential Analysis Software XML) by selecting **File:Export Workspace**. The DeCyder Differential Analysis Software XML file can then be opened in the XML Toolbox for further analysis.

Pick list data can be exported as a text file by selecting **File:Export Picking List From:Primary Spot Map or Pick Spot Map**. The text file can then be utilized by Ettan Spot Picker.

5 Spot picking

5.1 Overview

DeCyder Differential Analysis Software provides the user with the functionality to generate a spot pick list which assigns proteins for picking by Ettan Spot picking robot.

The generation of the pick list in DeCyder Differential Analysis Software involves the following processes:

- Assigning proteins of interest, which warrant identification on the analytical gels (i.e. gels analyzed during experiment).
- Spot detection on the preparative gel (gel prepared for picking - see section 3.3.1).
- Identification of reference markers on the preparative gel.
- Matching analytical and preparative gels.
- Confirming matches between analytical and preparative gels.
- Assign the proteins of interest as spots for picking.
- Confirming spots for picking.
- Exporting a pick list.

There are two principle means of generating a pick list within DeCyder Differential Analysis Software. The method adopted is dependent on the size of the experiment and the desired picking criteria. For both methods, processing the preparative gel in the DIA module is identical, and is described first.

5.2 Spot detection on the preparative gel

The preparative gel (gel picked from) is prepared and scanned as described in Ettan DIGE System User Manual.

The preparative gel is normally prepared after the analytical gels have been investigated. Therefore, it is often necessary to perform spot detection independently on the preparative gel. This is done as described in section 3.3.1, employing the single detection algorithm.

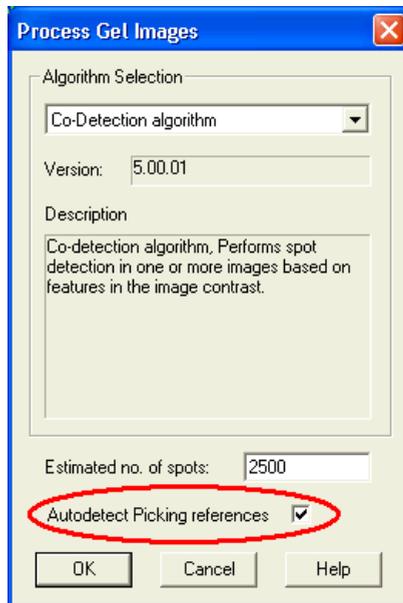
The DIA .XML file for the preparative gel must then be exported (**File:Export Spot Maps**) so that it can be loaded into the BVA module for matching against analytical gels. The .XML file is exported after the identification of reference markers, if the identification is performed in the DIA module.

5.3 Identification of reference markers

The reference markers placed on the preparative gels (seen as circles on the images) act as reference points for Ettan Spot picking robot.

Identification of the reference markers can be done in either the DIA or the BVA module, the process is identical in both cases.

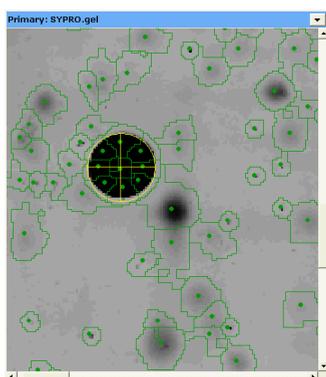
The reference markers can be detected automatically within the DIA module during the spot detection process. Once a fluorescently post-stained gel image is loaded into the DIA workspace, select **Process:Process Gel Images** to open the Process gel images window. Ensure the **Autodetect Picking references** check box is selected then click OK. The reference markers will then be detected automatically during the spot detection process.



Alternatively the reference markers can be assigned manually. Click on the **Magnify Gel View** icon to show the gel images only and select **Edit:Define Picking Reference**. Using the circular mouse cursor that appears in the Image View click on the center of the left reference

marker of the primary image. Zoom into the area of this reference marker by holding down the left mouse button to draw a rectangular area around the marker.

Clicking on the **Properties** icon and selecting the **Image View** tab can alter the size of the target. The value entered in the Default radius of picking references can be set so that the circle in the Image View is sufficiently large enough to accurately surround the reference marker. The position of the circle can be altered by selecting **Edit>Edit picking References**. Place the hand shaped cursor that now appears on the circle, then drag the circle over the center of the reference marker. Move the target so that it fits exactly around the reference. Magnifying the area of the gel around the reference marker by selecting the **Zoom in** icon may make accurate alignment easier. Select **Edit>Edit picking References** to exit this mode.



When the first reference has been selected, the same procedure is used to select the reference on the right hand side of the preparative gel image.

5.4 Identifying Proteins of Interest

Protein spots are first analyzed then filtered using various spot data parameters. The spots are filtered on the basis of user defined criteria, which result in the successfully filtered spot being assigned as a **Protein of Interest (POI)**. These proteins of interest can be the evaluated then given a **Pick** status in order to generate a pick list.

Spots can be identified as a protein of interest using either the DeCyder Differential Analysis Software DIA or BVA module.

5.4.1 Identifying proteins for picking using the DIA module

The DIA module is used to identify protein of interest for picking when performing small-scale experiments utilizing two samples on a single gel e.g. control-treated samples.

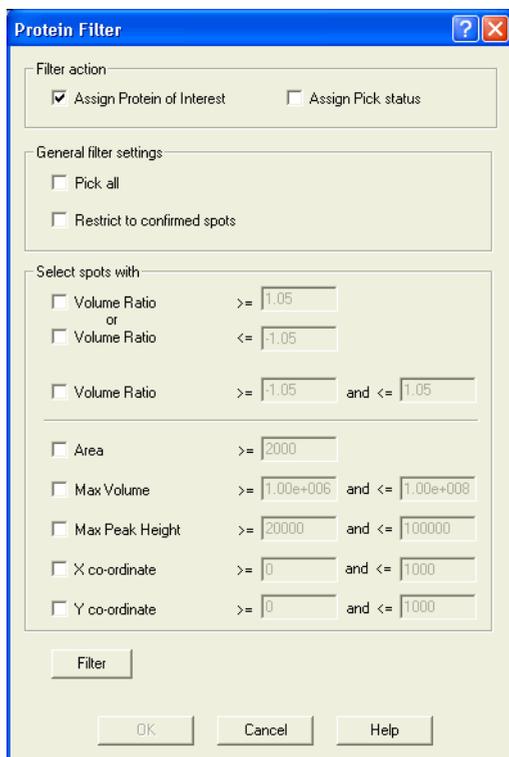
Assignment of POI is performed on all spot maps loaded into the DIA module. Proteins assigned as protein of interest can then be further evaluated in the BVA module, prior to assigning a Pick status to those proteins. The simplest means to assign POI status is to manually choose spots then select the **Protein of Interest** check box.



A screenshot of a software interface showing a set of checkboxes. The checkboxes are: 'Pick', 'PTM', 'Protein of Interest', and 'Exclude'. The 'Protein of Interest' checkbox is checked and is circled in red. To the right of these checkboxes is a 'Confirm' button.



The spots can also be selected using various criteria via the Protein Filter. Click on the **Protein Filter** icon to open the **Protein Filter** dialog box.



A screenshot of the 'Protein Filter' dialog box. The dialog has a blue title bar with a question mark and a close button. It contains several sections:

- Filter action:**
 - Assign Protein of Interest
 - Assign Pick status
- General filter settings:**
 - Pick all
 - Restrict to confirmed spots
- Select spots with:**
 - Volume Ratio >= 1.05
 - or
 - Volume Ratio <= -1.05
 - Volume Ratio >= -1.05 and <= 1.05
 - Area >= 2000
 - Max Volume >= 1.00e+006 and <= 1.00e+009
 - Max Peak Height >= 20000 and <= 100000
 - X co-ordinate >= 0 and <= 1000
 - Y co-ordinate >= 0 and <= 1000

At the bottom of the dialog are buttons for 'Filter', 'OK', 'Cancel', and 'Help'.

In the **Protein Filter** select the **Assign Protein of Interest** check box and deselect the **Assign Pick Status** check box. In this way the proteins that meet the filter criteria will be given a **Protein of Interest** status.

All spots detected can be assigned with protein of interest status by selecting the **Pick All** check box. Alternatively, spots can be filtered on specific spot properties.

Protein spots exhibiting a change in expression are commonly analyzed. Therefore selecting proteins of interest on the basis of volume ratio is useful for such spots. Spots can be selected for decreases or increases in order to select spots that have diverged notably from each other. Alternatively, spots can be selected for decreases and increases allowing selection of spot volume ratios within a finite limit.

Spots can also be selected on the basis of their physical characteristics (i.e. area, max volume, max peak height) to ensure that spot selection only occurs on spots that can potentially be successfully picked, digested then analyzed by mass spectroscopy. The location of the spots on the X and Y axes can also be filtered in order that spots near or on the edge of the gel are not present in the pick list, or to limit the eventual pick list to high or low molecular weight proteins. The various filters can be used simultaneously to select spots on multiple features.

This filtering can take place on the whole spot population or only those confirmed by selecting the **Restrict to confirmed spots** check box.

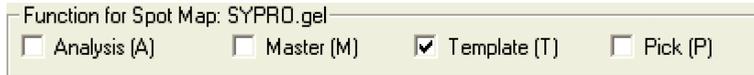
After entering the filter criteria, select the **Filter** button to ascertain the number of spots that will be picked. If the number of resultant spots is unsuitable, the stringency of the filter can be adjusted to produce an optimal number of spots for selection and subsequent picking. Click **OK** to accept the spot filter parameters. If spot confirmation was not performed prior to the picking filter, then the spots selected for picking can be visually inspected and confirmed.

The protein of interest status of spots can be removed by selecting **Process:Unassign all Protein of Interest**.

The DIA file containing the protein of interest selection information can then be exported as an XML file (**File:Export Spot Maps**) so that it can be loaded into the BVA module for matching against a preparative gel.

Both the preparative gel and analytical gel XML files are loaded into the BVA module (see section 4.3.1) and matched. Landmarking may be required to accurately match preparative gels (see section 4.6).

One of the analytical spot map must be assigned as a template spot map in the Spot Map Table mode of the BVA module by selecting the **Template** check box in the data control panel (see section 4.5).



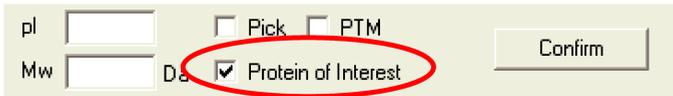
Function for Spot Map: SYPRD.gel

Analysis (A) Master (M) Template (T) Pick (P)

Spots assigned as protein of interest are denoted by the presence of the letter “I” in the Function column of the Protein Table and have the Protein of Interest check box selected.

5.4.2 Identifying proteins of interest using the BVA module

The simplest means to assign POI status is to manually choose spots then select the **Protein of Interest** check box when in Spot Map Table mode.



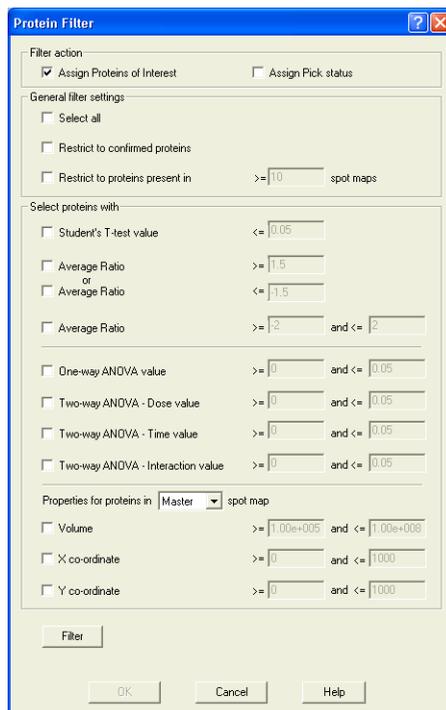
pl Pick PTM

Mw Protein of Interest

Alternatively identification of proteins for picking can be based on the statistical information generated from a BVA workspace containing several spot maps that have been subjected to statistical analysis within the BVA module.



Proteins can be selected using various criteria via the protein filter. Click on the **Protein Filter** icon to open the **Protein Filter** dialog box (see section 4.9 for further information).



All protein spots detected can be assigned by selecting the **Select All** check box. Alternatively, proteins can be filtered for picking based on the statistical analysis results produced (i.e. based on the statistical criteria described in section 4.8).

This filtering can take place on the whole spot population or a subset of the proteins or spot maps. Selecting the **Restrict to confirmed proteins** check box ensures that only proteins with a confirm status are filtered. Alternatively, selecting the **Restrict to protein present in** check box ensures that only spots that appear on the designated number of spot maps are filtered.

Filtering can be limited to spots of a specified volume by selecting the **Volume** check box then entering the volume values required.

The location of the spots on the X and Y axes can also be used for filtering in order that spots near or on the edge of the gel are not selected or to limit selecting too high or low molecular weight proteins.

The selection criteria for volume and gel region can be based on either the Master gel or the Pick gel (if present) by selecting the appropriate spot map in the pull-down menu.

After entering the filter criteria, select the **Filter** button to ascertain the number of spots that will be assigned as proteins of interest. If the number of resultant spots is unsuitable, the stringency of the filter can be adjusted to produce an optimal number of spots for picking. Click **OK** to accept the spot filter parameters.

Spots assigned as protein of interest are denoted by the presence of the letter “I” in the Function column of the Protein Table and have the Protein of Interest check box selected.

The protein of interest status of spots can be removed by selecting **Process:Unassign all Protein of Interest**.

The preparative gel XML file, which has to be first generated in the DIA module (section 5.2) can then be imported into the BVA module. Select **File:Import Spot Map(s)**, browse to locate the XML file then click **Open**. The preparative gel can then be matched (see section 4.6.2). Landmarking may be required to accurately match preparative gels (see section 4.6.1).

5.5 Assigning Spots for Picking

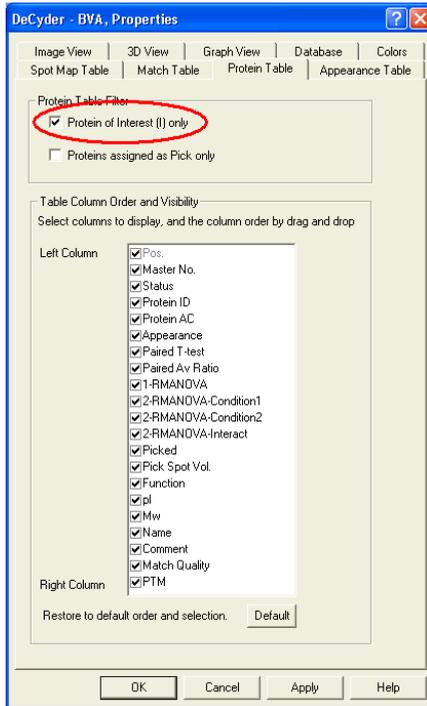
As can be seen from the previous section, a BVA workspace can be created using either the DIA module or the BVA module, to identify a set of proteins as protein of interest, in a workspace containing a preparative gel, that has been matched to the analytical spot maps.

The preparative gel should then be assigned as the Pick Spot Map (using the function check boxes in the Spot Map Table mode).



The spot map Pick status denotes that the co-ordinates for generating the pick list are based on the designated preparative gel.

It is recommended that the spots assigned as POI are visually inspected to remove those not suitable for picking and subsequent analysis. This can be performed in the Protein Table mode of the BVA module. The Protein Table must initially be organized so that POI proteins are exclusively displayed in the Protein Table. This is accomplished by selecting the **Protein of Interest (I) only** check box in the **Protein Table** tab of the **Properties** dialog window.

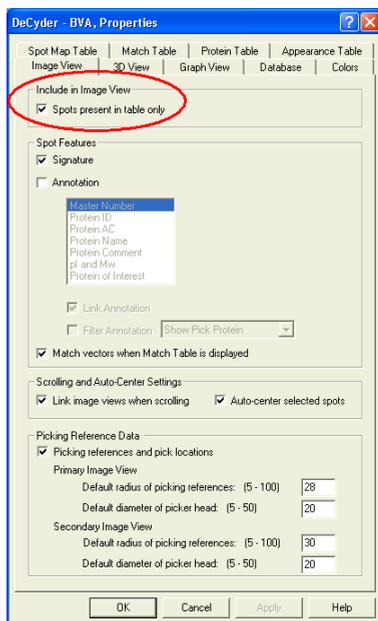


The **Function** column header in the **Protein Table** then displays the proteins with protein of interest status (possessing “I” status in the function column). The POI assigned proteins can then be reviewed sequentially to ensure that they are correctly matched and suitable for picking. Once the spot is deemed suitable for picking select the **Pick** check box in the data control panel.

Once the spot is assigned with a pick status it will have a black spot boundary in the image view and the letter “P” in the Picked column of the Protein Table. The pick location will then need to be reviewed on the spot as described below, before reviewing the next spot in the table.

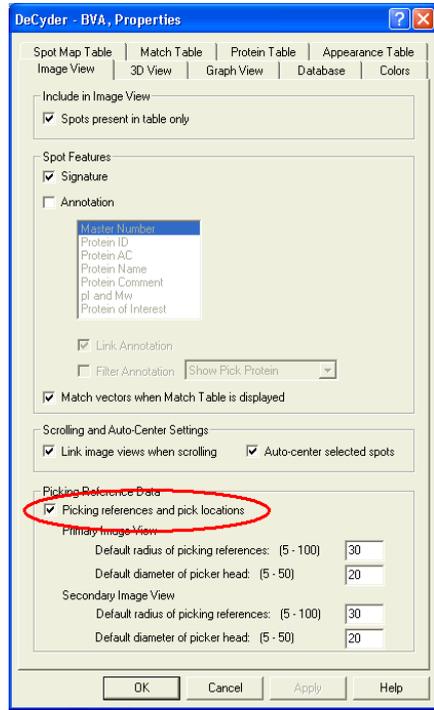
5.6 Editing Pick Locations

In the BVA module proteins assigned with pick status can be shown exclusively in the image view when in **Protein Table** mode. This is accomplished by opening the **Properties** dialog window and ensuring that **Spots present in table only** check box in the **Image View** tab is selected and the **Proteins assigned as Pick only** check box is selected in the **Protein Table** tab.

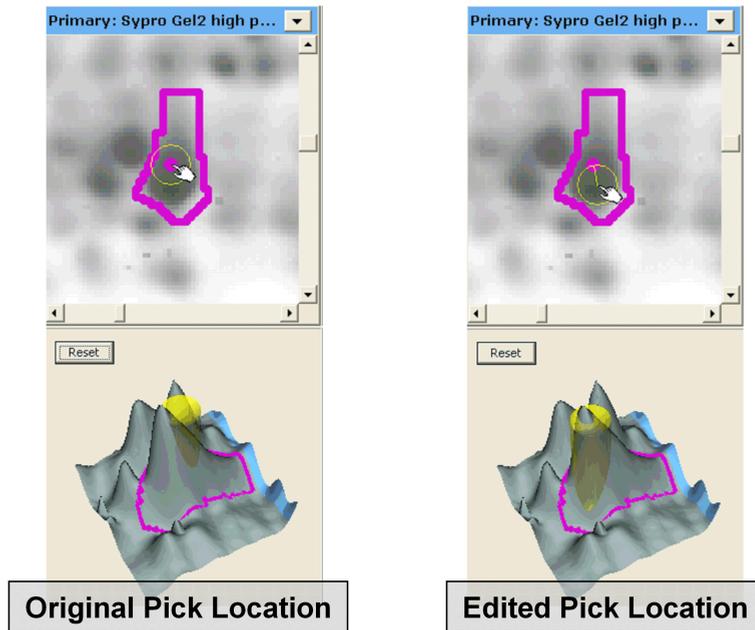


The Protein Table and the image view will then display only the proteins that have been assigned with a pick status. Using the **Zoom in** icon the individual spots selected for picking can be seen. Each of the protein spots have a black dot within the spot boundary denoting the centre of mass of the spot and hence the centre at which a picking head will pick from the gel. The centre of mass represents the optimal picking location for a vast majority of spots. However, it may be advantageous to edit the pick location when two spots are in very close proximity in order to minimize the possibility of cross contamination. The picking location can therefore be edited in these cases.

The picking location must first be displayed in both the image view and 3-D View when in Protein Table mode. The pick location is only displayed on the spot map assigned as pick (providing it has the picking references defined). To display the picking locations ensure that the **Picking references and pick locations** check box is selected in the **Image View** tab of the **Properties** dialog window.



The pick locations are displayed as a yellow (by default) transparent cylinder and a yellow circle in the 3-D View and image view, respectively when a picked spot on the pick gel is selected. The pick locations can be edited by zooming in on the selected spot in the image view, then selecting **Edit>Edit Pick Locations**. Place the hand shaped cursor that appears over the centre of the pick location that requires editing in the image view then drag the pick circle to the desired location. This can be repeated for pick locations that require editing. Select **Edit>Edit Pick Locations** to exit this mode.



To return all pick locations to the initial position of centre of mass select **Edit:Restore Default Pick Locations**.

*Note: Editing of pick locations can also be performed in the DIA module. To display only picked spots in the image view select **Picking references and pick locations** check box in the **Spot Display** tab of the DIA module **Properties** dialog window. Editing pick locations is then performed in an identical manner to the BVA module.*

5.7 Generating a Pick List

Pick lists can be generated for either Ettan Spot Picker or Ettan Spot Handling Workstation in both the BVA and DIA modules.

Select **File:Export Picking List From:Pick Spot Map** in the BVA module to generate a pick list for the preparative gel. A pick list is generated in the DIA module by selecting **File:Export Picking List**. The pick list can then be saved as either a text file or an XML file by selecting the appropriate file extension to the file name (*.txt* or *.xml*). The text and XML file are used for Ettan Spot Picker or Ettan Spot Handling Workstation, respectively.



6 Batch processor

6.1 Overview

The Batch Processor links both the DeCyder Differential Analysis Software DIA and BVA modules to perform all stages of the 2-D DIGE analysis process. Once the Batch Processor has been set up, the gels are processed sequentially without user intervention. The Batch Processor is, therefore, a means of automation and does not introduce further processing and analysis to the spot map data. All the concepts for processing the gel images have been discussed in the DIA and BVA sections of the manual and are not further elaborated upon in this section. References to the appropriate sections are made where necessary.

6.2 Batch list creation



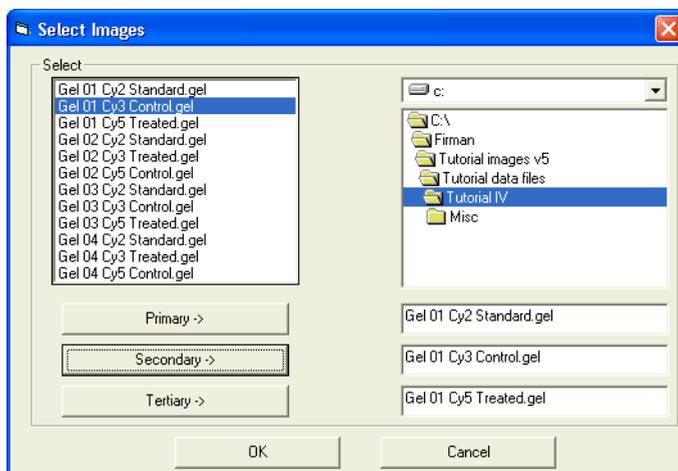
Double click the **DeCyder Differential Analysis Software Batch Processor** icon on the desktop to open the Batch Processor.

The creation of the batch list can be broadly sub-divided into two activities:

- DIA batch list set up. Loading the spot map pairs for co-detection in the DIA module.
- BVA batch list set up. Stipulating spot map attributes, setting up the statistical analysis and subsequent pick list in the BVA module.

6.2.1 Setting up the DIA Batch List

When setting up a DIA batch list each gel is processed sequentially. To create a batch list select **File:New Batch**.



In the dialog box that appears browse using the right panel to locate the folder containing the gel images to be processed. The image files contained in the selected folder appear in the left panel. Click to highlight the image designated to be the primary image in the first gel to be processed (conventionally the primary image is the standard image in an experiment employing an internal standard). Click the **Primary** button to select this image as the primary image.

Repeat the process for the secondary and tertiary image if present using the appropriate selection buttons, then click **OK**.

The **Gel Image Information and Result Files** dialog box automatically appears.

The CyDye DIGE Fluor dyes for the images can be selected in the appropriate pull down menus. If the image file names contain the text “Cy2”, “Cy3” or “Cy5” these fields are automatically assigned.

The DIGE labelling chemistry can be selected (**Min** is selected as default).

The spot co-detection process in the Batch Processor generates both a .DIA and an .XML file for each gel processed. The file name of these files can be stipulated in the text boxes in the Results File area of the dialog box. Both file names are automatically given identical names and are distinguishable by their file extension. It is recommended that the number “1” is appended to the end of the file name of the first gel.

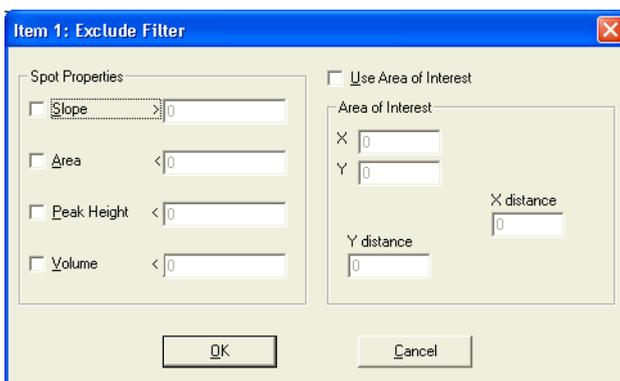
Subsequent gels will then be named automatically by incrementing the number at the end of the file name. If the number “1” is not appended to the end of the file name then the user is prompted to enter a file name for each gel. Click **OK** when all the data is correctly entered.

The **Estimated no of spots and BVA Processing** dialog box automatically appears. The estimated number of spots can then be entered; this estimation is used for all subsequent spot map pairs. See section 3.3.1 for further details.

If the gel is a pick gel, containing picking references which require autodetection, select the check box for **Autodetect pick reference markers**. If BVA inter-gel matching is also required, select the check box **Include in BVA batch list**. Click **OK**.

Note: *If the spot map of the preparative gel image contains the text “pick” or “preparative” in the image file name, the software will automatically assign the spot map with a Pick status and will autodetect the references.*

The **Exclude Filter** dialog box automatically appears.



The **Exclude Filter** dialog box can either be left empty (in which case click **OK**), or a set of known filter parameters can be entered followed by clicking **OK**. The filtering values are best pre-determined in the DIA using one of spot map pairs to be batch processed. See section 3.5.1 for further details.

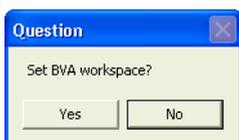
The Exclude Filter values stipulated are used on all subsequent gels.

The next dialog box allows a second gel to be entered. All subsequent sets of images are loaded in an identical manner to the first. When all the images have been loaded, click **Cancel**.

The batch information is then entered into the DIA batch list (example below).

DIA Batch List							BVA Batch List											
No.	Status	Primary	Label	Secondary	Label	Tertiary	Label	Type	DIA File	XML File	Estimated	Auto	Found	Filter	Area of interest	Pr		
1	Pending	Gel01	Standard	Cy2	Gel01	Cy3	gel	Cy3	Gel01	Cy5.gel	Cy5	Min	Gel 1.dia	Gel 1.xml	2500		S:0:A:0:P:0:V:0	Whole Image
2	Pending	Gel02	Standard	Cy2	Gel02	Cy3	gel	Cy3	Gel02	Cy5.gel	Cy5	Min	Gel 2.dia	Gel 2.xml	2500		S:0:A:0:P:0:V:0	Whole Image
3	Pending	Gel03	Standard	Cy2	Gel03	Cy3	gel	Cy3	Gel03	Cy5.gel	Cy5	Min	Gel 3.dia	Gel 3.xml	2500		S:0:A:0:P:0:V:0	Whole Image
4	Pending	Gel04	Standard	Cy2	Gel04	Cy3	gel	Cy3	Gel04	Cy5.gel	Cy5	Min	Gel 4.dia	Gel 4.xml	2500		S:0:A:0:P:0:V:0	Whole Image
5	Pending	Gel05	Standard	Cy2	Gel05	Cy3	gel	Cy3	Gel05	Cy5.gel	Cy5	Min	Gel 5.dia	Gel 5.xml	2500		S:0:A:0:P:0:V:0	Whole Image
6	Pending	Gel06	Standard	Cy2	Gel06	Cy3	gel	Cy3	Gel06	Cy5.gel	Cy5	Min	Gel 6.dia	Gel 6.xml	2500		S:0:A:0:P:0:V:0	Whole Image
7	Pending	Gel07	Standard	Cy2	Gel07	Cy3	gel	Cy3	Gel07	Cy5.gel	Cy5	Min	Gel 7.dia	Gel 7.xml	2500		S:0:A:0:P:0:V:0	Whole Image
8	Pending	Gel08	Standard	Cy2	Gel08	Cy3	gel	Cy3	Gel08	Cy5.gel	Cy5	Min	Gel 8.dia	Gel 8.xml	2500		S:0:A:0:P:0:V:0	Whole Image
9	Pending	Gel09	Standard	Cy2	Gel09	Cy3	gel	Cy3	Gel09	Cy5.gel	Cy5	Min	Gel 9.dia	Gel 9.xml	2500		S:0:A:0:P:0:V:0	Whole Image
10	Pending	Gel10	Standard	Cy2	Gel10	Cy3	gel	Cy3	Gel10	Cy5.gel	Cy5	Min	Gel 10.dia	Gel 10.xml	2500		S:0:A:0:P:0:V:0	Whole Image
11	Pending	Gel11	Standard	Cy2	Gel11	Cy3	gel	Cy3	Gel11	Cy5.gel	Cy5	Min	Gel 11.dia	Gel 11.xml	2500		S:0:A:0:P:0:V:0	Whole Image
12	Pending	Gel12	Standard	Cy2	Gel12	Cy3	gel	Cy3	Gel12	Cy5.gel	Cy5	Min	Gel 12.dia	Gel 12.xml	2500		S:0:A:0:P:0:V:0	Whole Image
13																		
14																		
15																		

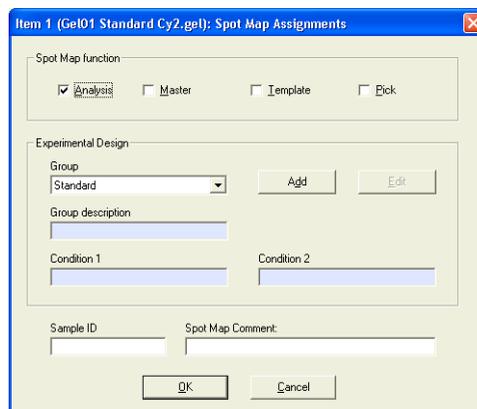
If inclusion of a BVA batch list was selected the user is prompted to set the BVA workspace.



Click **Yes** to set the BVA workspace in order to assign spot map attributes and set up statistical analysis and generate a pick list if required.

6.2.2 Setting up the BVA Batch List

The Spot Map Assignment dialog box appears when setting up the BVA workspace.



As with the Spot Map Table mode in the BVA module, attributes (i.e. spot map function, group, conditions, sample ID and spot map comments) can be assigned to individual spot maps. Attribute assignment is performed in the above dialog box.

Attributes are assigned sequentially to each spot map by selecting the appropriate spot map function or entering the desired experimental design criteria. Click **OK** once all the attributes are assigned to the first spot map. The Spot Map Assignment dialog box for the second spot map will automatically appear. The spot map attributes for all spot maps can therefore be assigned in this iterative process.

All attributes are assigned as in the Spot Map Table mode (section 4.5.1 and 4.5.2). In the Batch Processor, groups are added by selecting the **Add** button in the Spot Map Assignment dialog box.

Note: *As with the BVA module, if no master image is stipulated the internal standard spot map, (if present) with greatest number of detected spots is automatically assigned as the master image when running the batch list.*

Once all spot maps attributes have been defined the **Protein Statistics** dialog box automatically appears. This allows the user to perform statistical analysis on the experiment. The appropriate statistical analysis can be set up as described in section 4.8. Click **OK** to confirm

the statistical analysis entered. Alternatively, if no statistical analysis is required click **Cancel**.

The **Protein Filter** dialog box then appears automatically. The filter allows either the highlighting of interesting proteins for further investigation or the direct identification of proteins for picking and subsequent generation of a pick list, if a pick gel is present in the batch list (see Chapter 5).

Note: It is recommended that the Protein Filter is used to highlight interesting proteins for further investigation. These proteins can subsequently be confirmed by the user in the BVA module prior to generating a pick list (see section 5.4.2).

6.3 Editing the batch list

6.3.1 DIA batch list

To edit the DIA batch list select **View:DIA Batch List** to ensure the DIA batch list is displayed.

Adding

To add further spot map pairs to a current DIA batch list, select the next empty row in the batch list and select **File:Add DIA batch item**. Extra gels are added as described in section 6.3.2.

Removing

To remove an item, select the item from the DIA batch list, select **File:Remove DIA batch item** then confirm that the item is to be deleted. The item will then be automatically removed.

Editing

If an incorrect image is selected while loading the DIA batch list, continue loading the rest of the images then amend the desired image at the end.

To amend the batch list, select the desired item from the DIA batch list, and then select **Edit>Edit Item**.

6.3.2 BVA batch list

To edit the BVA batch list select **View:BVA Batch List** to ensure the BVA batch list is displayed.

Adding

To add a spot map(s) to the BVA batch list, select **File:Add Spot Map(s)**. Browse to locate and select the XML files corresponding to the spot map(s) processed in a previous DIA analysis. Click **Open**. The spot map(s) will automatically be added to the batch list.

Removing

To remove an item, select the item from the BVA batch list, select **File:Remove BVA batch item** then confirm that the item is to be deleted.

Editing

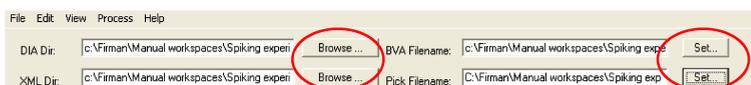
To amend the spot map attributes of an item in the BVA batch list, select the desired item from the BVA batch list, then select **Edit:Edit Item**.

6.4 Saving Data

The batch list, containing all the information entered into the Batch Processor, can be saved by selecting **File:Save As** (files are saved as text files). The batch list text files can then be re-opened later by selecting **File:Open**.

The folder that the DIA and XML files generated when processing the DIA batch list can be selected by clicking the **Browse...** buttons adjacent to the respective directory text box.

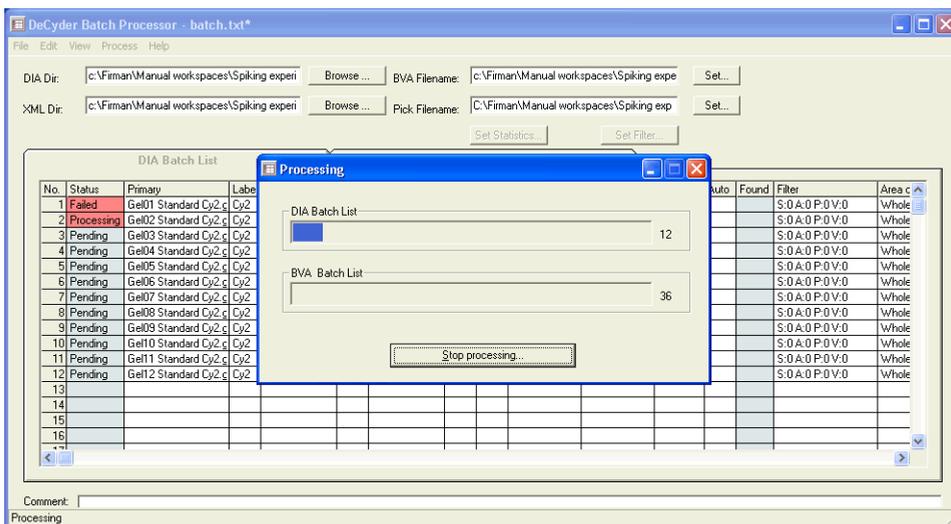
Similarly, if the BVA batch list is also performed, the names and destinations for saving the subsequent BVA file and pick list (if requested) can be stipulated by selecting the **Set...** buttons adjacent to the respective file name text box.



6.5 Running the batch processor

To run the Batch Processor select **Process:Run batch**. If the default BVA filename has not been changed the user will be given the option to change this.

The **Run batch** dialog box, which now appears, gives the user the option to begin processing immediately or introduce a delay period. The batch processor must be left open during the delay period. To begin the countdown to processing or to start processing, click **OK**.



7 XML Toolbox

7.1 Overview

Large amounts of data are generated when a workspace is created and analyzed in DeCyder Differential Analysis Software. It is useful to be able to save this data in a format that can be efficiently stored and queried. Data is exported from the different DeCyder Differential Analysis Software modules using a common XML format called DeCyder Differential Analysis Software XML. XML is a structured universal tagged language with tags that can be custom designed, making it easy to access data from DeCyder Differential Analysis Software analysis workspaces for other applications. The XML format is partly used to transfer data between the different modules of the DeCyder Differential Analysis Software but it is also used to make data available for post processing by other software packages such as database building packages.

The DeCyder Differential Analysis Software XML Toolbox is a toolbox shell housing different tools for extraction of the DeCyder Differential Analysis Software data from the different XML files produced within DeCyder Differential Analysis Software. This enables users to create tools to convert their data into text files or html files (potentially other data formats can be supported for conversion). Two basic tools are supplied to create Tabbed text files and Web tables.

7.2 Opening the XML Toolbox module



Click on the **Toolbox** icon to open the XML Toolbox.

The XML Toolbox can extract data in two formats: Tabbed text files and Web tables.

Tabbed text files

The Tab Separated Tables tool is used to export DeCyder Differential Analysis Software XML data in a tabbed text format for further data processing in other software that can import this data (e.g. Microsoft Excel or SpotFire™). The data in the resulting output table is configured such that the top row consists of an identifying line of text, containing the name of the source XML file and the type of data extracted from the file. The second row in the output table is a tab-separated list of column headers, and the following rows consist of the actual tab separated raw data.



To utilize the Tab Separated Tables tool click on the **Tab Separated Tables** button.

Web tables

The Web Tables tool is an example of how the DeCyder Differential Analysis Software XML format in combination with XSL and other web technology can be used to customize reports in HTML format. The Web Tables tool can thus be used as a template for user designed report tools.



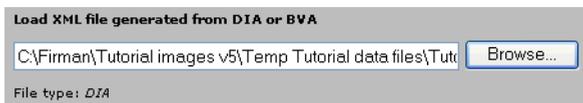
To utilize the Web Tables tool click on the **Web Tables** button.

The interface for both tools is similar.

7.3 Extracting data

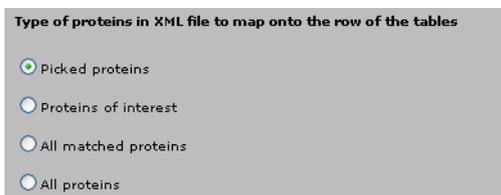
Loading the XML files

XML files can be loaded at the top of the tool page, by either browsing for a file then clicking **Open** or by typing the file pathway directly into the list box then pressing enter on the keyboard. The tool verifies whether the selected file is a valid XML file exported from DeCyder Differential Analysis Software-DIA or DeCyder Differential Analysis Software-BVA. The user is alerted if the selected file is not valid.



Type of proteins in the output table

Once the DeCyder Differential Analysis Software XML file has been opened, select the types of proteins to be included in the output table. For XML files exported from BVA, four selections are available, **Picked proteins**, **Proteins of interest**, **All matched proteins**, and **All proteins**. For XML files exported from DIA, only the **Picked proteins** and **All proteins** selections are available.



Category of data in table columns

The data is categorized into four groups:

- 1 **Experiment data.** The table contains general experimental set up parameters valid for the current DIA or BVA experiment.
- 2 **Image data.** The data in this table contains the different parameters that are unique to each spot in the individual gel images contained in the selected DIA or BVA experiment.
- 3 **Gel data.** The data displayed in this table is unique for each detected spot pair within each set of gel images that originate from the same gel.
- 4 **Protein data.** All protein data that is unique to a set of matched spots across the different gels in the current BVA experiment is displayed in this table.



The interfaces diverge when selecting categories of data:

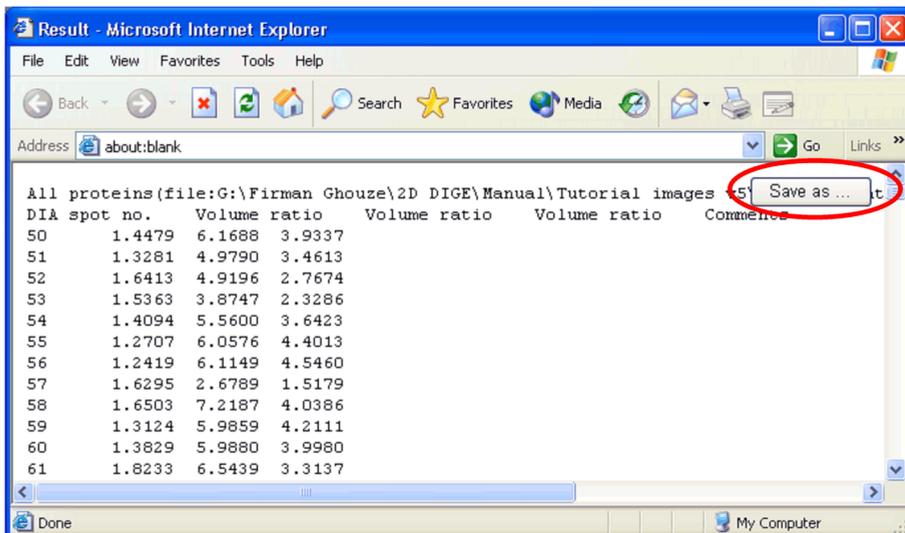
- **Web tables.** Each of the four lists of data items represents a type of result table. Thus if at least one item in each list is selected, the resulting output contains four different result tables. Multiple selections are possible in all lists. The content of the four data lists also changes, depending on whether an XML file is exported from the DIA or BVA module.
- **Tabbed text files.** Select a category of DeCyder Differential Analysis Software data to include in the result output from the pull down menu. Depending on the selection of category and type of XML file selected, a number of items are displayed in the list box below. For the Experimental data and Protein data categories, multiple selections are possible in the list of data items, but for the Image data and Gel data categories, only a single selection is possible.

Extracting Data

Once the data output has been specified, the data can be extracted by clicking the **Extract data** button.

The data extraction process can take from a few seconds to several minutes depending on the size of the selected file, the amount of data extracted and the performance of the computer. Whilst the data is being

extracted, a progress window displays the size of the XML file. Once extracted, the Result Table is displayed in a separate browser window. A floating **Save As** button is present in the upper right hand corner of the output window. This button can be used to save the extracted output directly to a designated file for use by other applications.



7.4 Tag definitions

7.4.1 BVA parameters

Experimental data

Condition 1. Numerical condition 1 assigned to the spot map group.

Condition 2. Numerical condition 2 assigned to the spot map group.

DIA experiment. File name of the corresponding DIA workspace.

Dye chemistry. Indicates whether minimal or saturation labelling was used.

Dye label. Indicates the fluor used to label the protein.

Group Description

Group. Name of the experimental group that the spot map has been assigned to.

Image name. Gel image filename.

No. of included spots. Number of spots included in the data set.

No. of matched spots. Number of spots matched to the master image.

Sample ID. Numerical identifier assigned to spot maps for paired statistical analysis.

Spot map comment. User defined comment on selected spot map.

Spot map function. Spot map function in DeCyder Differential Analysis Software BVA. M, A, T and P represent master, analysis, template and pick respectively.

Image data

Volume. Sum of the detected pixel values above background within a spot boundary.

Peak height. Pixel value at the X,Y position of the spot. This represents the highest pixel value within the spot boundary.

Normal volume. Normalized volume obtained by dividing the calculated volume for each included spot by the center of volume of the corresponding spot map.

Gel data

Volume ratio. Ratio of the normalized volumes of a pair of spots from a spot map pair. A value of 2.0 represents a two-fold increase while -2.0 represents a two-fold decrease, whilst a value of 1.00 represents an unchanged spot.

Spot comment. User defined comment on selected spot.

Picked status. Indicates whether the protein has been assigned for picking.

X-pos. Position of a spot peak along the horizontal axis of the spot map.

Y-pos. Position of a spot peak along the vertical axis of the spot map.

Protein ID. Unique protein identifier that can be used to search databases providing it is in the correct format for the specific database.

Area. Area as expressed in pixels within a spot boundary.

Match confidence. Indicates the type of match achieved i.e. Unmatched, Auto Level1, Auto Level2, Manual or Added (to Master Spot Map).

Match status. Indicates whether match has been confirmed.

Spot number. DIA spot number as given by the spot detection algorithm.

Protein data

Picked status. Indicates whether the protein has been assigned for picking.

Protein ID. Unique protein identifier that can be used to search databases providing it is in the correct format for the specific database.

Protein AC. The protein accession number as designated by the selected database.

Protein Name. Protein name given by selected database.

Protein comment. Note typed in the protein comment box

Protein function. Proteins assigned Protein of Interest are marked with an “I” in this column.

pI. Isoelectric point of the protein.

pI Status. Indicates whether displayed isoelectric point was entered for reference purpose of the algorithm, calculated from the selected algorithm or manually edited (designated Template, Calculated and Manual respectively).

Mw. Molecular weight of protein.

Mw Status. Indicates whether displayed molecular weight was entered for reference purpose of the algorithm, calculated from the selected algorithm or manually edited (designated Template, Calculated and Manual respectively).

T-test value. Student’s T-test p value.

Average ratio. Average ratio between the groups selected for protein analysis.

ANOVA1. One-way ANOVA (ANalysis Of VARIation) p value.

ANOVA2 condition1. Two-way ANOVA for condition 1 p value for condition 1.

ANOVA2 condition2. Two-way ANOVA for condition 2 p value for condition 2.

ANOVA2 interaction. Two-way ANOVA interaction p value for the interaction between condition 1 and condition 2.

Paired T-test. Paired Student's T-test p value.

Paired average ratio. Paired average ratio between the groups selected for protein analysis.

RM ANOVA1 (Repeated Measures ANOVA). Paired One-Way ANOVA p value.

RM ANOVA2-condition1. Paired Two-Way ANOVA p value for condition 1.

RM ANOVA2-condition2. Paired Two-Way ANOVA p value for condition 2.

RM ANOVA2-interaction. Paired Two-Way ANOVA p value for the interaction between conditions 1 and condition 2.

7.4.2 DIA parameters

Experimental data

Primary Image. Refers to the left hand gel image present in the image view.

Secondary Image. Refers to the right hand gel image present in the image view.

DIA experiment. Gel name/designation.

No. of included spots. Number of spots designated as included for subsequent analysis.

Dye chemistry. Indicates whether minimal or saturation labelling was used.

Estimated no. of spots. User defined approximation of the number of spots present in spot map image entered prior to spot detection.

No. of similar spots. Number of spots within the threshold mode set.

No. of increased spots. Number of spots showing an increase in abundance, above the threshold mode set from the secondary image compared to the primary image.

No. of decreased spots. Number of spots showing a decrease in abundance, above the threshold mode set from the secondary image compared to the primary image.

Threshold mode. Value above or below which spots are classed as being differentially expressed.

2SD value. 2 standard deviation of the spot ratio distribution, 95% of the spots lie within this ratio for normally distributed data.

Image data

Volume. Sum of the detected pixel values above background within a spot boundary.

Peak height. Pixel value at the X,Y position of the spot. This represents the highest pixel value within the spot boundary.

Spot slope. Gradient associated with the 3 dimensional attributes of a spot map pair.

Gel data

Volume ratio. Ratio of the normalized volumes of a pair of spots from a spot map pair. A value of 2.0 represents a two-fold increase while -2.0 represents a two-fold decrease, whilst a value of 1.00 represents an unchanged spot.

Spot comment. User defined comment on selected spot.

Picked. Indicates whether the protein has been assigned for picking.

Coordinates. Values giving the position of the center of mass of a spot on the spot map.

Protein ID. Spot identifier, which after confirmation assignment will appear with the relevant spot in the table view.

Area. Area as expressed in pixels within a spot boundary.

Peak height ratio. Ratio of the normalized peak heights of a pair of spots from a spot map pair. A value of 2.0 represents a two-fold increase while -2.0 represents a two-fold decrease, whilst a value of 1.00 represents an unchanged spot.

Spot radius. Refers to the radius that the spot would have if it was circular and had the same area.

Excluded. Indicates whether spot has been excluded.

Confirmed. Indicates whether a spot has been manually confirmed by the user.

Confirm Status. Indicates whether the user has confirmed protein as real.

Auto Confirm Assignment. Refers to the spot status after application of the exclude filter. Spots are either designated as excluded or included.

8 LWS Integration

8.1 Overview

The XML Toolbox has been designed to enable the smooth transfer of files containing pick lists from DeCyder Differential Analysis Software to Ettan LWS software (product number 18-1164-06), and the transfer of protein identification results from Ettan LWS software, after MALDI mass spectrometry, back into DeCyder Differential Analysis Software.

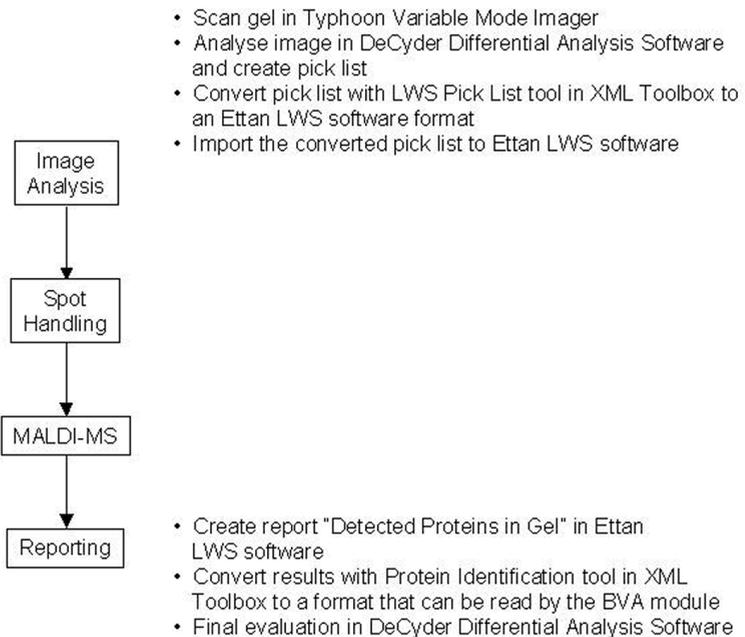


Fig 8-1. Ettan LWS - DeCyder XML Tools in Ettan DIGE system workflow.

To enable full tracking of the CyDye DIGE Fluor dye treated samples some extra information has to be added in the Sample Definition component of the Ettan LWS software.

The XML toolbox contains two tools to perform data conversion:

- LWS Pick List converts pick lists from a DeCyder Differential Analysis Software workspace to a format that Ettan LWS can import.
- Protein Identification converts results from Ettan LWS in *.xml* format to a format that the BVA module can import.

This section of the manual aims to:

- describe the deviation from the normal 2D-MS workflow as described in Ettan 2D-MS and LWS Software Laboratory Guide for samples generated within Ettan DIGE system
- describe how the LWS Pick List tool is used to convert exported workspace data from the DeCyder Differential Analysis Software evaluation into a pick list format Ettan LWS 1.0 can read
- describe how the Protein Identification tool is used to import the protein identification results from Ettan LWS data back into DeCyder BVA

8.2 Ettan LWS with samples generated within Ettan DIGE system

There are two alternative ways to enter a sample generated within Ettan DIGE system, into Ettan LWS workflow:

- as a protein tube or;
- as a gel with a pick list and an image file.

The two alternatives offer different levels of sample tracking. For a protein tube, tracking is available all the way through Ettan LWS workflow. For a gel, the tracking starts at spot handling. The workflows are also different, see Fig. 8-2.

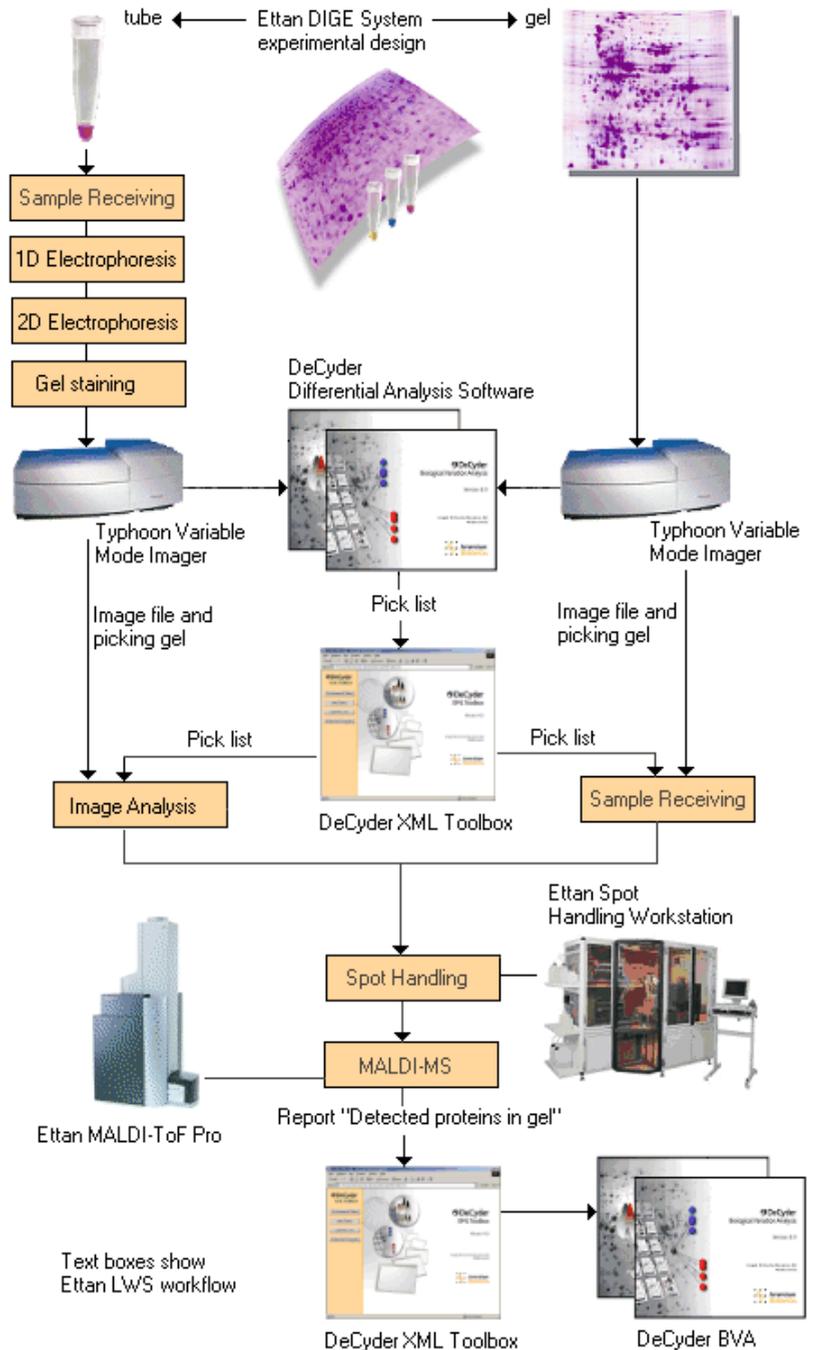


Fig 8-2. Overview of the workflow for a sample generated within Ettan DIGE system entered as a tube or as a gel.

8.3 Enter a protein tube generated within Ettan DIGE system, into Ettan LWS

The workflow for the samples should be set up the normal way in Ettan LWS. The only differences are within sample attributes template and sample definition.

A user defined sample attribute template must be set up to match the experimental design within Ettan DIGE system. By creating a **DIGE** sample template in Sample Attribute Templates it will be possible to later search for the information and to produce Ad hoc reports. For information on Ad hoc reports, please consult Ettan 2D-MS and LWS software Appendices.

- 1 In the **Management Workbench:Sample Template Editor**, define a sample type for a sample generated within Ettan DIGE system, by adding user defined attributes. An example is shown in Fig. 8-3.

Application:

Sample Types for Ettan 2D-MS:

- All Sample Types
- Protein**

Templates for Protein:

- Core Template for Protein
- Bluebird
- Crane
- DIGE**

Add Copy Edit Remove

Attributes for DIGE:

Name	Type	Default	Is Required?
Cy2	Text		<input type="checkbox"/>
Cy3	Text		<input type="checkbox"/>
Cy5	Text		<input type="checkbox"/>
Protein labelling	List	50 µg:400 pmol mi... ▾	<input type="checkbox"/>
Experiment	List	E.coli-benzoic-study1 ▾	<input type="checkbox"/>

OK Help

Fig 8-3. Example of attributes for a sample type generated within Ettan DIGE system.

The samples need to be defined in Sierra™ web.

- 2 In **Sierra web** select project name and **Define Samples**.
- 3 Scan or type sample ID, choose **Sample Type (Protein)**, **Sample**

Template (the user defined template created in step 1) and **Container (tube)** and enter values of the sample attributes. An example is shown in Fig. 8-4.

Note: The extra information (Sample attribute values) helps the user to keep track of which samples are in each gel.

Sample Definition | Sample Definition Import

Sample Definition - Ettan 2D-MS

ID:

Sample Type:

Sample Template:

Container:

Project Selection	
Add to Project	Project Name
<input checked="" type="checkbox"/>	E.Coli protein mapping

Comments:

Sample Attributes	
Name	Value
Concentration (mg/ml)	0
Name	tn413
Volume (ul)	0
Cy2	S1 - pool
Cy3	S1 - control 1
Cy5	S1 - treated 2
Protein labelling	50 µg:400 pmol ...
Experiment	E.coli-benzoic-stu...

Finish Activity | Clear Window | Close Window

Fig 8-4. Example of Sample Definition and entered values for the attributes when entering a sample generated within Ettan DIGE system as a protein tube.

Methods

Normally the standard methods included in Ettan LWS software and in Ettan MALDI-ToF Pro can be used. If non-standard methods are to be used these methods have to be added:

- In **Management Workbench:System:Method editor**, add staining method.
- In **Sierra LWS** select **Spot Handling** from the Workflow Graph. In **Tools:Spot Handling Method Management:Method Overview**, add Spot Handling methods.
- In **Ettan MALDI Control:Methods**, add MALDI methods.

8.4 Enter a gel, prepared for spot picking within Ettan DIGE system, into Ettan LWS

A preparative gel or an analytical gel with CyDye DIGE Fluor dye labelled proteins can be used as a picking gel. For details of the methods for the different routes to picking, refer to chapter 5.

The workflow for the gel should be set up the normal way in Ettan LWS. The only differences are within sample attributes template and sample definition.

- 1 Set up a sample attribute template as in 8.3, step 1.

The samples need to be defined in Sierra web.

- 2 In **Sierra web** select project name and **Define Samples**.
- 3 Scan or type sample **ID**, choose **Sample Type (Protein)**, **Sample Template** (the user defined template created in step 1) and **Container (Gel 2D)** and enter values of the sample attributes. An example is shown in Fig. 8-5. This extra information helps the user to keep track of which samples are in each gel.
- 4 Select **Glass backing**, fill in **backing thickness 3.2**.
- 5 Browse for the **image file** and the converted **evaluation file** (converted according to section 8.5) to be used.

Note: Image files to be included in the sample definition must have a file name of less than 45 characters, otherwise a work request for that sample cannot be generated and an error message will be shown.

Note: If samples generated within Ettan DIGE system are entered as a gel it will only be possible to track the samples in the spot handling and MALDI mass spectrometry steps.

Sample Definition | Sample Definition Import

Sample Definition - Ettan 2D-MS

ID:

Sample Type:

Sample Template:

Container:

Name	Value
Concentration (mg/ml)	0
Name	tn425
Volume (ul)	0
Cy2	unlabelled
Cy3	unlabelled
Cy5	unlabelled
Protein labelling	unlabelled
Experiment	E.coli-benzoic-stu...

Project Selection

Add to Project	Project Name
<input checked="" type="checkbox"/>	E.Coli protein mapping

Comments

Picking gel stained with SYPRO Ruby
E.coli-benzoic-study

Backing material:

Backing thickness (mm):

Image file:

Evaluation file:

Fig 8-5. Example of Sample Definition and entered values for the attributes when entering sample generated within Ettan DIGE system as a gel

Methods

Normally the standard methods included in Ettan LWS software and in Ettan MALDI-ToF Pro can be used. If non-standard methods are to be used these methods have to be added:

- In **Management Workbench:System:Method editor**, add staining method.
- In **Scierra LWS** select **Spot Handling** from the Workflow Graph. In **Tools:Spot Handling Method Management:Method Overview**, add Spot Handling methods.
- In **Ettan MALDI Control:Methods**, add MALDI methods.

8.5 XML toolbox - LWS Pick List

The LWS Pick List tool is used to transform DeCyder Differential Analysis Software data into an Ettan LWS 1.0 format. The tool takes a DeCyder *.xml* file exported from BVA or DIA, extracts the spots in the file that are selected for picking in DeCyder Differential Analysis Software and creates a new *.xml* file for import to Ettan LWS software.

8.5.1 Preparations before XML conversion

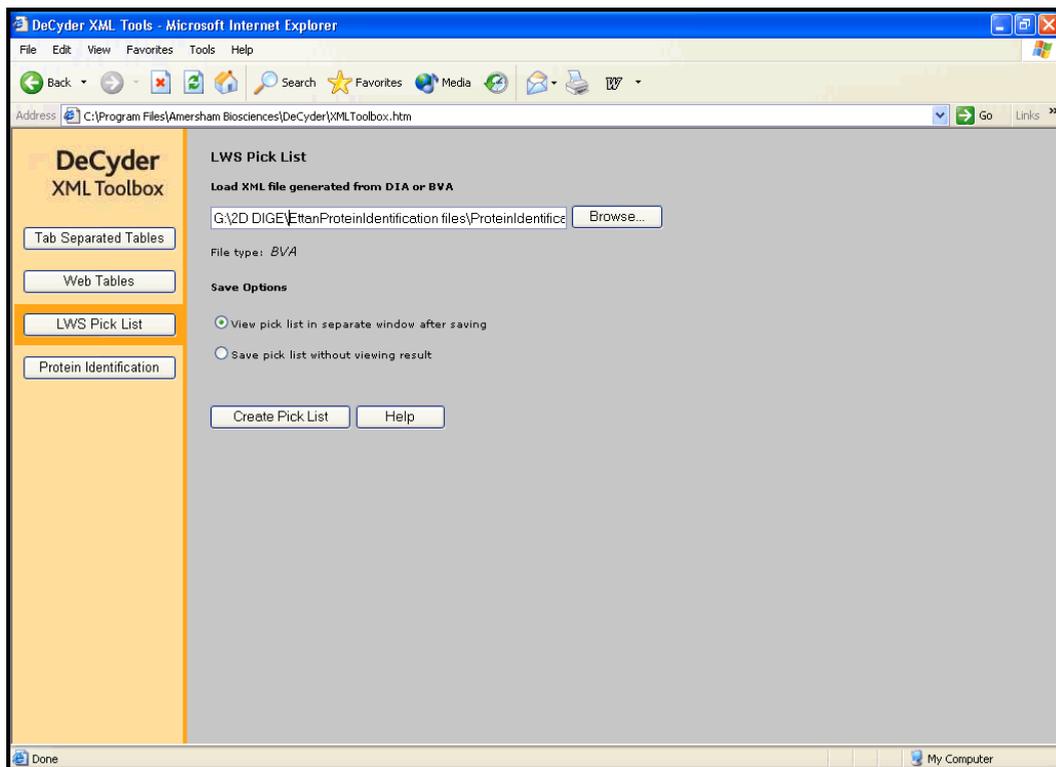
- 1 If possible, base pick lists on information in a BVA workspace rather than a DIA workspace. It is a lot easier to find relevant protein spots to pick in the BVA module as it includes an overview of all experimental conditions.

In BVA workspace: use a workspace including spot maps from a picking gel that contains real pick references. A spot map from the gel to be used for picking needs to be assigned as a pick spot map before exporting the XML workspace file.

In DIA workspace: use a workspace based on gel images from a picking gel that contains two real pick reference markers. If a DIA workspace is used to generate a pick list for LWS it will not be possible to import protein data for that pick list into the BVA module using the protein identification tool.

- 2 Assign the pick references in DeCyder Differential Analysis Software.
- 3 Export the workspace from BVA by selecting **File:Export Workspace** or in DIA by selecting **File:Export Spot Map**. The file must be saved to a location accessible by the XML Toolbox.

8.5.2 Using the LWS Pick List tool



- 1 Either browse for a file or type the path of the file into the **Load XML file generated from DIA or BVA** field at the top of the tool page.
- 2 Select either **View pick list in separate window before saving** or **Save pick list without viewing result**.
- 3 Click **Create Pick List** to start the conversion procedure. At the end of the conversion procedure a **Save As** dialog is displayed. Enter the name of the file where the pick list is saved. Type the extension *.xml* to the file name. Save the file in a location accessible by Sierra LWS.

8.5.3 After XML conversion

In Ettan LWS, the converted pick list file (from section 8.5.2) will now be used as the evaluation file together with the gel image file in sample definition for the 2D-gel (see section 8.4).

Note: If you have entered the sample generated within Ettan DIGE system as a tube (see section 8.3), follow the normal workflow in Ettan LWS using the converted pick list file as the evaluation file in Image Analysis on the Image Results tab.

After entering the pick lists in Ettan LWS do not modify the original DeCyder BVA workspace. If changes are made to the BVA workspace between exporting the pick list data and importing the annotated Ettan protein identification result, it cannot be guaranteed that the imported protein IDs are still consistent with the information in the BVA workspace. If changes are required, it is advisable to make these changes on a copy of the original.

8.6 XML Toolbox - Protein Identification

The Protein Identification tool is used to transform Ettan LWS data back into DeCyder Differential Analysis Software format. It is then possible to compare the original gel images in DeCyder BVA with the protein identification results from the MALDI mass spectrometry analysis in Ettan LWS.

The report “Detected proteins in gel” from Sierra LWS (saved in *.xml* format) is used as the input file for the Protein Identification tool. This file is converted to an *.xml* file that can be imported into the BVA module.

Data that can be imported from Ettan LWS to DeCyder BVA are:

- Protein ID. If multiple protein candidates are included in the LWS report for the different picked spots, the protein Id's of the different candidates are all appended in a single string, separated by semi-colons (;).
- Protein Name. If multiple protein candidates are included in the LWS report for the different picked spots, the protein names of the different candidates are all appended in a single string, separated by semi-colons (;).
- Molecular Weight (MW). The MW value of the first protein candidate in the list is added to Ettan Protein Identification data.

- Isoelectric Point (PI). The pI value of the first protein candidate in the list is added to Ettan Protein Identification data.

8.6.1 Preparations in Ettan LWS before XML conversion

After MALDI mass spectrometry is completed the “Detected proteins in gel” report can be generated.

- 1 In Sierra Web select **Tools:Reports** and generate the standard report “Detected proteins in gel” as described in Ettan 2D-MS and LWS Software Laboratory Guide.

Note: Select how much information to import into DeCyder BVA by altering the maximum rank and by including sub-rankings or not.

The report will appear on screen.

Sierra REPORTING HOME LOGOUT GO: [dropdown]

Detected proteins in gel

Showing which proteins are represented in a gel.

Gel ID:

Include sub-rankings: Yes No

Maximum Rank:

Home Separate Window All Results No Grouping Excel XML Comma Separated Tab Separated Page 1

Total number of spots in gel: 96 **Summary of Gel ID 1760023632112339389**

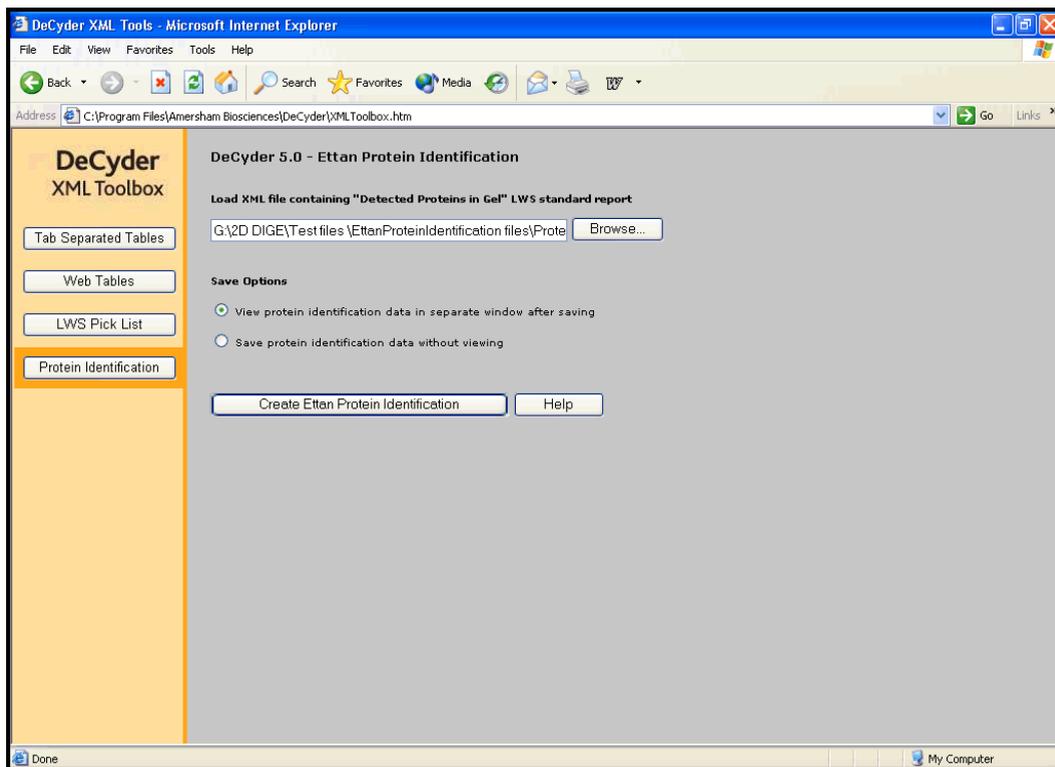
Upstream gel information

Title	Value
Gel ID	1760023632112339389
Created	Second Dimension Electrophoresis: 2002.11.11-16:11
1D Strip	1760024629378537448

Fig 8-6. Save the report “Detected proteins in gel” as XML.

- 2 Click **XML** in the middle of the dialog to save the file. Save it to a location accessible by XML Toolbox.

8.6.2 Using the Protein Identification tool



- 1 Either browse for a file or type the path of the XML file containing the “Detected Proteins in Gel” LWS standard report into the **Load XML file containing “Detected Proteins in Gel” LWS standard report** field in the middle of the tool page.
- 2 Select either **View template spot map in separate window before saving** or **Save template spot map without viewing result**.
- 3 Click **Create Ettan Protein Identification** to start the conversion procedure. At the end of the conversion procedure a **Save As** dialog is displayed. Enter the name of the file where the Ettan protein identification file is to be saved. Save the file in a location accessible by DeCyder Differential Analysis Software.

8.7 Importing protein data templates into the BVA module

- 1 Open the original BVA workspace and select **File:Import:Protein Data**.

Note: If the BVA module suspects that the file for importing originates from a different picking context than the current workspace a warning is displayed. It is, however, still possible to import the information if desired, but it is not recommended.

8.8 Database queries in DeCyder BVA

In DeCyder BVA it is possible to make online database queries based on the protein ID of the individual proteins. The protein ID:s generated by Ettan MALDI-ToF Pro follow the notation used in the NCBI nr (non-redundant) database.

To be able to use the database query function of DeCyder BVA based on the protein ID:s from Ettan MALDI-ToF Pro, a new query definition that matches the NCBI nr format must be added in the database property page.

When adding a new database query definition enter information in the two fields in the dialog **Edit web database label and address**:

- 1 In the **Label** field type **NCBI nr: Protein search based on ID** or something similar.
- 2 In the **Address** field type the following search string:

```
http://www.ncbi.nlm.nih.gov/entrez/eutils/  
efetch.fcgi?rettype=gp&retmode=html&db=protein&id=#BVA_PROTID#
```

- 3 Click **OK**.

Add new web database label and address

Label

NCBI nr: Protein Search based on ID

Address

A placeholder string has to be included in the address, and will be replaced by the protein's ID or AC.
Eg: http://www.expasy.ch/cgi-bin/get-sprot-entry?#BVA_PROTID#

#BVA_PROTID#	- Protein ID
#BVA_PROTAC#	- Protein AC
#BVA_PROTIDAC#	- Both Protein ID and AC

[_gov/entrez/eutils/efetch.fcgi?rettype=gp&retmode=html&db=protein&id=#BVA_PROTID#](http://www.expasy.ch/cgi-bin/get-sprot-entry?#BVA_PROTID#)

Test

Select a Protein ID or AC that will replace the placeholder string.

Test

OK Cancel Help

Fig 8-7. Information to enter in the dialog **Edit web database label and address** when adding the new database query definition.

For further information on using the database query function and adding new database query definitions see section 4.12 or DeCyder Differential Analysis online help.

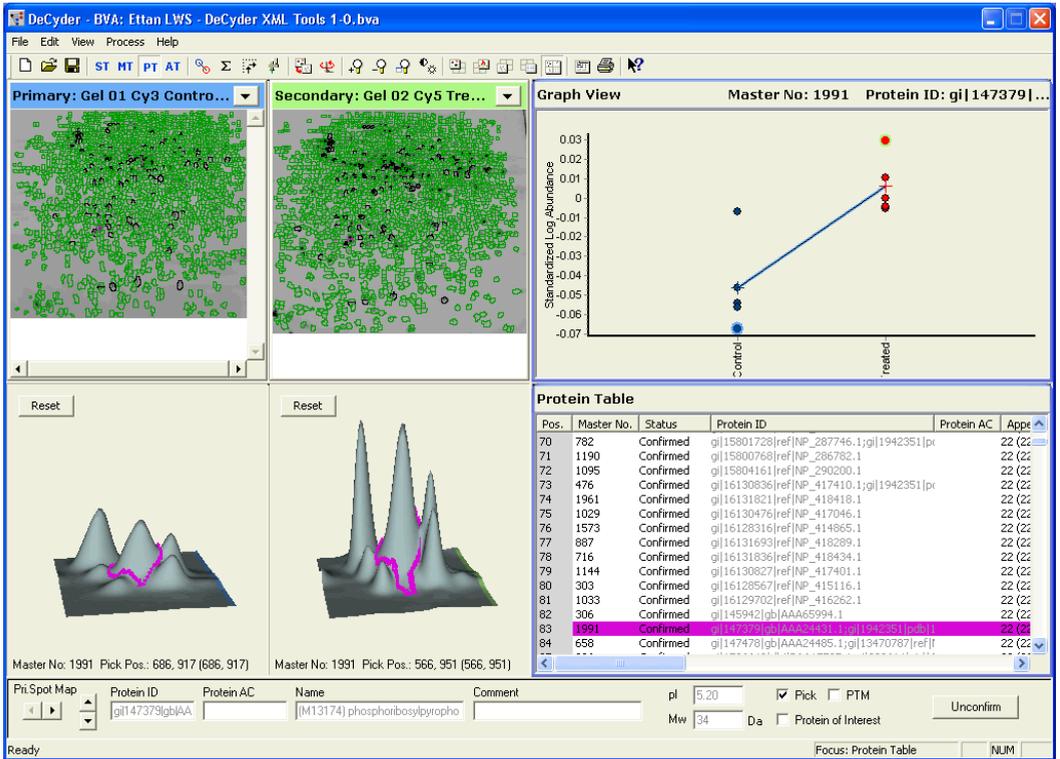


Fig 8-8. Example of results in DeCyder BVA after import and database query.

9 Tutorials Introduction

9.1 DeCyder Differential Analysis Software structure

The DeCyder Differential Analysis Software suite is made up of four user interfaces:

DIA (Differential In-gel Analysis) – Protein spot detection and quantitation on images, from the same gel.

BVA (Biological Variation Analysis) – Matches multiple images from different gels to provide statistical data on protein expression levels.

Batch Processor – Fully automated image detection and matching of multiple gels without user interaction.

XML Toolbox – Comprises tools for the extraction of data from the different XML files produced in DeCyder Differential Analysis Software.

DIA

DeCyder DIA processes a set of gel images (each saved as either 16-bit TIFF or customized *.GEL* files) from a single gel. Each image in the pair is generated from samples labelled with different fluors. Images must be processed in the DIA interface prior to data analysis in BVA. The DIA algorithms detect spots on a combined image derived from merging individual images from an in-gel set of images. This co-detection ensures that all spots are represented in all images. DIA then quantitates spot protein abundance for each image and expresses these values as a ratio thereby indicating changes in expression levels by direct comparison of corresponding spots.

The data can be saved as a *.DIA* file from which spot pick lists can be exported as a text file. In addition data can be exported in an XML format (which can be queried using the XML toolbox) for multi-gel analysis in DeCyder BVA or copied directly from DIA and pasted into applications such as Microsoft Word and Excel.

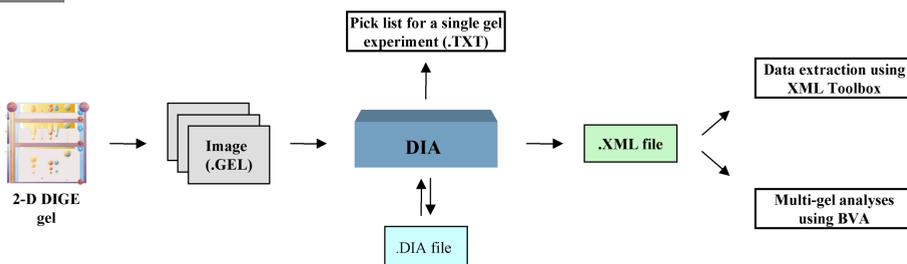


Fig 9-1. Schematic representation of DeCyder Differential Analysis Software DIA workflow.

BVA

DeCyder Differential Analysis Software BVA utilizes the DIA exported .XML files together with the original gel images to match protein spots on different gels. Data is then subjected to statistical analyses to accurately assess protein expression changes occurring in biological replicates comparing different conditions/treatments. This multi gel approach, which allows analysis of replicates, provides greater statistical validity to findings.

The final data can be saved as a .BVA file, from which spot pick lists and .XML files can be exported. In addition data can be copied directly from BVA and pasted into applications such as Microsoft Word and Excel.

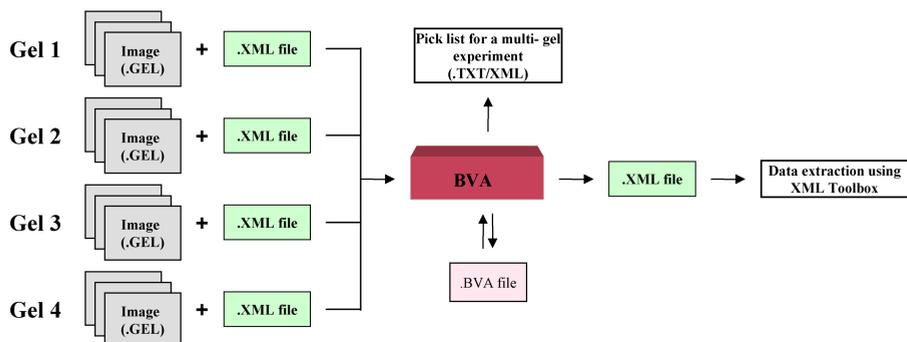


Fig 9-2. Schematic representation of DeCyder Differential Analysis Software BVA workflow.

Batch processor

The Batch Processor integrates both the DIA and BVA modules, enabling fully automated processing of multiple gels with minimal user intervention. The Batch Processor can be configured to analyze several gels in the DIA exclusively. Alternatively multiple gels can be processed through both interfaces to produce a final .BVA file.

XML Toolbox

The XML Toolbox enables the extraction of user specific data from the XML files generated in either the DIA or BVA modules. This data can be saved in either text or html format enabling users to access data in the DeCyder Differential Analysis Software workspaces for other applications. The XML Toolbox also provides an interface for linking Ettan DIGE system data with Ettan Laboratory Workflow System.

9.2 Scope of tutorials

The following tutorials are aimed at introducing the functionality of DeCyder Differential Analysis Software within the context of an actual experiment. The tutorials have been designed to be a step by step guide utilizing gel images and DeCyder Differential Analysis Software files which are co-installed with the tutorials. The four tutorials cover different aspects of the software suite. They are all self contained and can be undertaken independently.

To assist the user each tutorial includes a completed version of the DeCyder Differential Analysis Software file, which the tutorial is designed to generate. These files all include the word **finished** in their names.

Tutorial I

The DIA is used exclusively to demonstrate a small-scale experiment to assess protein changes in samples that are limited. See chapter 10.

Tutorial II

The second tutorial encompasses a similar experiment to that found in tutorial 1. However, this tutorial utilizes gels that have already undergone spot detection in the DIA and have been exported in an XML format. The processes involved in configuring the BVA, without the Batch Processor, are demonstrated. See chapter 11.

Tutorial III

The processes required to analyze a “picking” gel for subsequent spot picking, digestion and mass spectrometry analysis are demonstrated. See chapter 12.

Tutorial IV

The tutorial employs all three software interfaces to ascertain protein expression changes in an experiment containing replicate gels in order to generate a pick list for Ettan Spot Picker. The entire workflow is coordinated via the Batch Processor, thereby providing a fully automated process. See chapter 13.

The tutorials described above introduce the concepts and functionality of DeCyder Differential Analysis Software. It is therefore recommended that these tutorials are performed first to gain a preliminary understanding of the software. However it is recommended that various elements of the tutorials are used when analyzing an experiment incorporating an internal standard. A workflow describing the various steps involved in an entire experiment are described in detail in appendix A.

10 Tutorial I - Using DIA module for preliminary investigation of protein changes

10.1 Objective

This tutorial describes how to find protein abundance differences above experimental variation between a control and a treated protein lysate of bacterial cultures using DeCyder Differential Analysis Software DIA (Differential In-gel Analysis). The procedure in this tutorial can be applied to any two protein mixtures/lysates.

This experimental design can be used to rapidly identify proteins that show a substantial change in abundance from control to treated sample and is often used at the preliminary stages of an experiment.

The method is used to gain early information for the samples in question, and to check that the experimental procedures are appropriate for the new samples.

The major drawbacks of this approach are that only one control sample and one treated sample are studied. There are no gel replicates or biological replicates and therefore the biological relevance and statistical validity of an altered protein is impaired. This experimental design only addresses differences above system variation (see section 1.3 for different sources of variability). Since there are no biological replicates the inherent biological variation within populations is not considered. This is also true if the biological control and treated sample are pools of control and treated populations, as this approach is looking at averages but not variation associated with the different populations.

10.2 Overview

- 1 An experimental design using two gels is set up, “control-control” gel and “control-treated” gel.
- 2 The control-control gel is first analyzed in the DIA module. From this, an indication of overall system variation is determined by elucidating the fold change 95% of spots lie within, e.g. 95% of spots on control-control gel are within 1.5 fold.
- 3 The control-treated gel is subsequently analyzed in the DIA module. As the control-control gel gave a result that 95% are within X fold (e.g. 1.5 fold) then any spots with a fold change

greater than X fold (e.g. 1.5 fold) has a 95% confidence that it is a difference between the control and treated sample above experimental variation.

- 4 The spots greater than X fold on the control-treated gel are looked at in detail. These can be assigned with a protein of interest status and can be confirmed in preparation for spot picking. Spot picking from a preparative gel is covered in tutorial III.

10.3 Experimental design

Two gels were loaded with bacterial lysates as indicated in the table below.

<i>Gel</i>	<i>Cy3</i>	<i>Cy5</i>
1	Control	Control
2	Control	Treated

The control-control gel (gel 1) is analyzed to determine the level of experimental variation. The protein abundance variation calculated in DIA indicates the degree of similarity between the protein spots from two images from the same gel. Ideally there should be no difference between the proteins from the two control-control gel images (because the same sample is being run on the same gel). However due to different experimental factors (for example, pipetting variation) there is some intrinsic variability.

The data analysis is performed in three parts using the DeCyder Differential Analysis Software DIA module:

- Analysis of control-control gel
- Analysis of control-treated gel
- Creating a spot pick list

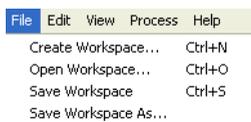
10.4 Analysis of control-control gel

10.4.1 Selecting gel images

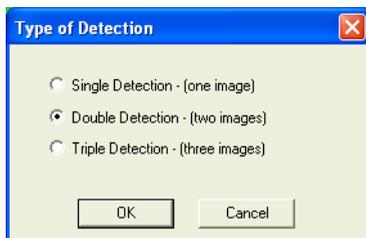


- 1 Double click the DeCyder Differential Analysis Software **DIA** icon on your desktop to open the DIA module.

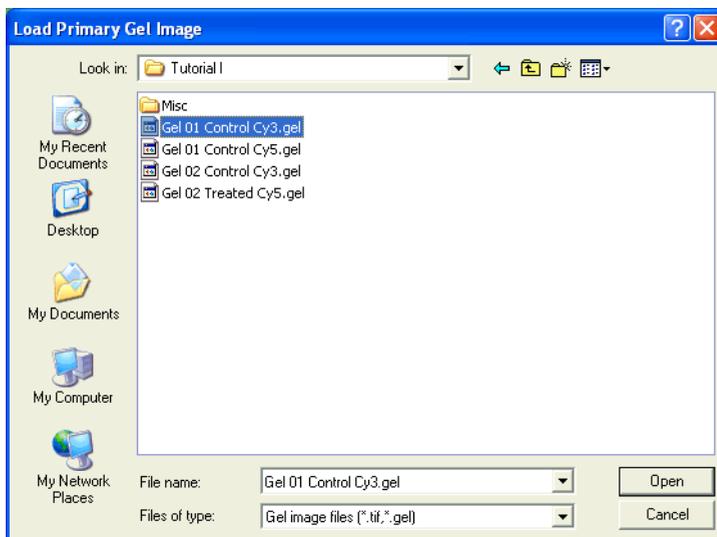
-
- 2 Select **File:Create Workspace**.



- 3 Select **Double Detection**, then click **OK**.



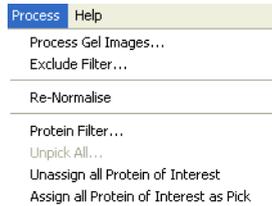
- 4 Browse to locate the folder entitled **Tutorial I** in the **Tutorial data files** folder. Double click on the image **Gel 01 Control Cy3** to select the primary gel image.



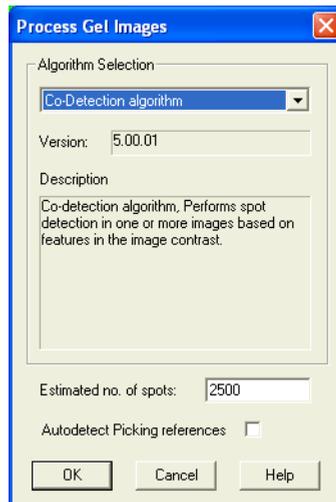
- 5 When the **Load Secondary Gel Image** window appears, double click on the image **Gel 01 Control Cy5**.

10.4.2 Spot detection

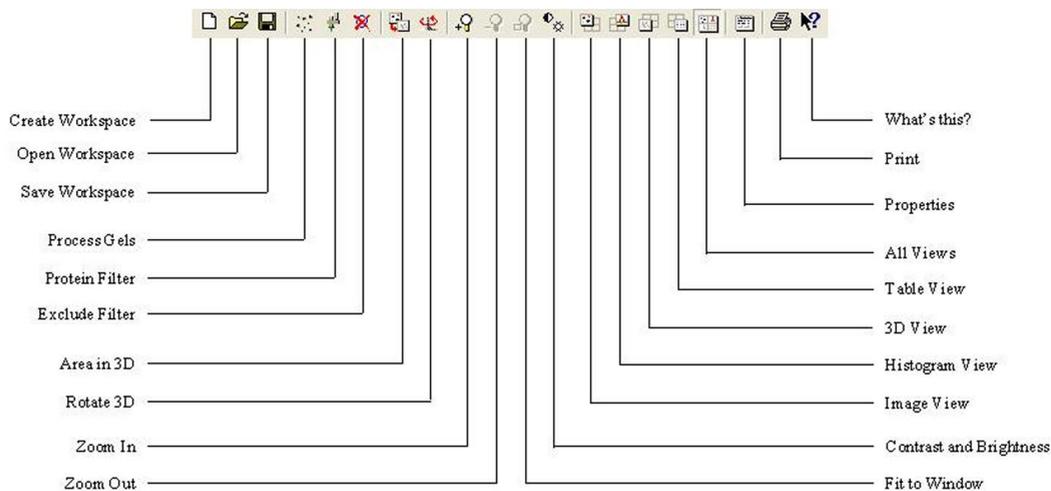
- 1 After the images have been loaded select **Process:Process Gel images**.



- 2 In the process gel images dialog box which now appears enter the value **2500** for the estimated number of spots (see help file for an explanation of this value). Click **OK** to begin spot detection on the images.



Before going any further, it is important to become acquainted with the main tool bar.



The user interface is divided into four main windows, the Image View, the Histogram View, the 3-D View and the Table View. In addition the Data Control Panel at the bottom of the screen allows user defined attributes to the spot selected in the four views.

The screenshot shows the DuCyder DIA software interface. It features four main view windows and a Data Control Panel at the bottom.

- Image View:** Shows two gel images: "Primary Gel 02 Control Cy3..." and "Secondary Gel 02 Treated...".
- Histogram View:** A scatter plot of "Number of Spots" vs "Log Volume Ratio". It includes a "Histogram selections" panel with parameters like "Scatter parameter: Max Volur", "Threshold mode: 2 model S", "Threshold: 2.41551", and "2 S.D.: 2.32553".
- 3-D View:** Two 3D surface plots of spot data. The left plot shows a peak with "Volume: 3.856e+005" and "Pick pos.: 1143, 1106". The right plot shows a peak with "Spot No: 2309", "Position: 1143, 1106/P", and "Pick pos.: A1".
- Table View:** A table listing spot data with columns for Status, Spot No., Abundance, Excluded, Volume Ratio, and Picked.

Status	Spot No.	Abundance	Excluded	Volume Ratio	Picked
Unconfirmed	2317	Similar		-1.19	
Unconfirmed	2316	Similar		-1.82	
Unconfirmed	2315	Decreased		-13.86	
Unconfirmed	2314	Similar		-1.30	
Unconfirmed	2313	Decreased		-5.26	
Unconfirmed	2312	Similar		-1.73	
Unconfirmed	2311	Similar		-1.34	
Unconfirmed	2310	Similar		-1.17	
Unconfirmed	2309	Similar		-1.19	
Unconfirmed	2308	Similar		-2.27	
Unconfirmed	2307	Similar		-1.45	
Unconfirmed	2306	Similar		-1.42	
Unconfirmed	2305	Decreased		-4.64	
Unconfirmed	2304	Similar		1.92	
Unconfirmed	2303	Similar		-1.21	
Unconfirmed	2302	Similar		1.26	
Unconfirmed	2301	Similar		-1.59	
Unconfirmed	9999	Similar		-1.18	
- Data Control Panel:** Located at the bottom, it includes fields for "Protein ID:" and "Comment:", and buttons for "Pick", "PTM", "Protein of Interest", "Exclude", and "Confirm".

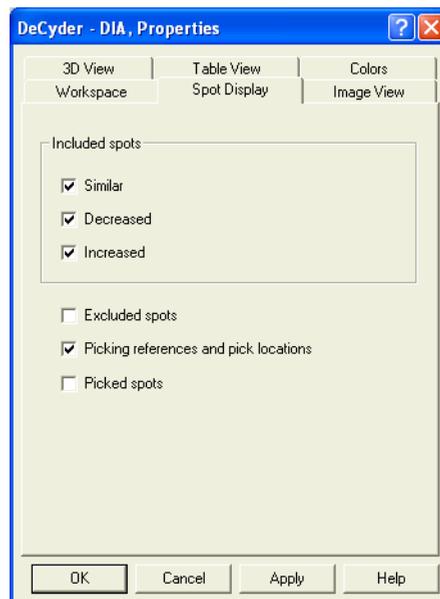
The four views are all linked. Clicking on a spot on the Image View highlights the spot in magenta in the Histogram View. The spot is also represented three dimensionally and is highlighted in gray in the Table View.



- 3 To display the entire gel image, click on the **Fit to window** icon.
- 4 After detection it is important to save the workspace. Select **File:Save Workspace As**, and place an appropriate name in the dialog box, which now appears.

The detected spots are of three types: increased, decreased and similar (colored blue, red and green, respectively, in the Histogram View)

- 5 If, after detection (the hourglass turns back to a cursor) the Table View remains empty click on the **Properties** icon to bring up the properties dialog box. By default this will open on the Workspace tab.
- 6 Change to the table options by clicking on the **Table View** tab.
- 7 By selecting and deselecting the check boxes the various categories of spots can be selectively displayed. Deselecting all the check boxes results in the Table View being blank. Similarly, spots can be selectively displayed in the image view using the Spot Display tab. The Table View tab works on OR logic, i.e a spot only needs to conform to one criterion to be shown in the table.

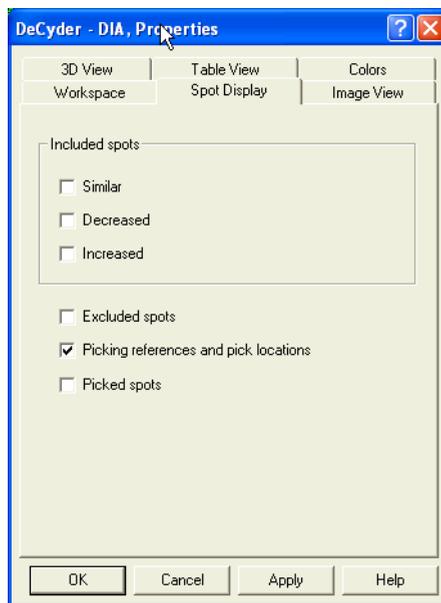


Selecting different protein spots on the images reveals that some of the spots that have been detected are in fact either dust particles or artifacts from the gel. To remove these artifacts an Exclude filter must be used on the images.

10.4.3 Assigning an area of interest

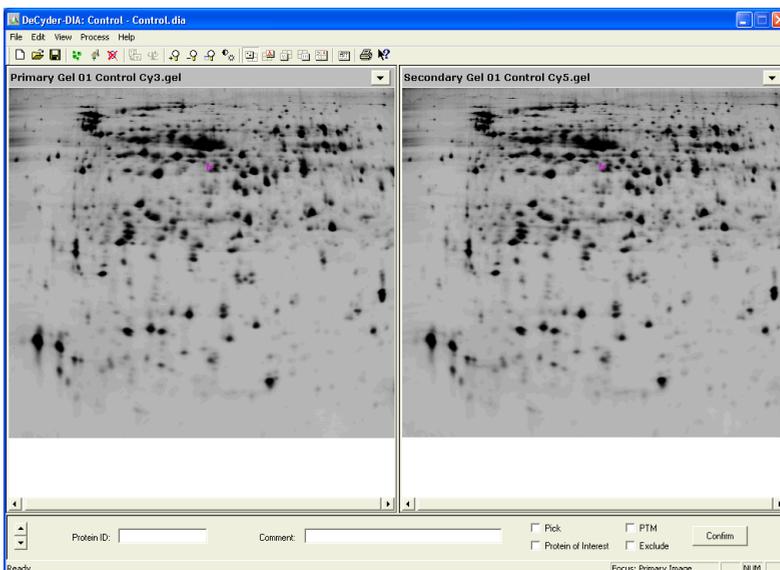
Due to gel heterogeneity at the edges of the images, there are artifacts that need to be removed. These can be removed by setting an area of interest. This function only works with the filter. If an area of interest was set before detection all the spots on the image will still be detected (even those outside the area of interest).

- 1 Click on the **Fit to window** icon on the tool bar to fit the images to the Image View.
- 2 Click on the **Image view** icon on the tool bar to have a full screen view of the gel images.
- 3 Next, click on the **Properties** icon to bring up the properties window.
- 4 Select the **Spot Display** tab and deselect **Similar**, **Increased** and **Decreased**, so that the check boxes are identical to those shown in the figure below. Click **OK**.



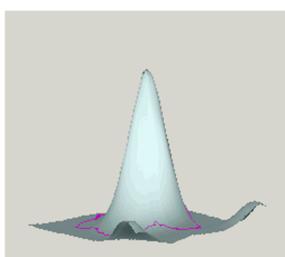
- 5 To set an area of interest select **Edit:Define area of interest**. Using the rectangular target pointer which now appears drag the mouse to draw a rectangle around the gel, ensuring that edge artifacts are

removed. It does not matter if the Reference Markers are inside or outside the Area of Interest.

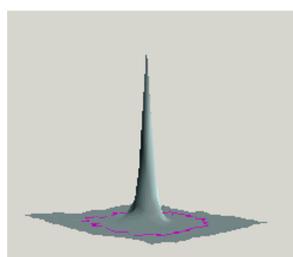


10.4.4 Gel artifact removal

A 3-D View clearly shows if a detected spot is a gel artifact rather than a protein spot due to the very steep sides and pointed top of an artifact compared to the smooth curve of a protein spot.



Actual Protein



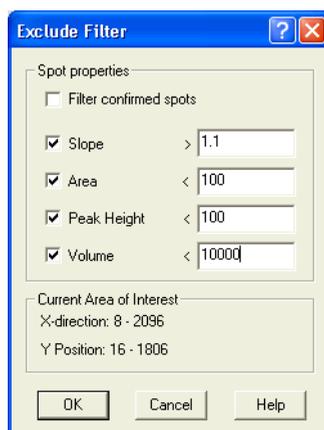
Dust Particle

In order to exclude these artifacts from subsequent analysis a set of Exclude filter parameters must be generated.

- 1 In the Table View click on the column header **Max slope** until the spot with the highest max slope is present at the top of the table.

Max Slope	Area	Max Peak...	Max Volume
-----------	------	-------------	------------

- 2 Scroll down the table from one spot to another until the next spot in the table is an actual protein spot and not a dust particle or gel artifact. Make a note of the Max slope value of the artifact immediately before the real protein spot. This value will be used for the Slope category in the Exclude filter.
- 3 Repeat this procedure for **Area**, **Max Peak Height** and **Max Volume**, but in each case ordering the table so that the smallest value is at the top of the table.
- 4 When all the values have been found, select **Process:Exclude Filter** to display the **Exclude Filter** dialog and enter the found values. Click **OK**.



An example of a non-stringent set of filter parameters that can be used to run a light filter is as follows:

Slope	1.1
Area	100
Peak Height	100
Volume	10000

10.4.5 Ascertaining the 2 S.D. (standard deviation) threshold value

The 2 S.D. value calculated for the control–control gel is the fold ratio that encompasses 95% of the spots and can be used as a guide to the level of experimental variation. This value should be noted and can be set as the threshold mode when analyzing the control-treated gel.

The protein differences being observed on the control-treated gel will be above the experimental variation.

In the Histogram View is a Histogram Selections toolbar with two drop down menus and two information boxes. Make a note of the value in the information box entitled **2 S.D.**

10.5 Analysis of control-treated gel

10.5.1 Selecting gel images

- 1 Select **File:Create workspace** (there will be an automatic prompt to save the current workspace. This is unnecessary in this tutorial, therefore click **NO** to continue).
- 2 Select **Double Detection**, then click **OK**.
- 3 Browse to locate the folder entitled **Tutorial I**, in the **Tutorial data files** folder. Double click on the image **Gel 02 Control Cy3** to select the primary gel image.
- 4 When the **load secondary gel image** window appears, double click on the image **Gel 02 Treated Cy5**.

10.5.2 Spot detection

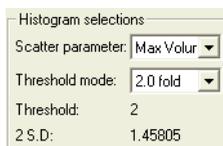
- 1 After the images have been loaded select **Process:Process gel images**.
- 2 In the process gel images dialog box which now appears enter the value **2500** for the estimated number of spots (see help file for an explanation of this value).
- 3 Click **OK** to begin spot detection on the images.

10.5.3 Spot filtering

- 1 Assign an Area of interest and Filter as described for the control-control gel, using the same filter parameters.

10.5.4 Setting a threshold

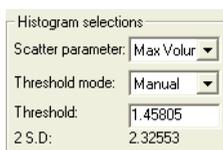
- 1 Using the **Threshold Mode** drop-down menu in the Histogram Selections toolbar of the Histogram View, thresholds can be set to highlight protein spots that differ in their abundance. For example, to select spots whose volume ratios differ between samples by 2 fold or more, select 2.0 fold from the drop down menu.



Histogram selections

Scatter parameter:	Max Volur
Threshold mode:	2.0 fold
Threshold:	2
2 S.D.:	1.45805

A variety of values in the Threshold Mode box can be selected, such as 2 S.D. which is based on the model histogram (blue line) in the Histogram View. Alternatively the threshold mode can be set manually by selecting **Manual** and then entering a figure in the **Threshold information** box. Enter the 2 S.D. value ascertained in the control-control gel into the information box. As previously discussed, this value represents the inherent experimental variation. It can therefore be hypothesized that spots with ratios which fall outside this threshold in a control/treated experiment have some level of significance. As a general rule, the higher the threshold set the more confidence that differences above the value is significantly different above experimental variation.



Histogram selections

Scatter parameter:	Max Volur
Threshold mode:	Manual
Threshold:	1.45805
2 S.D.:	2.32553

10.6 Assigning protein of interest

10.6.1 Selecting spots for picking



- 1 To select those proteins that have been found to be different on the control–treated gel, select the **Protein Filter Dialog** icon.

This process can be performed prior to or after spot confirmation. The amount of time spent on spot confirmation can be reduced by performing filtering first.

The Protein Filter selects protein spots based on user defined criteria such as volume ratio, above a maximum volume or peak height. The protein filter is therefore used as a filter to find proteins that are interesting which can then be confirmed.

- 2 Open the **Protein Filter** dialog window. Ensure that the **Assign Protein of Interest** check box is selected and that the **Assign Pick status** is not selected. This results in the proteins successfully filtered being given a **Protein of Interest** status.
- 3 For the purpose of this tutorial, select to assign protein spots with a volume ratio greater the 2 SD value from the control-control gel and a max volume between 10^5 and 10^8 , for example. Click on **Filter** to calculate the number of spots that meet the criteria. These values can be adjusted and the **Filter** button clicked again to see how many spots now meet the criteria. The purpose of this filtering is to narrow down the users search for interesting proteins that are significantly changing. Click **OK**.

Protein Filter

Filter action

Assign Protein of Interest Assign Pick status

General filter settings

Pick all

Restrict to Confirmed spots

Select spots with

Volume Ratio \geq 1.4

or

Volume Ratio \leq -1.4

Volume Ratio \geq 1.05 and \leq 1.05

Area \geq 2000

Max Volume \geq 1.00e+005 and \leq 1.00e+008

Max Peak Height \geq 20000 and \leq 100000

X co-ordinate \geq 0 and \leq 1000

Y co-ordinate \geq 0 and \leq 1000

Filter

OK Cancel Help

- 4 The subset of spots that met the criteria will now be assigned with a protein of interest in the protein table (denoted by the letter “I” in the Function column).
- 5 Select **View:Properties** or click on the **Properties** icon. Go to the **Table View** tab and select the **Protein of Interest** check box, and uncheck all of the other options. The table view will now display only those spots assigned as protein of interest in the protein filter. These are the spots that are of most interest and they can now be manually confirmed.

10.6.2 Spot confirmation

Three options are available when you start to confirm your spots:

- 1 Only confirm spots that are Increased (blue) and/or Decreased (red) in their abundance. This is the recommended option. Select **View:Properties**, and go to the **Table View** tab. Ensure that the **Decreased** and **Increased** options are selected, and that the **Excluded** box is not selected. Also ensure that **Confirmed** is selected, unconfirmed is not. Click **OK**.

- 2 Only confirm spots that have been assigned with a protein of interest status (recommended option when the protein of interest assignment option has been performed).
Select **View:Properties**, and go to the **Table View** tab. Ensure that only the **Picked spots** checkbox is selected.
- 3 All spots, Decreased (red), Increased (blue) and Similar (green) are manually verified. (This takes approximately 1.5 hours for 1000 spots - this is not recommended as only those spots that are increasing or decreasing in abundance are of interest).
Select **View:Properties**, and go to the **Table View** tab. Ensure that the **Similar, Decreased** and **Increased** options are selected, and that the **Excluded** box is not selected. Click **OK**.

When confirming spots, a decision is made whether the spots selected are true-protein spots or non-protein spots, such as a particle of dust that has not been removed by the Exclude Filter function. This can be determined by using the **Image View** and the **3-D View**.

Click **Confirm** if the spot is a genuine protein spot. The next spot in the list will now be automatically displayed. If it is non-protein spot, select **Exclude** in the **Spot Assignment** toolbar and then click **Confirm**.



Continue with this process until you have confirmed all the relevant spots.

10.6.3 Exporting the Pick List and physically excising spots from the gel

To excise the spots from the gel a separate preparative gel should be used. This is then matched to this analytical gel set in DeCyder Differential Analysis Software BVA module and the protein of interest status is transferred to the relevant matching spots on the preparative gel. The pick list is based upon these proteins. The pick list with x, y coordinates of spot positions is then exported based on the preparative gel. The pick list with the physical preparative gel is then taken to Ettan Spot Picker or Ettan Spot Handling Workstation for spot excision and later analysis by mass spectroscopy.

Generating a pick list is covered in more detail in Tutorial III.

11 Tutorial II - Employing an internal standard to Analyze Protein Changes

11.1 Objectives

Tutorial I describes the analysis of a simple experimental design using only the DIA module to evaluate samples where an internal standard is not possible due to sample scarcity. This tutorial extends to a more complex experimental design incorporating an internal standard with several replicate samples.

This tutorial describes how to find proteins that exhibit statistically significant changes between control and treated groups of bacterial cultures using DeCyder Differential Analysis Software. The procedure in this tutorial can be applied to any two groups of protein mixtures with either replicate gels or replicate biological samples. The tutorial outlines the various stages in experimental design, sample organization, protein detection and quantitation, gel matching and statistical analysis.

11.2 Overview

The following describes the various stages involved in identifying all proteins that are differentially expressed in a given system, and which therefore warrant further investigation. The stages are:

- 1 An experimental design is devised that will generate statistically significant results; a design which minimizes or eliminates in-gel and gel-to-gel system variation.
- 2 Eight sample lysates that form the basis of the experiment are prepared. This consists of four lysates derived from four bacterial cultures treated with benzoic acid and four control flasks that have not been treated. Since the gel to gel variation in this system is low, gel replicates are not compulsory if biological replicates are available.
- 3 Aliquots from each sample are taken and pooled to prepare a standard sample. All three sample types (standard, control and treated) are labelled with an appropriate CyDye DIGE Fluor minimal dye (Cy2, Cy3 or Cy5 dye) as described in the table overleaf. Each sample is then applied to the gels.

- 4 The samples are separated within 4 gels by 2-D electrophoresis. Three images are produced from each gel, by scanning at appropriate wavelengths for each of the three fluors. The recommended scanning instrument for this is the Typhoon 9000 series Variable Mode Imager.
- 5 The images from the first gel are loaded into the DIA module.
- 6 Spot detection and calculation of the spot properties are performed for all images from the same gel.
- 7 Protein spots are normalized using the in-gel linked internal standard.
- 8 XML formatted data for the first gel is exported, for subsequent matching and analysis in the BVA module.
- 9 This XML data for the first gel and prepared XML data from the remaining gels are loaded into the BVA module.
- 10 Spot maps (images) are assigned to groups and the experimental design is set up in the BVA module.
- 11 All gels or those gels that require it to aid in the matching process are landmarked. Matching is then performed.
- 12 Spots that conform to certain criteria such as statistical significance and above a certain ratio change are assigned as proteins of interest.
- 13 These spots are confirmed for eventual inclusion into a pick list.

The individual stages involved in the experiment are now described more fully.

11.3 Experimental design

Four replicate gels are loaded with bacterial lysates as indicated in the table below.

<i>Gel number</i>	<i>Cy2</i>	<i>Cy3</i>	<i>Cy5</i>
Gel 1	Standard	Control 1	Treated 1
Gel 2	Standard	Treated 2	Control 2
Gel 3	Standard	Control 3	Treated 3
Gel 4	Standard	Treated 4	Control 4

Each gel contains a standard sample to normalize control and treated samples against. The standard sample is derived from control and treated lysates, which are pooled in equal concentration. The set of three images generated from a single gel are sequentially processed using the DIA module, in order to perform spot detection and spot quantitation against the internal standard.

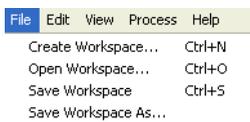
The data from DIA analysis is exported as an XML file for further analysis in the BVA module to evaluate protein abundance differences.

11.4 Spot detection and quantitation

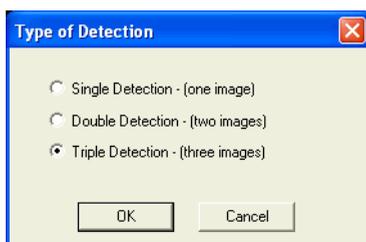
Spot detection and quantitation by normalizing against the internal standard is performed using the DeCyder Differential Analysis Software DIA (Differential In-gel Analysis) module. Gels are processed sequentially in the DIA module.

11.4.1 Selecting gel images

- 1 Double click the DeCyder Differential Analysis Software **DIA** icon on your desktop to open the DIA module. When the user interface appears, select **File:Create Workspace**.



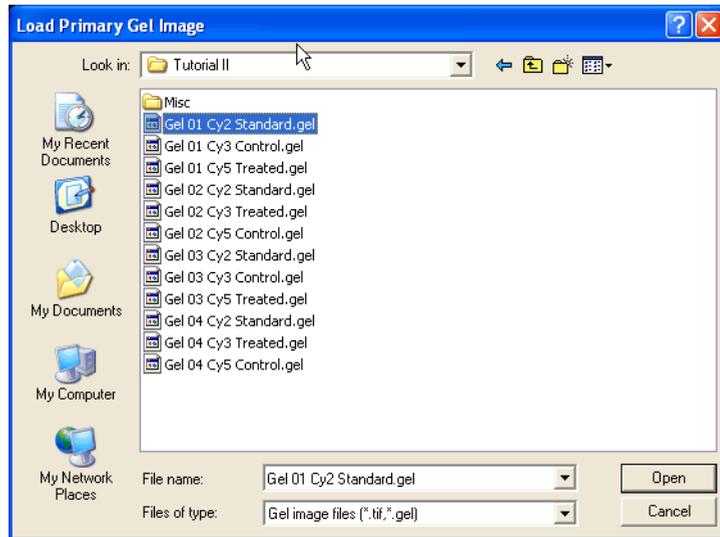
- 2 Select **Triple Detection**, then click **OK**



- 3 Browse to locate the folder **Tutorial II** in the **Tutorial data files** folder.

Note: All the images are named in the style **Gel ON Cy X Standard, Control or Treated**.

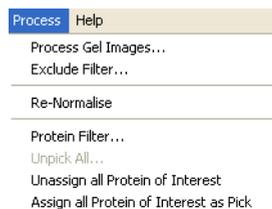
- 4 Select **Gel 01 Cy2 Standard**, then click **Open** to load the **Primary Gel Image**.



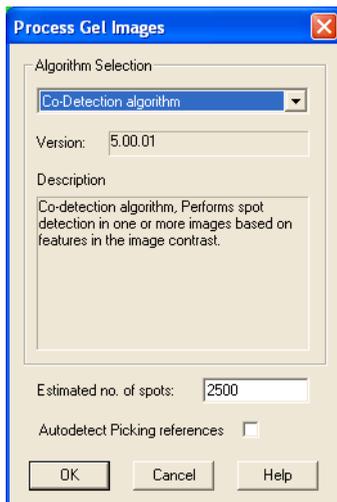
- 5 Select **Gel 01 Cy3 Control**, then click **Open** in the **Load Secondary Gel Image** window.
- 6 Select **Gel 01 Cy5 Treated**, then click **Open** in the **Load Tertiary Gel Image** window.

11.4.2 Spot detection and quantitation

- 1 After the images have been loaded select **Process:Process Gel Images**.



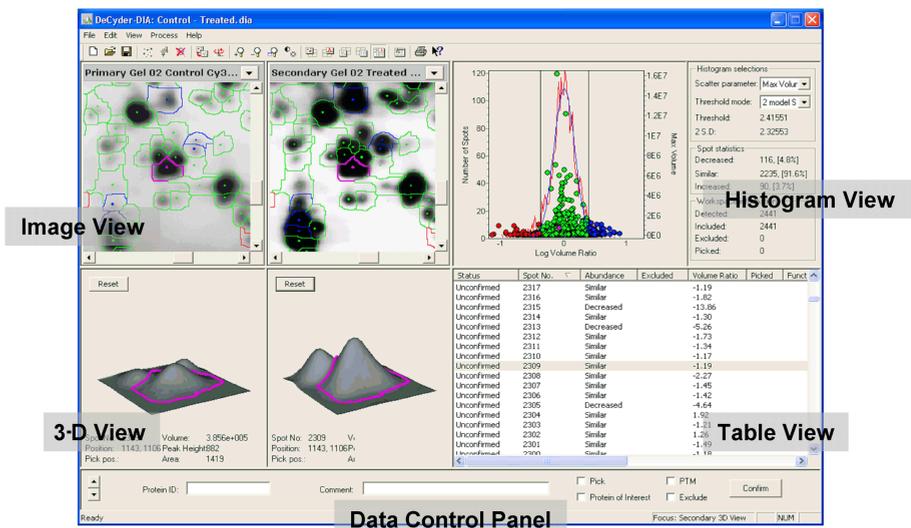
- 2 Enter the value **2500** for the estimated no. of spots (see section 3.3.1 for an explanation for this value). Then click **OK** to begin spot detection and quantitation on all three images.



The DIA analysis takes between 3 and 10 minutes depending on the specifications of the computer.

11.4.3 Viewing the DIA workspace

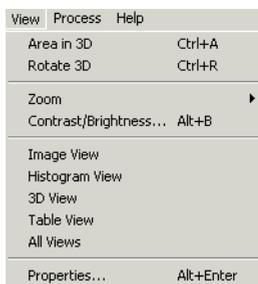
- 1 Upon completion of the DIA processing, the spot map data is displayed in the workspace (as shown below).



The user interface is divided into four main windows that show the Image View, the Histogram View, the 3-D View and the Table View.

The four views are linked. Clicking on any spot on the Image View

highlights the spot in magenta in the Histogram View. The spot is also represented three dimensionally and is highlighted in the Table View. Each of the four views can be expanded to fill the whole screen. To do this, click on the **View** button and the menu below is displayed.

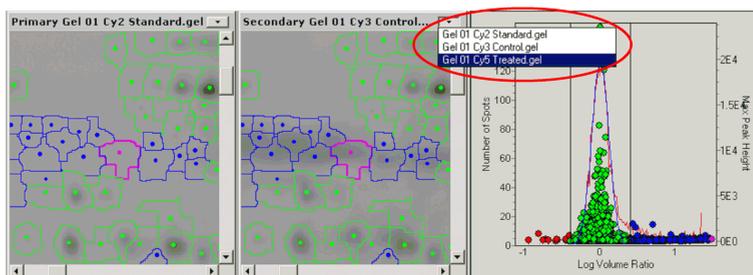


- 2 Click on any of the four views available to expand that view to fill the screen. After expanding all the views in turn, click on **View** and select **All Views** from the menu above to return to the **All views** display. Alternatively, use the icons displayed on the main toolbar.



*Note: If, after detection has been completed, the Table View remains empty, click on the **Properties** icon to bring up the properties dialog box. Alternatively click on **View:Properties**. By default this will open on the Workspace tab. Change to the Table View options by clicking on the **Table View** tab.*

- 3 The gel images displayed in the image view can be changed using the pull down menus in the image view title bar.



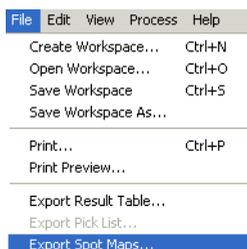
When using an experimental design that includes an internal standard, as in this case, it is conventional that the primary image view stays as the image of the internal standard whereas the secondary view displays the images of the analytical gels.

-
- 4 Spot quantitation, expressed as a volume normalized against the internal standard, is calculated in the DIA module then displayed in the Table View under the **Volume Ratio** column. This column is automatically amended when selecting between analytical gels in the secondary gel view.

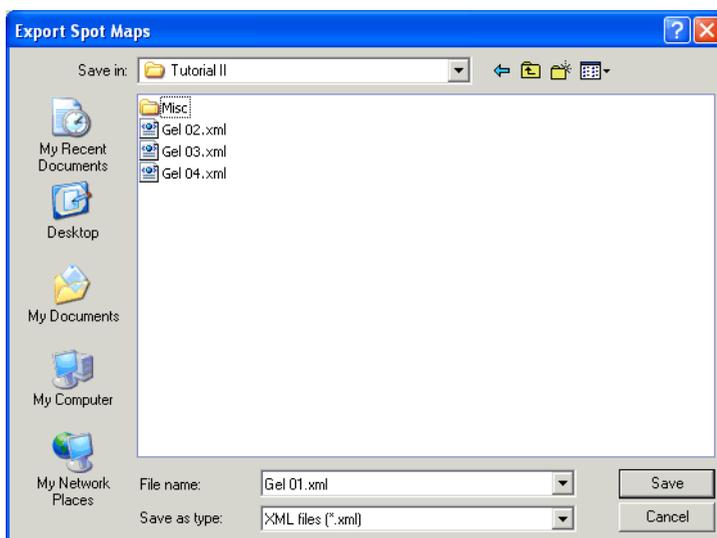
11.4.4 Exporting spot map data

The data associated with the spot boundaries and spot quantitation are employed by the BVA module to perform inter-gel matching and further statistical analysis. This data is therefore exported from the DIA module in an XML format.

- 1 Select **File: Export Spot Maps...**



- 2 Browse to locate the folder entitled **Tutorial II** in the **Tutorial data files** folder. Name the exported file **Gel 01**, then click **Save**.

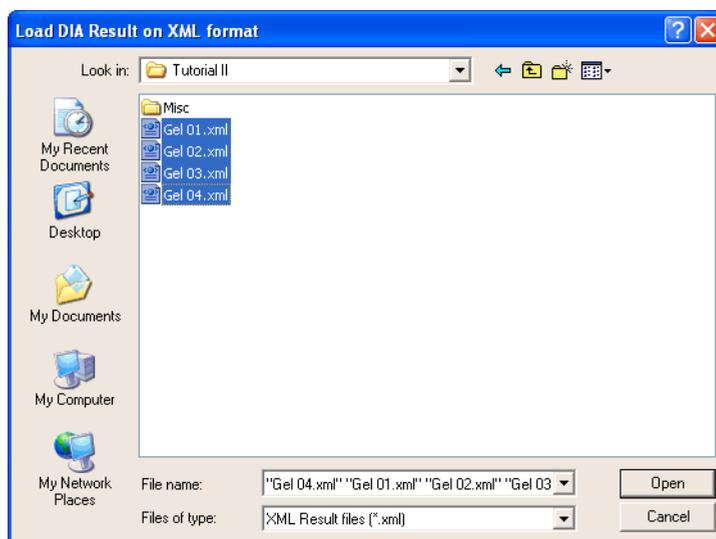


- 3 The remaining gels must be subjected to similar analyses using the DIA module. For the purposes of this tutorial, the gels have been prepared and placed in the **Tutorial II** file, ready for analysis in the BVA module.
-

11.5 Creating the BVA workspace

11.5.1 Selecting gels

- 1 Double click the DeCyder Differential Analysis Software BVA icon on your desktop to open the BVA workspace. Select **File:Create workspace** and browse to locate the XML files listed above, in the folder **Tutorial II** located in the **Tutorial data files** folder.



- 2 Select all four XML files by holding down the shift key and clicking on each image using the left mouse button then click **Open**. The XML and image files will automatically be loaded into the BVA workspace.

The BVA interface consists of four different modes: Spot Map Table (**ST**), Match Table (**MT**), Protein Table (**PT**) and the Appearance Table (**AT**). Each screen in the BVA interface possesses a similar layout, consisting of:

- **Image View:** Displays the gel images making up the experimental set, with the currently selected spot shown in magenta.
- **Table View:** Depending on the screen selected, displays either information on the images used in the experiment, different properties for a single spot across the gels in the analysis set or protein spot specific statistical values.
- **3-D View:** Representation of the selected spot in 3 dimensions (not displayed in Spot Map Table mode).

- **Graph View:** Graphical representation of the protein abundance data for a single spot across the different images in the analysis set.

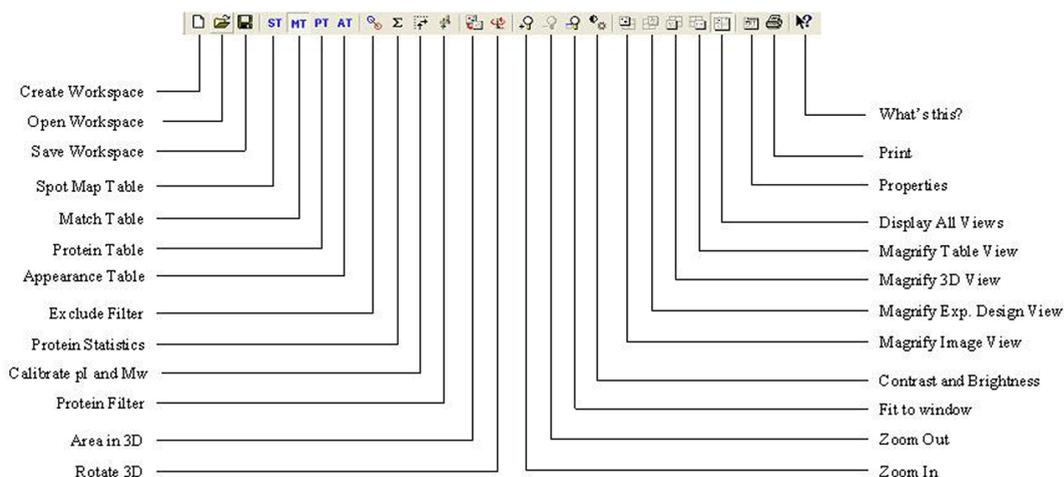
Note: The graph view is only displayed in Protein Table and Appearance Table (AT) mode. In Spot Map Table (ST) mode the graph view is substituted for Experimental Design view (see section 4.5.2).

All the tables in BVA can be ordered by clicking on the headers at the top of each column.

The screenshot displays the BVA software interface with several key components labeled:

- Image View:** Shows two spot maps: 'Primary: Gel 04 Cy3 Contro...' and 'Secondary: Gel 05 Cy5 Tre...'. Both show a grid of green spots.
- 3-D View:** Shows two 3D surface plots of the spot maps, with 'Spot No: 940 Position: 638, 402' and 'Spot No: 798 Position: 682, 377' respectively.
- Graph View:** A line graph titled 'Master No: 973' showing 'Standardized Log Abundance' on the y-axis (ranging from -0.3 to 0.3) and 'Standard', 'Control', and 'Treated' on the x-axis. A red dot is visible at the 'Treated' point.
- Protein Table:** A table titled 'T-test and Av.Ratio: Treated / Control' with columns: Pos., Master No., Status, Protein ID, Protein AC, Appearance, and T-test. Row 7 is highlighted in pink.
- Data Control Panel:** Located at the bottom, it includes fields for 'Protein ID', 'Protein AC', 'Name', 'Comment', 'pl', 'Mw', 'Da', and checkboxes for 'Pick', 'PTM', and 'Protein of Interest'. A 'Confirm' button is also present.

Other than the icons to move between the different screens the main BVA tool bar is identical to the DIA tool bar.



11.5.2 Function assignment

- 1 Ensure that the **Spot Map** table icon is selected. This screen consists of an Image, Experimental Design and Table View. Select **View:Table View** or click on the **Magnify Table View** icon to display the Table View only. The Table View shows the images that have been loaded into the workspace and the number of spots that have been detected on them.



Note: *A set of images from the same gel will have the same number of spots since the DIA detection algorithm is designed to detect the same number of spots on images from the same gel.*

No.	Status	Image	Gel No.	Type	Label	No. of Spots	Matched	Function	Group	Group Description	Condition1	Condition2	Sample ID
1	Pending	Gel 04 Cy2 Standard.gel1		DIGE Min	Cy2	2298	0	A	Standard				
2	Pending	Gel 04 Cy3 Treated.gel1		DIGE Min	Cy3	2298	0	A	Unassigned				
3	Pending	Gel 04 Cy5 Control.gel 1		DIGE Min	Cy5	2298	0	A	Unassigned				
4	Pending	Gel 01 Cy2 Standard.gel2		DIGE Min	Cy2	2254	0	A	Standard				
5	Pending	Gel 01 Cy3 Control.gel 2		DIGE Min	Cy3	2254	0	A	Unassigned				
6	Pending	Gel 01 Cy5 Treated.gel2		DIGE Min	Cy5	2254	0	A	Unassigned				
7	Master	Gel 02 Cy2 Standard.gel3		DIGE Min	Cy2	2572	2572	M, A	Standard				
8	Matched	Gel 02 Cy3 Treated.gel3		DIGE Min	Cy3	2572	2572	A	Unassigned				
9	Matched	Gel 02 Cy5 Control.gel 3		DIGE Min	Cy5	2572	2572	A	Unassigned				
10	Pending	Gel 03 Cy2 Standard.gel 4		DIGE Min	Cy2	2507	0	A	Standard				
11	Pending	Gel 03 Cy3 Control.gel 4		DIGE Min	Cy3	2507	0	A	Unassigned				
12	Pending	Gel 03 Cy5 Treated.gel4		DIGE Min	Cy5	2507	0	A	Unassigned				

- Under the column **Function** all the images will be labelled with the function Analysis (A) and one of the analysis images (the image with the largest number of detected spots) will be labelled with the function Master (M). The analysis function designation indicates that each of the images will be included in post matching statistical analysis. The Master function can be assigned to a different image, by selecting the image to be assigned as master using the left mouse click. Tick the box entitled **Master** at the bottom of the table in the area entitled **Function for Spot Map**.

11.5.3 Experimental group assignment

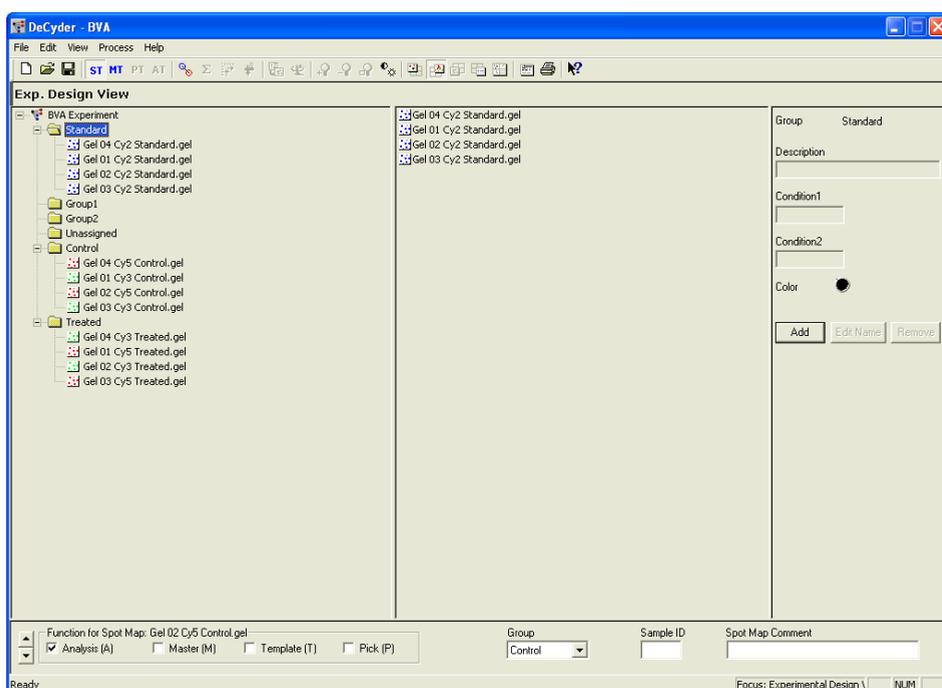
To generate statistical data for the proteins which differ between the control and treated groups it is necessary to identify the different experimental groups within the software. The three groups involved in this experiment are Control, Treated and Standard. Group assignment is performed via the Experimental Design view.



- Select **View:Experimental Design view** or select the **Magnify Experimental Design View** icon.
- The standard images will automatically be assigned to the group Standard by the software, if the word **standard** or **std** appears in the image name. The standard images will therefore be located in the

folder entitled **standard**. By default non-standard spot maps are assigned as **Unassigned**. To assign the control and treated images to their respective groups, these groups must first be created.

To create a control group select **Add** button, then enter **Control** in the **Group** text box and click **Confirm**. The new group folder, **Control** is subsequently displayed. The treated group is similarly generated. Spot maps are assigned to either the control or treated group folders by selecting the **Unassigned** folder, then dragging and dropping spot maps from the center panel of the Experimental Design view to the appropriate folders on the left of the screen.



11.5.4 Landmarking

Landmarking allows the user to manually define matched protein spots in order to improve the accuracy of the gel-to-gel matching process, when the samples are particularly complex or the gels have not run well.

- 1 In some cases images from different gels may vary sufficiently to require landmarks to be set. To set a landmark, select the **Match Table (MT)** by clicking the **MT** icon.
- 2 Click on the **Magnify Image View** icon to display the gel images only. The image on the left represents the Master image while the image



on the right represents one of the images to be matched to the master.

- To begin setting landmarks click on the button entitled **Landmark mode**, at the bottom of the screen (this button will now become yellow). From the **Select Match Gel** scroll down menu (circled in the following image), select the first image.



All the spots should appear red to indicate that they are unmatched. If you select an image and the spots appear green, both images are from the same DIA image analysis and thus landmarks are not necessary, therefore select another image.

- If the spots on the images are not clearly visible it may be necessary to alter the contrast/brightness settings of the images. Click on the **contrast/brightness** icon and alter the position of the bars to alter the contrast and brightness of the images until only the most intense spots are visible.
- To set a landmark click on a clearly defined spot on the master image (the spot boundary should become magenta). Now select the spot on the match image which corresponds to the master image spot, so that this spot becomes magenta, a few seconds later a vector line should appear showing that the spots have been matched and the landmark has been set.
- Repeat this procedure until approximately 10 landmarks have been set. It is recommended that landmarks are evenly distributed across the image as this aids the matching process. After landmarks have

been set on the first image, select the second spot map from the **Select Match Gel** scroll down menu.

The landmarks which were set on the previous match image will be colored yellow on the Master image, thus the same landmarks can be set on the next match standard spot map by finding the spots which correspond to the yellow master image spots.

Note: Landmarks are only required on the standard image from a set of images that have been simultaneously processed in a DIA analysis.

- 7 Repeat this procedure until landmarks have been set on all the standard images from each DIA analysis. When landmarking is completed click on the **Landmark mode** button to exit the landmark mode. Click on the **Match** icon, ensure that **Match all** is selected and click **Match** to commence the matching process.



Note: It is usually only necessary to set landmarks on those images that differ significantly from the Master image. In such cases it may be necessary to set more landmarks. In addition landmarks may have to be set after matching if the number of matched spots is low or if there are a high number of wrongly matched spots, in which case matching will have to be repeated.

11.6 Assessing matching

- 1 Double click on the Match Table icon or select **View:Match Table**. This screen displays data on the gel to gel matching of the different protein spots and allows match confirmation

The match table consists of an Image View, 3-D View and Table View. Each of the views are linked, in such a way that selecting a spot in the Image View will display that spot in 3 dimensions and highlights that spot in the table.

The matches in the Table View are of two types, Auto Level 1 (high probability matches) and Auto Level 2 (lower probability matches). By default the two types of matches are displayed differently in the Image View. Auto level 1 and Auto level 2 spots are represented by a darker and lighter spot boundary, respectively. The boundaries of the unmatched spots are colored red (these colors can be altered by clicking on the properties icon and selecting the colors tab. The Colors window allows the user to define colors to the different matches).



The vector lines in the Image View indicate the positional difference for the same protein spot on different gels. The accuracy of the matching process can now be checked within the match table.

- 2 Order the table by clicking on the column header **Type** so that all the Auto Level 1 matches are at the top of the table.
- 3 To confirm a match in the Match Table View click on the first spot in the table (ensure that it is an actual protein spot). The spot should be selected in the Image View in magenta. Zoom into the area of the selected spot by holding down the left mouse click and drawing a square around the selected area.
- 4 Decide whether the spot in the match image corresponds to the spot on the Master image. Deciding whether a match is accurate can be aided by viewing the selected spot and the surrounding cluster in the 3-D View as well as looking at the matched spots in the Image

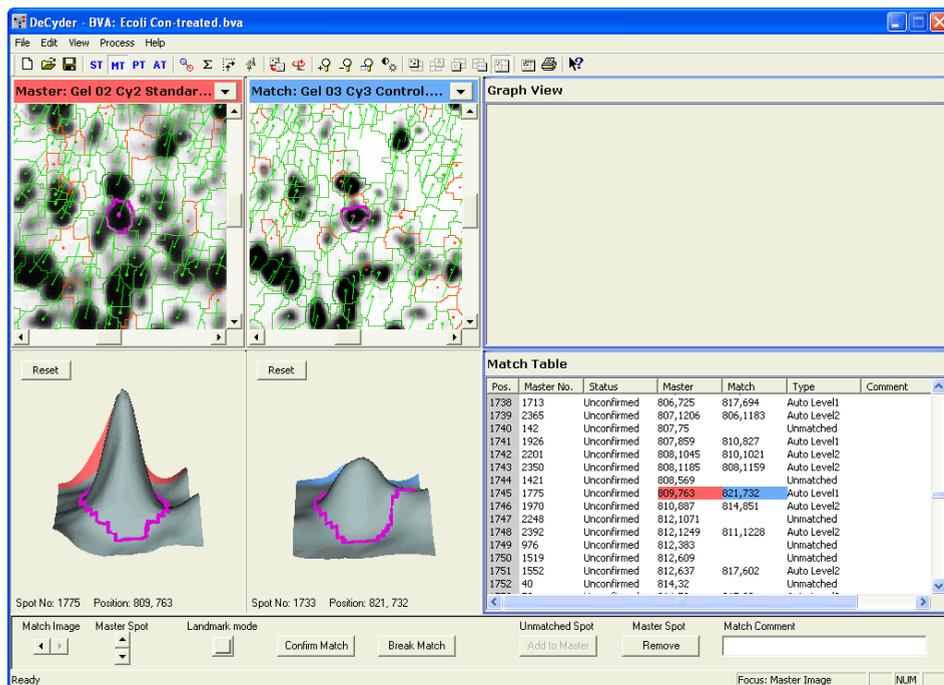
View. If the match is correct, change the Match gel image using the **Select Match Gel** scroll down menu to test the matching accuracy on the other gel images.

- 5 If the match is incorrect, click on the **Break Match** button (at the bottom of the screen), which now becomes **Add Match**. Click the correct spot on the match image.
- 6 If the spot on the match image has also been wrongly matched break this match as before. Select the corresponding spots on the master image and the match image and click on the button entitled **Add Match**.
- 7 This procedure can be carried out, for example, using 5 randomly chosen Level 1 spots and 10 level 2 spots (these spots should be real protein spots). If the matches are accurate in each case move to the **Protein Table (PT)**.

(If the majority of these 15 matches were incorrect, it may be necessary to perform landmarking and a further round of matching.)

The purpose of the above step is to get an overall impression of the matching accuracy. This can be done in several ways. One way is to look at the match vectors for each gel, if the match vectors are not oriented in the same direction over the gel, this may indicate an area of mismatches. For example if in one area of a gel the match vectors are at right angles to the other match vectors on the gel then this area should be landmarked and re-matched.

Rather than choosing a random auto level 1 match and flicking through each gel for that protein on a gel by gel basis, click on random spots on the image, spread around the gel and see if they are correctly matched. If the majority are matched correctly then you can move on to the next gel.



11.7 Post-matching landmarking

If rematching is necessary, a post-matching landmark procedure can be performed.

- 1 The landmarking process here is similar to landmarking before matching. To set the first landmark select the first spot to be set as a landmark on the master image. If the spot selected on the master image is incorrectly matched, click on **Break Match**. Then select the spot on the match image that corresponds to the correct match on the master image. If this spot is also incorrectly matched click on **Break Match**. Now select the two corresponding spots so that both spots become magenta and click the **Add Match** button.

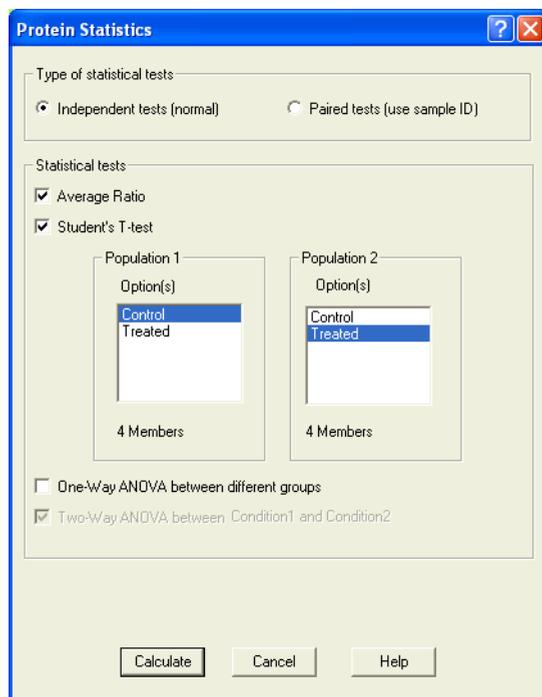
Note: It is only necessary to set landmarks on those images which are poorly matched. The number of landmarks set depends on the accuracy of matching.

11.8 Statistical analysis

- 1 Click on the **Protein Table** icon or select **View:Protein Table**. The Protein Table lists each of the matched spots, across all the images in the experiment. The Protein Table mode consists of an Image, 3-D, Graph and Table View. Each of the views are linked, in such a way that selecting a spot in the Image View will display that spot in three dimensions, highlight that spot in the table and graphically display the abundance values of that spot across all the images on which the spot occurs.

The Protein Table mode allows you to view those proteins that show a significant difference in expression between the control and the treated group. The statistical level of significance is calculated using the T-test (if more than two groups, the ANOVA test is used).

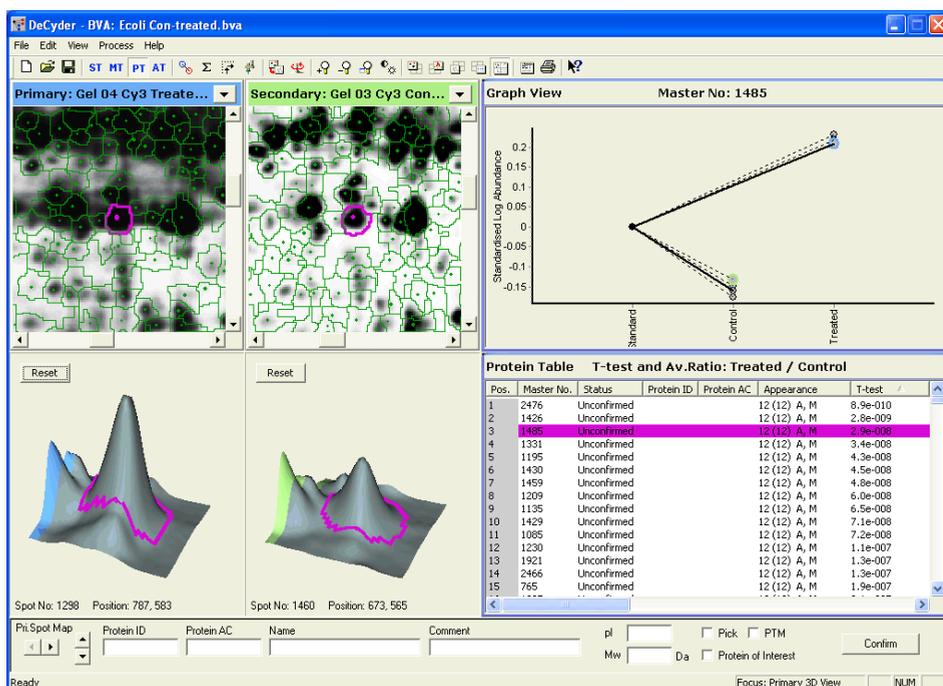
- 2 To perform statistical calculations select **Process:Protein statistics**. In the box entitled population 1 select **Control**, in the box entitled population 2 select **Treated**. Ensure that the **Average Ratio** and **Student's T-test** boxes are selected as below. Click **Calculate**.



The spot specific statistics along with the average spot volume ratio now appear in the table.

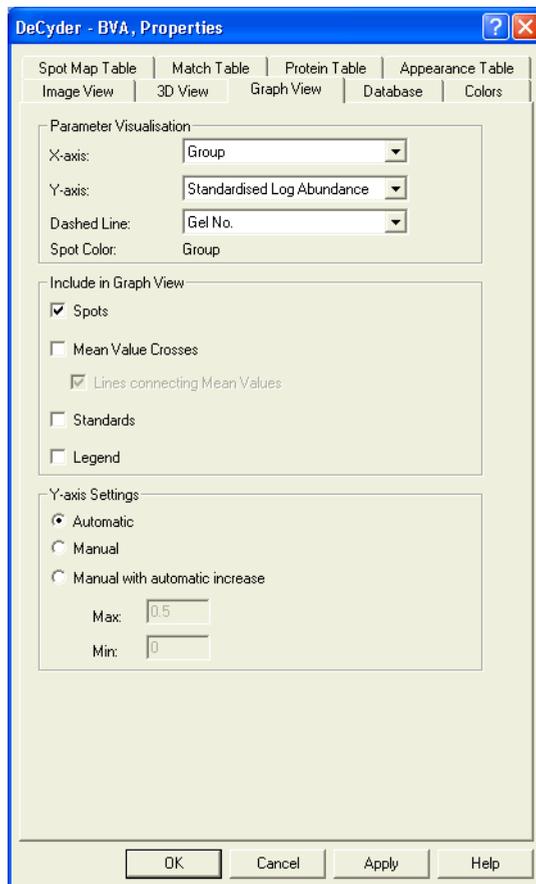
The statistical calculations are performed using the null hypothesis that the differences in protein standardized abundance (i.e. the abundance relative to intra-gel standard and/or biological variation within groups) of control and treated samples is not significant, therefore a low value indicates a high level of significance.

- Click on the **T-test** column header to order the table so that the lowest value is at the top of the table. Click on the spot with the lowest value in the Table View. This spot will now be displayed in all four views.





- 4 To alter the Graph View open the Properties window. Click on the **Graph View** tab and ensure that the view parameters are identical to those shown below. Click **OK**.



In the Graph View, the dashed lines link data points from sample spots to their respective standard data point. Therefore, in this experiment there are four matched control spots and four matched treated spots (associated with four standard data points).

The standards are the same on each gel, and hence the normalization calculates how samples change with respect to their in-gel standard. A ratio of sample to standard of for example 2.5, means that the sample protein is 2.5 times greater than its standard protein, or a ratio of 2.5 to 1. When this methodology is used, all standards are 1. Thus the standard is displayed graphically as 1 and the samples displayed relative to their standard. The log of 1 is zero and therefore when the log standardized abundance is displayed, all standards are zero.

The protein table shows average ratio (i.e. Average abundance of control/Average abundance of treated) as well as the statistical significance of the differences.

11.9 Assigning spots as proteins of interest

11.9.1 Selecting proteins of interest

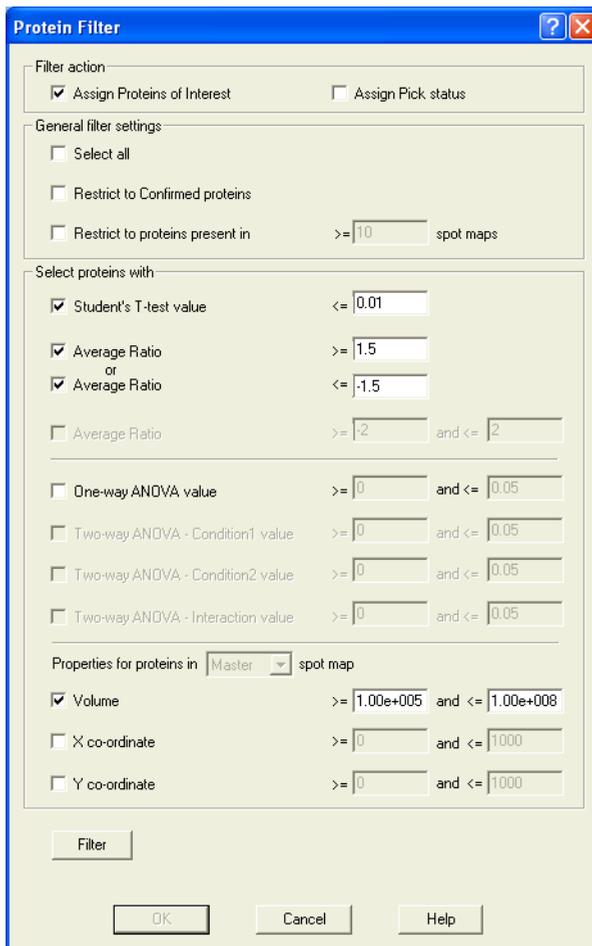
This process can be performed prior to spot confirmation or after spot confirmation. In some ways, doing this first reduces the amount of time spent on spot confirmation.



- 1 To select those proteins that have been found to be statistically different on the analytical gels, select the **Protein Filter Dialog** icon.

The Protein Filter selects protein spots based on user defined criteria such as Student's T-test value, average ratio or volume. The protein filter is therefore used as a filter to find proteins that are interesting which can then be confirmed.

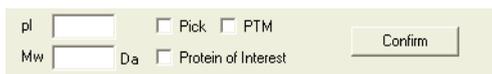
- 2 Ensure that the **Assign Proteins of Interest** check box is selected and that the **Assign Pick status** is not selected. This results in the proteins successfully filtered being given a **Protein of Interest** status. For the purpose of this tutorial, assign for selection protein spots with a T-test score less than 0.01 and with an average ratio greater than or equal to 1.5 or less than or equal to -1.5 and a gel volume between 10^5 and 10^8 . Click on **Filter** and it will tell you how many spots meet the criteria. These values can be adjusted and the filter button clicked again to see how many spots now meet the criteria. The purpose of this filtering is to narrow down the users search for interesting proteins that change significantly. Click **OK**.



Note: The volume filtering criteria is applied to spot volumes on the Master gel. If a preparative gel is assigned the parameters can be applied to this pick gel (see tutorial III).

- 3 The subset of spots that meet the criteria will now be assigned with a protein of interest status in the protein table.

Alternatively, individual spots can be assigned manually by clicking on the spot of interest and clicking in the **Protein of Interest** check box at the bottom of the **Protein Table** screen.



-
- 4 Select **View:Properties** or click on the **Properties** icon. Go to the Protein Table tab and check the **Protein of Interest (I) only** box. Click **OK**. Now in the Protein Table are only those spots that are of most interest and can now be manually confirmed.

11.9.2 Spot confirmation in Protein Table

A spot should be confirmed when a significant difference has been studied, and the user is certain that it is a real protein spot and it has been correctly matched on all gels.

- 1 To confirm the first spot in the **Protein Table**, select the spot in the **Table View**. The spot is automatically selected in all the other views. In the **Secondary image** scroll down menu above the image, select the image that corresponds to the Master image. Zoom into the area of the spot in the Image View as described previously to determine whether the correct spots have been matched correctly between the primary and the secondary images. If this is the case select another image using the **Primary image** scroll down menu and repeat the match assessment.
- 2 If the protein spots are incorrectly matched move to the Match Table View by clicking **MT** and alter the match as described previously.
- 3 When all the matches are correct the protein spot can be confirmed by clicking the **Confirm** button. This records that the spot has been manually checked.

This confirmation should be done for all the significant differences that have been assigned with a protein of interest status. If after manual scrutiny the user decides that the spot should not be selected, deselect the protein of interest status for the given spot by checking the box marked **Protein of Interest** at the bottom of the Protein Table.

11.9.3 Exporting the Pick List and physically excising spots from the gel

To excise the spots from the gel a separate preparative gel should be used. This is matched to the analytical gel set and the pick status is eventually transferred to the relevant matching spots on the preparative gel. The pick list with x, y co-ordinates of spot positions is then exported based on the preparative gel. The pick list along with the physical preparative gel is then taken to Ettan Spot Picker for spot excision and later analysis by mass spectrometry. For a detailed description of the workflow associated with pick list generation, see Chapter 3.

Generating a pick list is covered in more detail in Tutorial III.

12 Tutorial III - Processing the Preparative Gel and Generating a Pick List

12.1 Objective

As can be seen in tutorials I and II, proteins of interest, which warrant identification and further investigation, can be obtained in both the DIA and BVA modules. The Protein Filter tool is applied to the analytical gels to rapidly ascertain the proteins that are affected by the experimental conditions. When using either CyDye DIGE Fluor minimal dyes or CyDye DIGE Fluor saturation dyes, the highlighted proteins of interest have to be matched to a preparative gel, from where the protein spots can be directly excised using Ettan Spot Picker or Ettan Spot Handling Workstation (see Ettan DIGE system user manual). The workflow for generating a pick list from a preparative gel in both CyDye DIGE Fluor dye types is identical. Here in tutorial III a preparative gel for a CyDye DIGE Fluor minimal dye experiment is used to demonstrate the various steps in generating a pick list.

The aim of this experiment is to generate a pick list, on a SYPRO stained gel, relating to the most significant differentially expressed proteins in control and treated bacterial lysates determined from Ettan DIGE system analytical gels.

When using the CyDye DIGE Fluor minimal dyes, a SYPRO stained pick gel is required rather than picking directly from an Ettan DIGE system gel because the CyDye DIGE Fluor minimal dye label only 5% of the protein. This results in the majority of the (unlabelled) protein migrating slightly different and hence not ending at exactly the same place as the (labelled) protein seen on the gel. Therefore a post stain (total stain) is used on a preparative gel for spot excision. The image generated is then matched to the analytical set. The positions of the significant differences are transferred to the SYPRO stained gel and a pick list is created. This pick list, along with the SYPRO stained gel, is transferred to the automated Ettan Spot Picker or Ettan Spot Handling Workstation for spot excision and subsequent digestion, spotting and mass spectrometry analysis.

12.2 Overview

- 1 This Tutorial follows on from Tutorials I and II. The first two tutorials concentrate on finding differences between different biological groups. This tutorial describes how to excise these spots for further protein identification analysis.
- 2 An experimental design is determined as described in tutorial II, with the addition of a separate preparative gel containing two reference markers and stained with SYPRO Ruby.
- 3 The analytical gels are analyzed in DeCyder Differential Analysis Software to ascertain the proteins of interest, as described in Tutorial I.
- 4 The preparative gel is analyzed in DeCyder Differential Analysis Software DIA. The reference markers are assigned in DIA (alternatively this can be done in BVA).
- 5 The preparative gel spot map is exported as an XML file and added to the BVA workspace that contains the analytical gels.
- 6 The preparative gel spot map is landmarked and matched to the master gel of the analytical set.
- 7 The protein of interest assignments are transferred from the analytical set to the preparative gel.
- 8 The proteins of interest are assigned a Pick status, then the pick locations are visually inspected (and edited if necessary).
- 9 A pick list is exported from the preparative gel and the gel with pick list is then taken to Ettan Spot Picker or Ettan Spot Handling Workstation for spot excision (See the relevant User Manual for instructions).

12.3 Experimental design

Ettan DIGE system and DeCyder Differential Analysis Software are used to investigate differential protein expression in experimental groups. These identified proteins can then be picked from a preparative gel for digestion and subsequent mass spectrometry analysis.

Gels 1 to 4 were generated to ascertain those proteins that were differentially expressed in control and treated bacterial samples. These gels have been pre-analyzed in DeCyder Differential Analysis Software and saved as a BVA file.

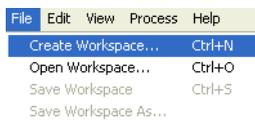
A SYPRO Ruby stained preparative gel (gel 5), loaded with both control and treated samples was prepared for spot picking. The spots on this gel are first detected within the DIA module, then matched to the analytical gels (gels 1 to 4) in DeCyder Differential Analysis Software BVA.

<i>Gel number</i>	<i>Cy2</i>	<i>Cy3</i>	<i>Cy5</i>	<i>SYPRO Ruby</i>
Gel 1	Standard	Control 1	Treated 1	
Gel 2	Standard	Treated 2	Control 2	
Gel 3	Standard	Control 3	Treated 3	
Gel 4	Standard	Treated 4	Control 4	
Gel 5		–	–	Control + Treated

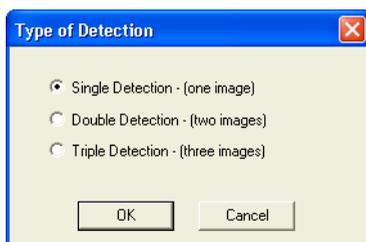
12.3.1 Identifying protein spots on SYPRO Ruby stained gel



- 1 Double click the DeCyder Differential Analysis Software **DIA** icon on your desktop to open the DIA module.
- 2 Select **File:Create Workspace**.

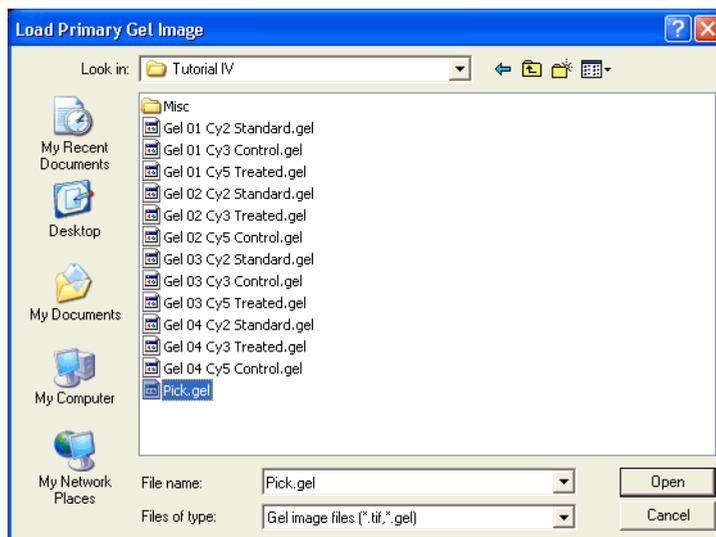


- 3 Select **Single Detection**, then click OK.

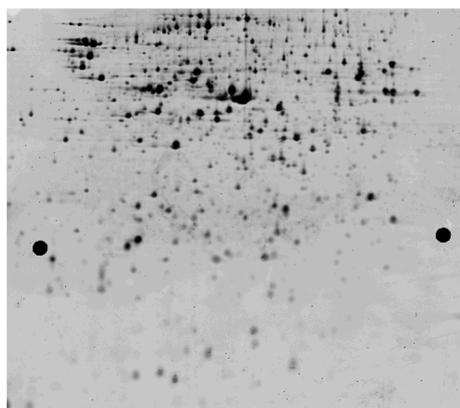


- 4 Browse to locate the folder entitled **Tutorial III** in the **Tutorial data files** folder. Select the SYPRO gel image, then click **Open** to load the

image.



- 5 As the image was stained with the SYPRO Ruby stain the initial image contrast may require adjusting. Select the **contrast/brightness** icon and alter the bars to make the image clearer.



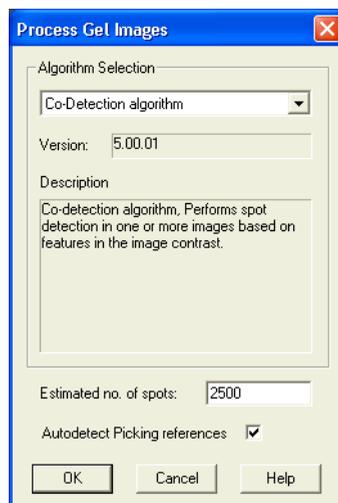
12.3.2 Spot detection

- 1 After the image has been loaded select **Process:Process Gel Images**.

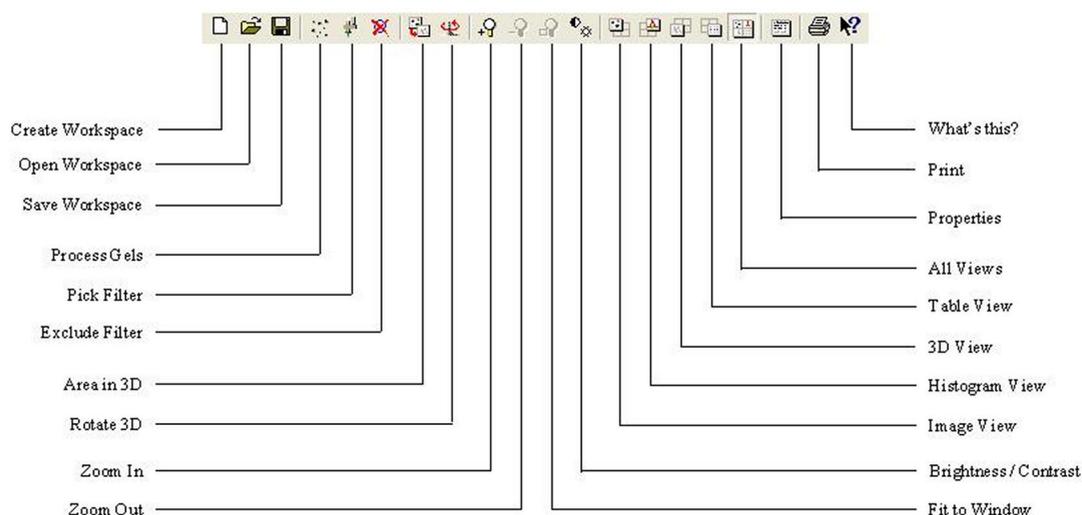


- 2 In the Process Gel Images dialog box which now appears enter the value **2500** for the estimated number of spots (see section 3.3.1).

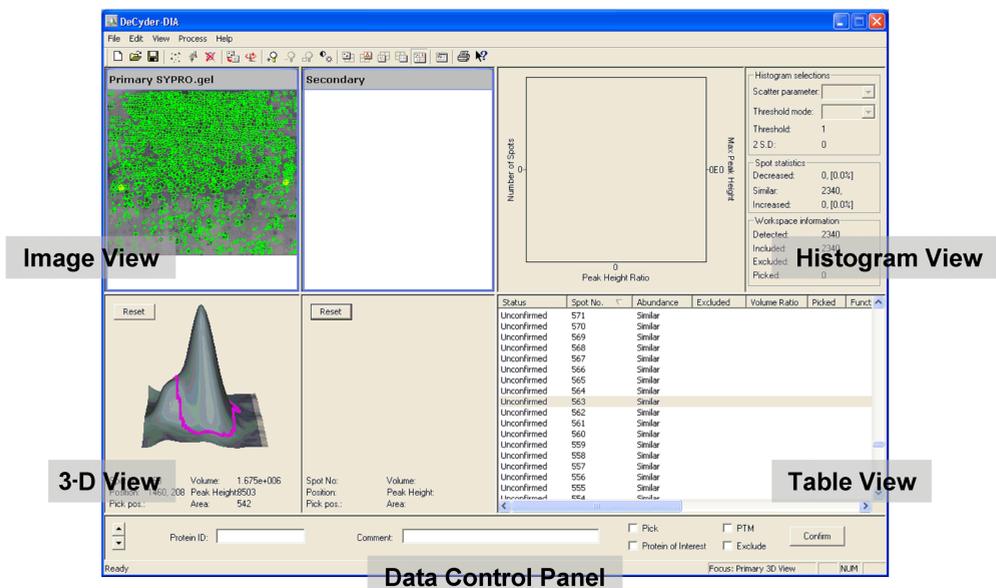
Select the **Autodetect Picking references** check box. Click **OK** to begin spot detection on the images.



Before going any further, it is important to become acquainted with the main tool bar.



The user interface is divided into four main windows, the Image View, the Histogram View, the 3-D View, and the Table View.



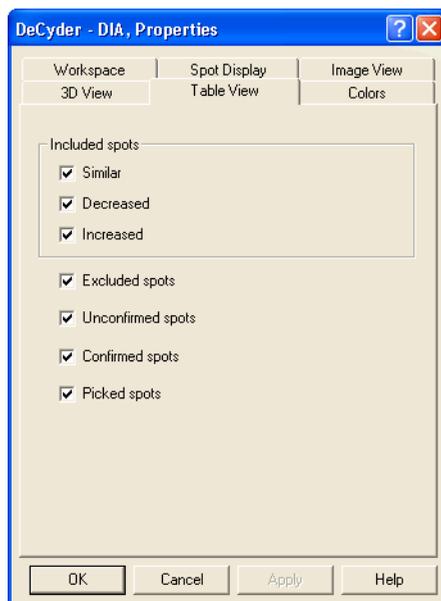
The four views are all linked. Clicking on a spot on Image View highlights the spot in magenta in the Histogram View (the Histogram View is disabled when analyzing single images such as SYPRO stained gel images). The spot is also represented three dimensionally and is highlighted in gray in the Table View.



- 3 To display the entire gel image, click on the **Fit to window** icon.

The detected spots are of three types: increased, decreased and unchanged in expression (colored blue, red and green, respectively, in the Histogram View).

- 4 If, after detection (the hourglass turns back to a cursor) the Table View remains empty, click on the **Properties** icon to bring up the Properties dialog box. By default this will open on the Workspace tab.
- 5 Change to the table options by clicking on the **Table View** tab.
- 6 By selecting and deselecting the check boxes the various categories of spots can be selectively displayed. Deselecting all the check boxes results in the Table View being blank. Similarly, spots can be selectively displayed in the image view using the Spot Display tab. The Table View tab works on OR logic, i.e a spot only needs to conform to one criterion to be shown in the table.



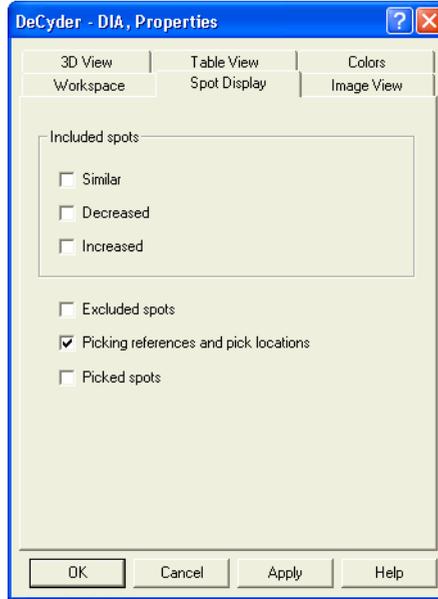
Selecting different protein spots on the images shows that some of the spots that have been detected are in fact either dust particles or artifacts from the gel. To remove these artifacts an Exclude Filter must be used on the images.

12.3.3 Assigning an area of interest

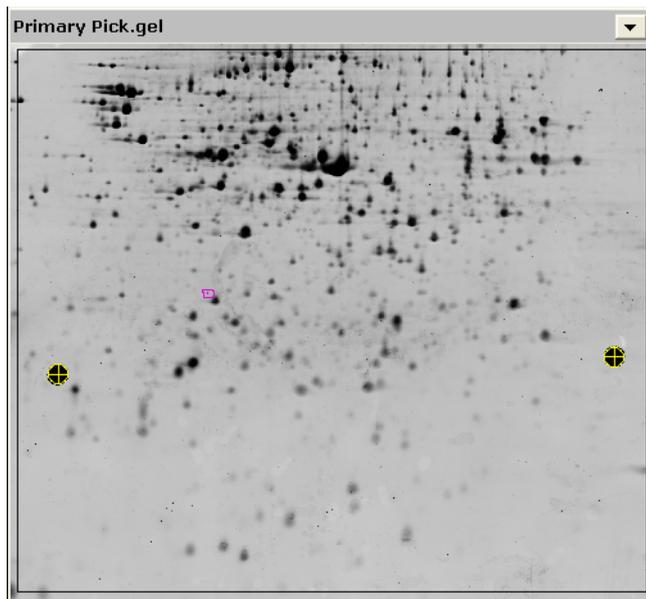
Due to gel heterogeneity at the edges of the images, there are artifacts that need to be removed. These can be removed by setting an area of interest. This function only works with the filter. If an area of interest was set before detection all the spots on the image will still be detected (even those outside the area of interest).

- 1 Click on the **Fit to window** icon on the tool bar to fit the image to the Image View.
- 2 Click on the **Image view** icon on the tool bar to have a full screen view of the gel image.
- 3 Next, click on the **Properties** icon to bring up the properties window.

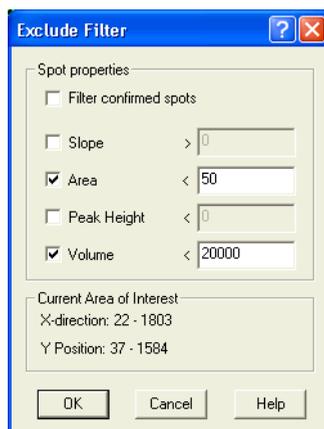
- 4 Select the **Spot Display** tab and deselect **Similar**, **Increased** and **Decreased**, so that the check boxes are identical to those shown in the figure below. Click **OK**.



- 5 To set an area of interest select **Edit:Define area of interest**. Using the rectangular target pointer which now appears drag the mouse to draw a rectangle around the gel, ensuring that edge artifacts are removed. It does not matter if the Reference Markers are inside or outside the Area of Interest as these are set manually.



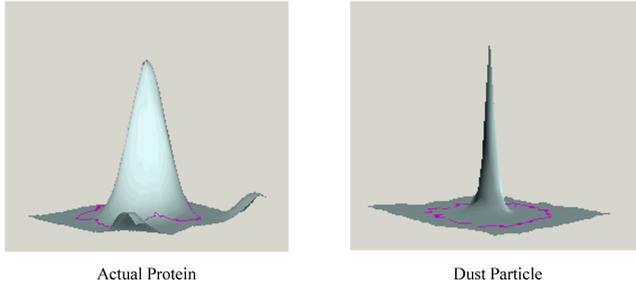
- 6 Click on the **Properties** icon to bring up the properties window. Select the **Spot Display** tab, reselect the **Similar, Increased** and **Decreased** options and click **OK**.
- 7 Select **Process:Exclude Filter** and use the same parameters to filter the images again.



All the spots on the outside of the area of interest will automatically be removed.

12.3.4 Gel artifact removal

A 3-D View clearly shows if a detected spot is a gel artifact rather than a protein spot due to the very steep sides and pointed top of an artifact compared to the smooth curve of a protein spot.

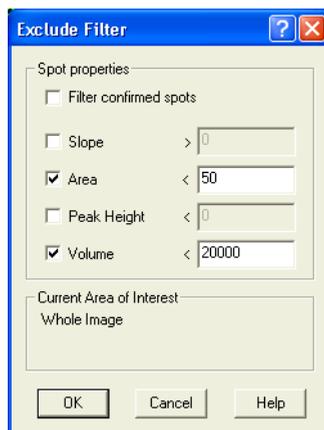


In order to exclude these artifacts from subsequent analysis a set of Exclude filter parameters must be generated.

- 1 In the Table View click on the column header **Area** until the spot with the smallest area is present at the top of the table.

Max Slope	Area	Max Peak...	Max Volume
-----------	------	-------------	------------

- 2 Scroll down the table from one spot to another until the next spot in the table is an actual protein spot and not a dust particle or gel artifact. Make a note of the Area value of the artifact immediately before the real protein spot. This value will be used for the Area category in the Exclude filter.
- 3 Repeat this procedure for **Volume**.
- 4 When both values have been found, select **Process:Exclude Filter** to display the Exclude Filter dialog box. Select the **Area** and **Volume** check boxes and enter the found values. Click **OK**.



An example of a non-stringent set of filter parameters that can be used to run a light filter is as follows:

Area	50
Volume	20000

- 5 Select the **All views** icon to display all four views.
- 6 Export the data from the DIA file (as an XML file) for gel to gel matching in BVA by selecting **File:Export Spot maps**.
- 7 Save the exported file in the folder entitled **Tutorial III** in the **Tutorial data files** folder.

At this point the DIA workspace can be closed (without saving the workspace).

12.4 Matching to analytical gels

The SYPRO Ruby stained picking gel has to be matched to analytical gels in order to identify the proteins on the picking gel that show differential expression and hence require picking.

12.4.1 Preparing the workspace



- 1 Double click the **DeCyder Differential Analysis Software BVA** icon on your desktop to open the BVA workspace.
- 2 Select **File:Open workspace** and browse to locate the BVA file entitled **Ecoli Control treated_for picking**, in the folder **Tutorial III** in the **Tutorial data files** folder.
- 3 Double click on the BVA file to open the workspace (this may take a few minutes depending on the specifications of the computer). This BVA file contains the experiment described in the experimental design section of this tutorial. This file has been pre-analyzed in the BVA module and several proteins that demonstrate a statistically significant change upon treatment have been given a **Protein of Interest** status and hence will require picking and subsequent identification.

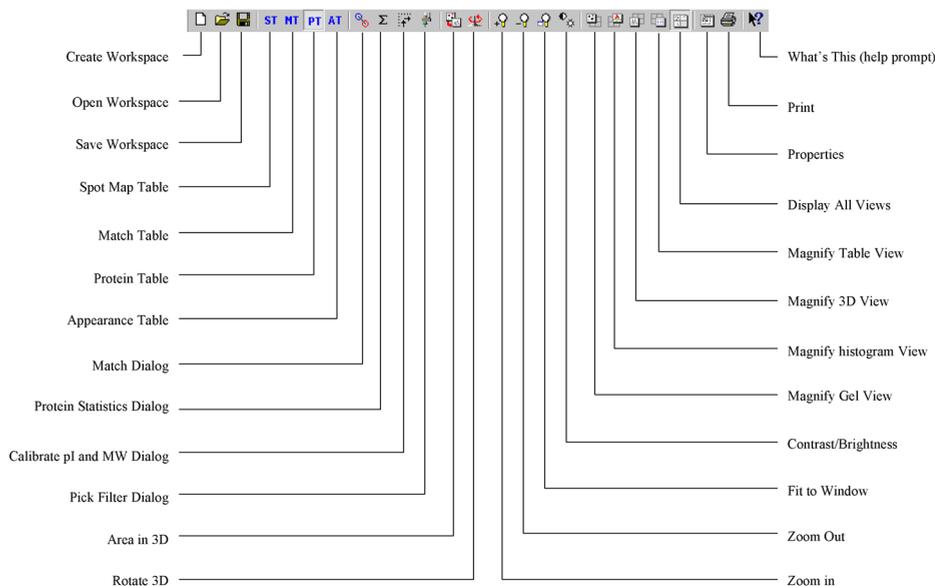
Note: If this tutorial is being performed directly after completing Tutorial II, the created tutorial II workspace can be used instead of the pre-analyzed BVA file.

The BVA interface consists of four different modes, Spot Map Table (ST), Match Table (MT), Protein Table (PT) and the Appearance Table (AT). Each screen in the BVA interface possesses a similar layout, consisting of:

- **Image View:** Displays the gel images making up the experimental set, with the currently selected spot shown in magenta.
- **Table View:** Depending on the screen selected, displays either information on the images used in the experiment, different properties for a single spot across the gels in the analysis set or protein spot specific statistical values.
- **3-D View:** Representation of the selected spot in 3 dimensions.
- **Graph View:** Graphical representation of the protein abundance data for a single spot across the different images in the analysis set.

All the tables in BVA can be ordered by clicking on the headers at the top of each column.

Other than the icons to move from one mode to another the main tool bar is very similar to that seen in DIA.



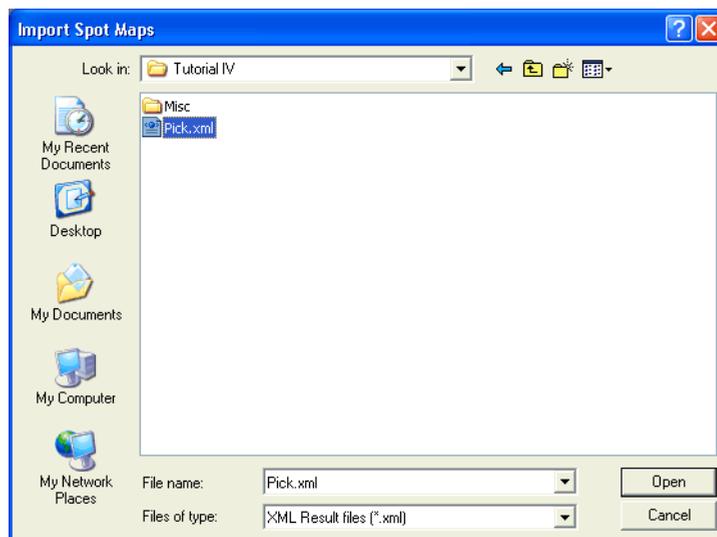
- 4 Ensure that the **Spot Map Table** icon is selected. This screen consists of an Image and a Table View.
- 5 Select **View:Table View** or click on the **Magnify Table View** icon to display the Table View only. The table shows the images that have been loaded into the workspace and the number of spots that have been detected on them.

Spot Map Table													
No.	Status	Image	Gel No.	Type	Label	No. of Spots	Matched	Function	Group	Group Description	Condition1	Condition2	Sample ID
1	Matched	Gel 04 Cy2 Standard.gel1		DIGE Min	Cy2	2298	1081	A	Standard				
2	Matched	Gel 04 Cy3 Treated.gel1		DIGE Min	Cy3	2298	1081	A	Treated	BA treated			
3	Matched	Gel 04 Cy5 Control.gel 1		DIGE Min	Cy5	2298	1081	A	Control	Untreated			
4	Matched	Gel 01 Cy2 Standard.gel2		DIGE Min	Cy2	2254	1130	A	Standard				
5	Matched	Gel 01 Cy3 Control.gel2		DIGE Min	Cy3	2254	1130	A	Control	Untreated			
6	Matched	Gel 01 Cy5 Treated.gel2		DIGE Min	Cy5	2254	1130	A	Treated	BA treated			
7	Master	Gel 02 Cy2 Standard.gel3		DIGE Min	Cy2	2572	2572	M, A	Standard				
8	Matched	Gel 02 Cy3 Treated.gel3		DIGE Min	Cy3	2572	2572	A	Treated	BA treated			
9	Matched	Gel 02 Cy5 Control.gel 3		DIGE Min	Cy5	2572	2572	A	Control	Untreated			
10	Matched	Gel 03 Cy2 Standard.gel4		DIGE Min	Cy2	2507	1780	A	Standard				
11	Matched	Gel 03 Cy3 Control.gel 4		DIGE Min	Cy3	2507	1780	A	Control	Untreated			
12	Matched	Gel 03 Cy5 Treated.gel4		DIGE Min	Cy5	2507	1780	A	Treated	BA treated			

Note: A pair of images from the same gel will have the same number of spots since the DIA detection algorithm is designed to detect the same number of spots on image pairs from the same gel.

Under the column Function all the images will be labelled with the function Analysis (A) and one of the analysis images (the image with the largest number of detected spots) will be labelled with the function Master (M). The analysis function designation indicates that each of the images will be included in the post matching statistical analysis.

- 6 The Master function can be assigned to a different image, by selecting the image to be assigned as master using the left mouse click. Tick the box entitled **Master** at the bottom of the table in the area entitled function for spot map. However, the Master should not be changed here or all the matching will have to be repeated.
- 7 To add the SYPRO stained preparative gel spot map to the BVA workspace select **File:Import Spot Map(s)** and double click on the XML file exported from the DIA workspace of the Preparative gel.



Once loaded, the SYPRO gel spot map can be assigned as a Pick gel, ensuring that the eventual pick list will be based on the spot co-ordinates of this SYPRO gel image. Select the SYPRO spot map in the table view then select the **Pick** check box in the data control panel. De-select the default **Analysis** check box.

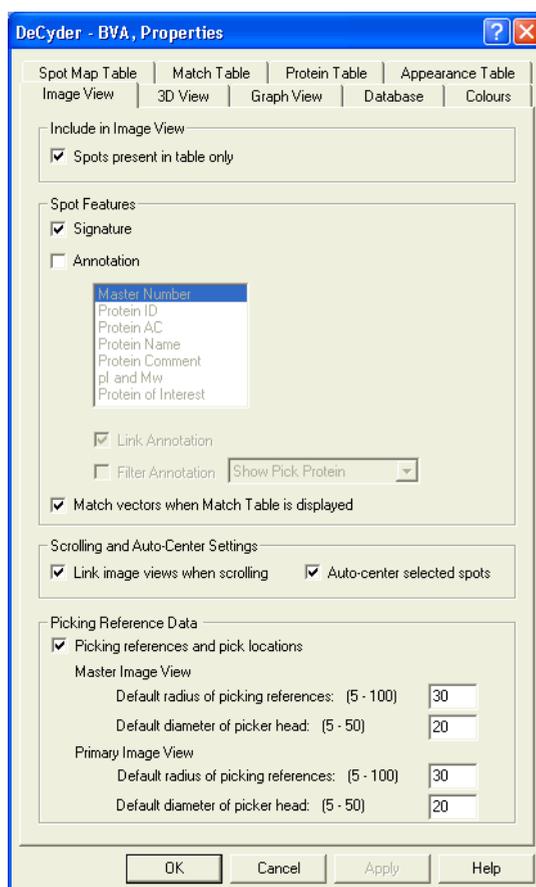


12.4.2 Editing picking reference markers

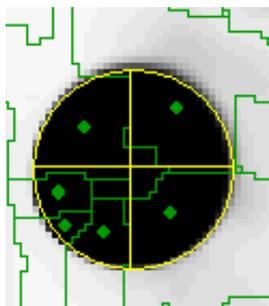
The software, that controls the picking robot requires two reference points in order to locate the spots for picking from the preparative gel. The reference markers placed on the preparative gels (seen as perfect circles on the images) act as these reference points.

The position of the reference markers are automatically detected during the spot detection process. However, it is advisable to review the position of the reference markers and edit them if necessary.

- 1 Zoom into the area of this reference marker by holding down the left mouse click to draw a rectangular area around the marker.
- 2 The size of the target can be altered by clicking on the properties icon and selecting the **Image View** tab. A value of approximately 30 should be entered in the **Primary Image View** box of the picking references area of the window.



- 3 The position of the target can be altered by selecting **Edit>Edit picking Reference**. Place the hand shaped cursor that now appears on the target and hold down the left mouse click. Move the target so that it fits exactly around the reference. Magnifying the area of the gel around the reference marker by selecting the **Zoom in** icon may make accurate alignment easier. After reviewing the left reference marker, zoom into the right reference marker then check and edit its position in an identical manner. Select **Edit>Edit picking References** to exit this mode.



When the picking references have been reviewed it is necessary to match the preparative gel to the master image, thereby matching the preparative image to all the analytical images.

12.4.3 Matching and landmarking



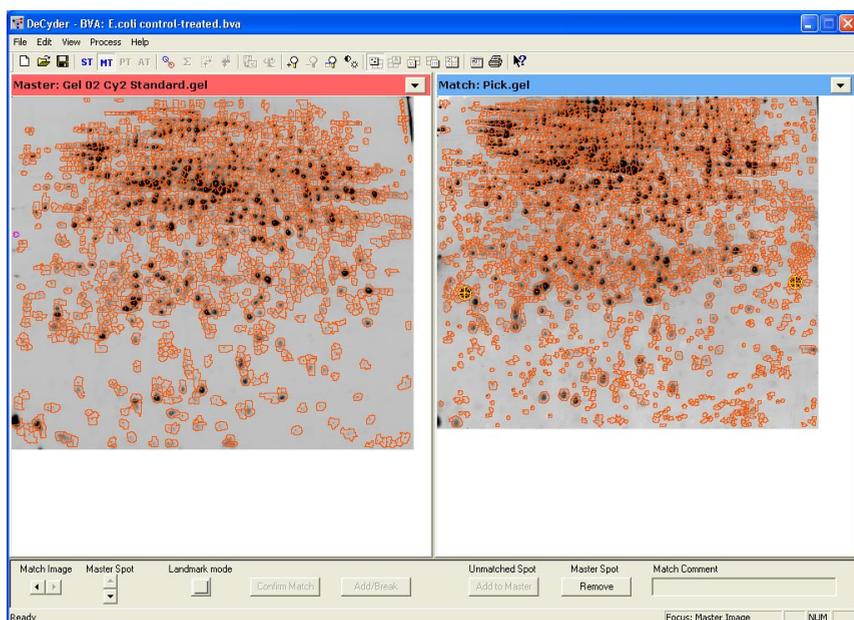
- 1 To match the images first move to the Match Table, by clicking on the **Match Table** icon.

From the images it is clear that the SYPRO image is slightly different from the analytical images. It is therefore necessary to set manual matches known as landmarks. Landmarking allows the user to manually define matched protein spots in order to improve the accuracy of the gel-to-gel matching process.

- 2 Click on the **Magnify Image View** icon to display the gel images only. The image on the left represents the Master image while the image on the right represents one of the images to be matched to the master. Ensure that the image on the right is the SYPRO image using the **Select Match Gel** scroll down menu.
- 3 To begin setting landmarks, click the button entitled **Landmark mode**, at the bottom of the screen (this button will now become yellow). All the spots should appear red to indicate that they are unmatched.
- 4 If the spots on the images are not clearly visible it may be necessary to alter the contrast/brightness settings of the images. Click on the

contrast/brightness icon, ensuring the **Apply to all gel images** check box is **not** selected. Alter the position of the bars to alter the contrast and brightness of the images until only the most intense spots are visible.

- To set a landmark, click on a clearly defined spot on the master image (the spot boundary should become magenta). Now select the spot on the match image which corresponds to the master image spot, so that this spot becomes magenta. A few seconds later a vector line should appear showing that the spots have been matched and the landmark has been set.



- Set approximately 10–20 landmarks on the images, spread evenly across the images. As the images are slightly different it will be necessary to scroll the two images separately. This can be done by clicking on the **Properties** icon and selecting the **Image View** tab. Deselect the **Link image views when scrolling** option.

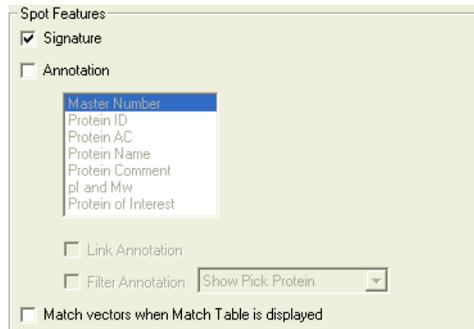


- After landmarking click on the **Match** icon and select the **Match Pending and Landmarked** option in the window which now appears.

Click **Match**.



- 8 When matching has completed, the vector lines which appear will be very long due to the difference in size of the analytical and preparative gel. These vector lines can be hidden by clicking on the properties icon and selecting the Image View tab. Deselect the **Match vectors when Match Table is displayed** option.

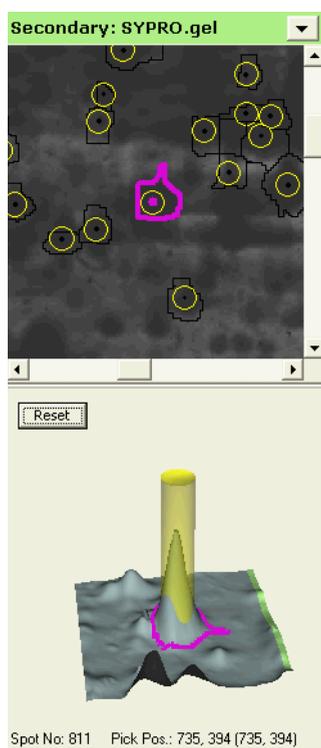


- 9 It is only necessary to determine whether the matches between the master image and the preparative image are accurate. The only matches that need to be confirmed are those that match to proteins assigned with a pick status. Therefore, instead of confirming matches in the match table, they should be confirmed in the protein table.
- 10 Click on the **PT** icon. Go to the **Protein Table** tab of the properties dialog box and select the **Protein of Interest (I) only** check box. Click **OK**.

PT



- 11 In the table only proteins assigned as protein of interest will be present. In the **Image View** have the master gel as the primary image and the preparative gel as the secondary image. Start at the top of the table and click on the first spot, check if this spot is matched to the master, and if so, whether it is matched correctly. If this is the case, select the **Pick** check box in the data control panel. However, if there is an incorrect match, switch to the Match Table and correct the match. Then move back to the protein table and look at the next spot in the list. If the spot is not present on the preparative gel and hence cannot be picked move to the next protein in the table (this protein will not have a **Pick** status). Whilst reviewing the match accuracy of the preparative gel, the pick locations should be verified simultaneously.
- 12 Spots assigned with a **Pick** status have a yellow transparent cylinder and a yellow circle denoting the picking location in the 3-D and image views, respectively. The picking location of all picks is reviewed to assess whether the pick location requires editing (see section 5.6).



-
- 13 To edit a pick location, zoom in on the selected spot in the image view. Select **Edit:Edit Pick Location** then place the hand shaped cursor that appears in the image view over the centre of the pick locations that require editing, then drag the pick circle to the desired location. Select **Edit:Edit Pick Location** to exit this mode. To return pick locations to their original position select **Edit:Restore Default Pick Locations**.

12.4.4 Exporting the Pick List

Now that the SYPRO stained gel has been matched in and the pick proteins have been reviewed, the Pick list has to be exported for use on the appropriate Ettan picking system.

Select **File:Export Picking List from** and select the **Pick Spot Map** to export a pick list. Type in an appropriate file name for the pick list. The pick list can be saved as either a text file or an XML file by appending the appropriate file extension to the file name (*.txt* or *.xml*). The text and XML file are used for Ettan Spot Picker or Ettan Spot Handling Work station, respectively. Click **OK**.

The pick list file and the actual SYPRO Ruby stained gel are taken to the appropriate picking system. The pick list contains the relevant information to instruct the picker to excise the spots. For more details refer to the relevant User Manual.

13 Tutorial IV - Fully automated identification of differentially expressed proteins

13.1 Objective

This tutorial describes how to find statistically significant proteins that are differentially expressed between control and treated groups of bacterial cultures using DeCyder Differential Analysis Software in a fully automated manner. The methodology of this tutorial can be applied to any two groups of protein mixtures with either replicate gels or replicate biological samples. The tutorial outlines the various stages in experimental design, sample organization, protein detection and quantitation in a gel, gel matching and statistical analysis.

13.2 Overview

The following list describes the various stages involved in identifying all the proteins that are differentially expressed in a given system, and which may therefore be worthy of further investigation.

The stages are:

- 1 An experimental design is devised that will generate statistically significant results; a design which minimizes or eliminates in-gel and gel-to-gel system variations.
- 2 Eight sample lysates that form the basis of the experiment are prepared. This consists of four lysates derived from four bacterial cultures treated with benzoic acid and four control flasks that have not been treated. Since the gel to gel variation in this system is low, gel replicates are not compulsory if biological replicates are available.
- 3 Aliquots from each sample are taken and pooled to prepare a standard sample. All three sample types (standard, control and treated) are labelled with an appropriate CyDye DIGE Fluor minimal dye (Cy2, Cy3 or Cy5 dye) as described in the table overleaf. Each sample is then applied to the gels.
- 4 The samples are applied to the four replicate Ettan DIGE system gels, and the gels are then run. Three images are produced from each gel, by scanning at appropriate wavelengths for the three fluors. The recommended scanning instrument for this is the Typhoon Variable Mode 9000 series Imager.

-
- 5 Two representative images from a single gel are initially processed using DeCyder Differential Analysis Software DIA (Differential In-gel Analysis) module. The DIA module performs spot detection using all images loaded, generating spot information for all images simultaneously (in this case standard, control and treated images).
 - 6 The data is examined to identify appropriate values for a number of variables that are set in the Exclude filter (or artifact filter). These values are automatically applied later in the experiment to remove all non-protein spots from all the gel images.
 - 7 The Batch Processor is run to perform a fully automated comparative analysis of the control and treated samples. This analysis involves the following steps:
 - Running a series of four sequential DIA analyses (for each of the four gels).
 - Performing spot detection and calculation of the spot properties.
 - Removing non-protein spots using the previously ascertained exclusion filter parameters.
 - Normalization of the protein spots using the in-gel linked internal standard.
 - Loading of the derived spot maps and associated data into the BVA module.
 - Inter-gel matching of spot maps.
 - T-test analysis of proteins in control and treated samples.
 - Generation of a pick list for Ettan Spot Picker.

13.3 Experimental design

In order to generate statistically valid data, a minimum of three replicates must be used. These can be either biological replicates from each group or, if no biological replicates are available then replicate gels must be employed. For biologically relevant answers to a hypothesis, it is strongly advised that biological replicates of each group are used. The greater the number of biological replicates, the more biological variation is taken into consideration and therefore, the more biologically relevant the results are to the system being investigated. For the purpose of this tutorial, we describe the analysis of four replicate gels loaded with bacterial lysates as indicated in the following table.

Each gel contains a standard sample to allow normalization of all control and treated samples. The Standard sample is obtained by mixing aliquots of control and treated lysates, pooled in equal concentration.

Gel Number	Cy2	Cy3	Cy5	SYPRO Ruby
Gel 1	Standard	Control 1	Treated 1	
Gel 2	Standard	Treated 2	Control 2	
Gel 3	Standard	Control 3	Treated 3	
Gel 4	Standard	Treated 4	Control 4	
Gel 5				Control+Treated

The data analysis is performed in two parts, using the various DeCyder Differential Analysis Software modules:

- Determining the protein spot filtering parameters.
- Fully automated multi-gel processing using the Batch Processor module.

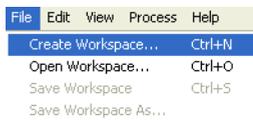
13.4 Protein spot filtering (optional)

In order to identify and exclude artifacts from subsequent analyses, the DIA module is initially run on a representative pair of images from the same gel. This is performed using the DeCyder Differential Analysis Software DIA (Differential In-gel Analysis) module, which identifies protein spots and allows subsequent filtering to remove dust particles and other gel artifacts. The user can examine the results, and specify a series of parameters that will automatically filter all spots that fail to meet the set parameters. This step is optional. Analyses can be performed without filtering workspaces to remove dust particles.

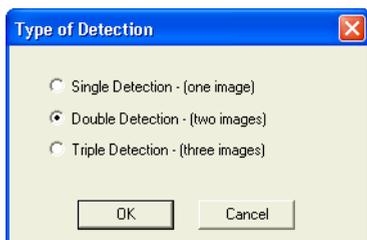
It is possible to omit filtering because a dust particle on a specific gel is very unlikely to be in the same position in other gels and therefore will not be matched and will not be in the final analysis. However, filtering does “clean up” the gel and may make the matching algorithm faster with possible greater accuracy.

13.4.1 Selecting gel images

- 1 Double click the DeCyder Differential Analysis Software **DIA** icon on your desktop to open the DIA module. When the user interface appears, select **File:Create Workspace**.



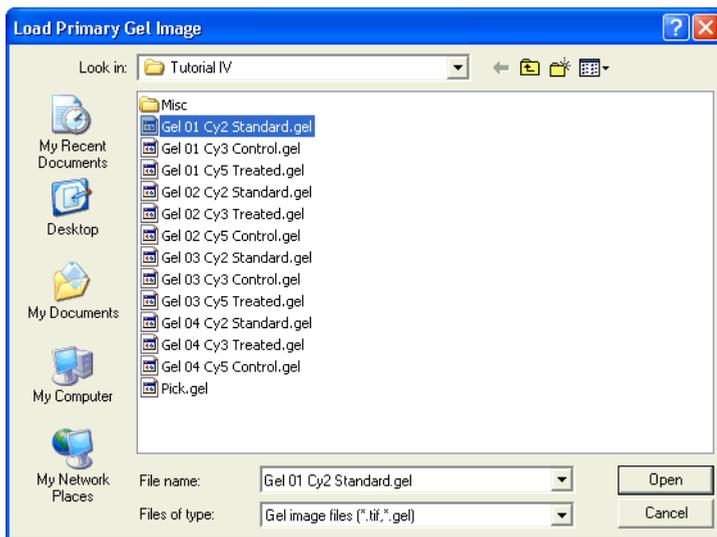
- 2 Select **Double Detection**, then click OK.



- 3 Browse to locate the folder **Tutorial IV** in the **Tutorial data files** folder.

*Note: All the images are named in the style **Gel OX Cy X Standard, Control or Treated.***

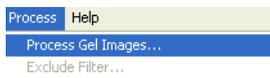
- 4 Double click on image **Gel 01 Cy2 Standard** to load the **Primary Gel Image**.



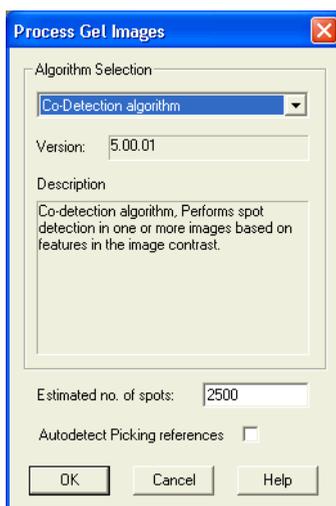
-
- 5 When the **Load Secondary Gel Image** window appears, double click on the image **Gel 01 Cy3 Control**. At this point the two images are displayed in the top left area of the screen. Having loaded the images, the next stage is to perform Spot Detection.

13.4.2 Spot detection

- 1 After the images have been loaded select **Process:Process Gel Images**.

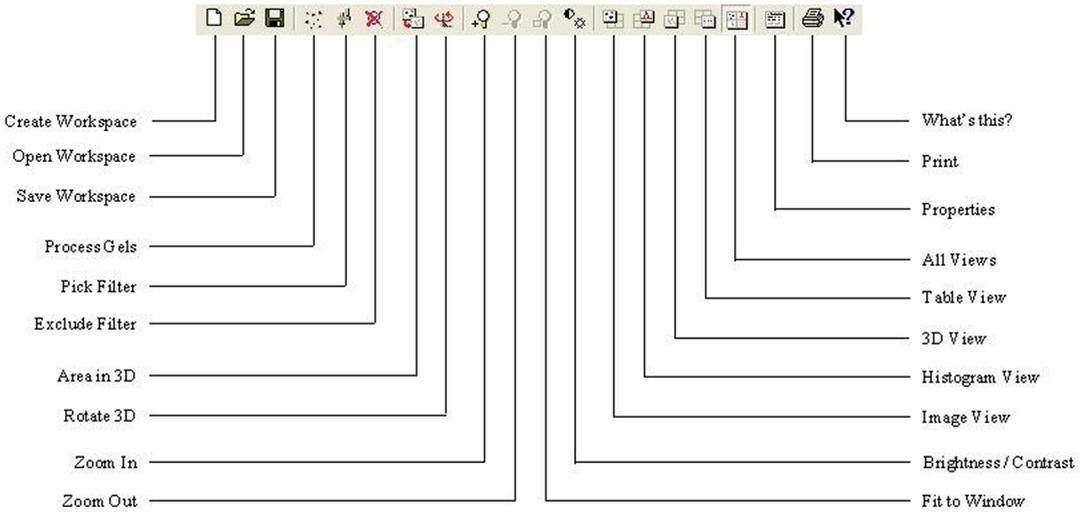


- 2 When the **Process Gel Images** menu option is selected, the following dialog box appears:



- 3 Enter the value **2500** for the Estimated no. of spots (see help file for an explanation of this value). Then click **OK** to begin spot detection on the images. The software then processes the two images and calculates a series of values for each spot on the two images. The DIA analysis takes between 1 and 5 minutes, depending on the specifications of the computer.

Before going any further, it is important to become acquainted with the main tool bar.



4 When the DIA analysis has finished, the screen below appears.

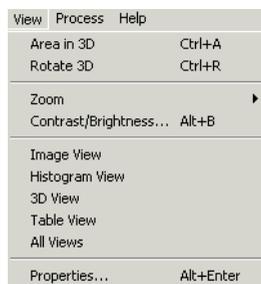
The screenshot shows the DeCyder software interface with the following components:

- Image View:** Displays two gel images: "Primary Gel 02 Control Cy3..." and "Secondary Gel 02 Treated...".
- Histogram View:** A scatter plot of "Number of Spots" vs "Log Volume Ratio". It includes a "Histogram selections" panel with parameters like "Scatter parameter: Max Volur", "Threshold mode: 2 model 5", "Threshold: 2.41551", and "2 S.D.: 2.32953".
- 3-D View:** Shows two 3D surface plots of spot volumes. The left plot has a volume of $3.896e+005$ and the right plot has a volume of $1.143e+006$.
- Table View:** A data table with columns: Status, Spot No., Abundance, Excluded, Volume Ratio, Picked, and Funct. The table lists various spots with their respective abundance and volume ratios.
- Data Control Panel:** Located at the bottom, it includes fields for "Protein ID:", "Comment:", and buttons for "Pick", "Protein of Interest", "PTM", "Exclude", and "Confirm".

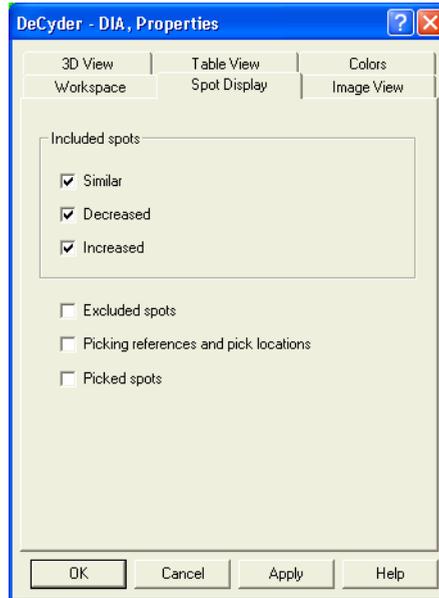
The user interface is divided into four main windows which show the Image View, the Histogram View, the 3-D View and the Table View.

The four views are linked. Clicking on any spot on the Image View highlights the spot in magenta in the Histogram View (the Histogram View is disabled when analyzing single images such as SYPRO stained gel images). The spot is also represented three dimensionally and is highlighted in gray in the Table View.

Each of the four views can be expanded to fill the whole screen. To do this, click on the **View** button and the menu below is displayed.



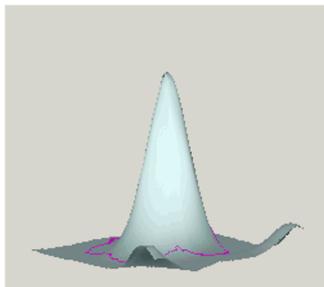
- 5 Click on any of the four views available to expand that view to fill the screen. When you have finished, click on **View** and select **All Views** from the menu above to return to the **All views** display. Alternatively, use the icons displayed on the main toolbar.
- 6 If, after detection has been completed (the hour glass turns back to a cursor), the Table View remains empty, click on the **Properties** icon to bring up the Properties dialog box. Alternatively click on **View:Properties**. By default this will open on the Workspace tab.
- 7 Change to the Table View options by clicking on the **Table View** tab.
- 8 By selecting and deselecting the check boxes the various categories of spots can be selectively displayed. Deselecting all the check boxes results in the Table View being blank. Similarly, spots can be selectively displayed in the Image View using the **Spot Display** tab. The Table View tab works on OR logic, i.e. a spot only needs to conform to one criterion to be shown in the table. Select the **Similar**, **Decreased** and **Increased** check boxes, as indicated on next page.



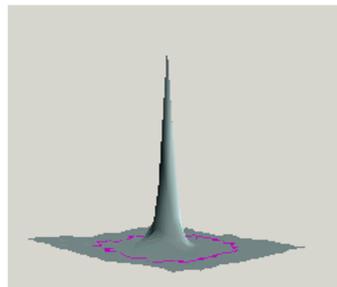
Selecting different protein spots on the images shows that some of the spots that have been detected are in fact either dust particles or artifacts from the gel. To remove these artifacts an **Exclude Filter** must be used on the images.

13.4.3 Gel artifact removal

A 3-D View clearly shows if a detected spot is a gel artifact rather than a protein spot due to the very steep sides and a pointed top of an artifact compared to the much smoother curve of a protein spot.



Actual Protein



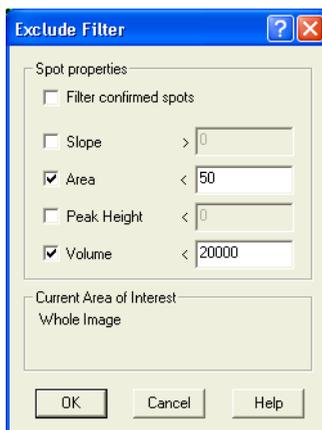
Dust Particle

In order to exclude these artifacts from subsequent analysis a set of Exclude Filter parameters must be generated. This is the main reason for running the DIA module at this point.

- 1 In the Table View click on the column header **Area** until the spot with the smallest area is present at the top of the table.

Max Slope	Area	Max Peak...	Max Volume
-----------	------	-------------	------------

- 2 Scroll down the table from one spot to another until the next spot in the table is an actual protein spot and not a dust particle or gel artifact. Make a note of the Area value of the artifact immediately before the real protein spot. This value will be used for the Area category in the Exclude filter.
- 3 Repeat this procedure for **Volume**.
- 4 When both values have been determined, select **Process:Exclude Filter** to display the Exclude Filter dialog box. Select the **Area** and **Volume** check boxed and enter the found values. Click **OK**.



An example of a non-stringent set of filter parameters that can be used to run a light filter is as follows:

Area	50
Volume	20000

The four gels used in this experiment were run under identical electrophoretic conditions, therefore the filter parameters generated for a representative pair of images performs equally effectively on the images from the other gels in the experiment. If comparing gels from different runs it may be wise to perform a single DIA analysis for each run and set filter parameters accordingly. Once the gel artifacts have been successfully filtered, note the exclude filter parameters then close the DIA module.

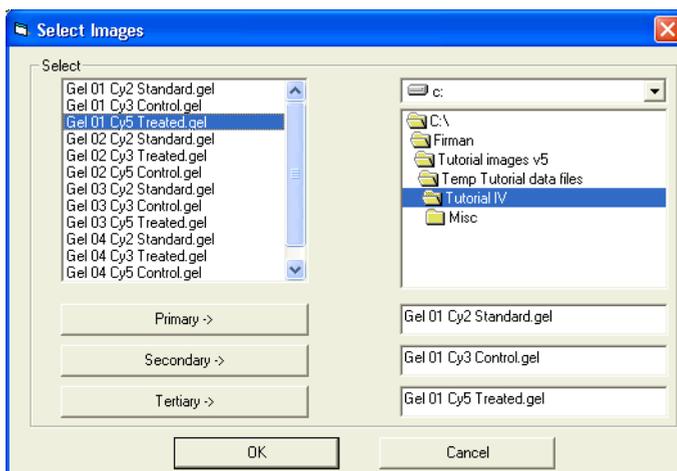
13.5 Processing multiple images

Detection and subsequent matching of protein spots from the images derived from the four gels must now be performed. Rather than analyzing a large number of individual gels using the DIA module, an automated module, the **Batch Processor**, can be used. This automatically runs a series of sequential DIA analyses with no need for user intervention and can be set up to then automatically match the individual spot maps, perform statistical analysis then generate a pick list in the BVA module.

13.5.1 Setting up the batch processor

The first stage of the process is to create a DIA batch list by defining which gel images are to be processed, and in what order. The Primary Image in every case is the specific standard image for that gel, and the secondary and tertiary images are the control and treated sample images (the order is arbitrary).

- 1 Double click on the DeCyder Differential Analysis Software **Batch Processor** icon on the desktop to open the Batch Processor interface.
- 2 Select **File:New Batch**. In the dialog box that now appears browse to locate the folder entitled **Tutorial IV** in the **Tutorial data files** folder.
- 3 Select the image **Gel 01 Cy2 Standard** in the left hand panel, then click the **Primary ->** button. Repeat the procedure for the **Gel 01 Cy3 Control** and **Gel 01 Cy5 Treated** using the **Secondary ->** and **Tertiary ->** buttons, respectively. Click **OK**.



- 4 In the Result Files section of the box that now appears, ensure that the DIA and the XML file names end in the number one (by typing "1" as indicated above), as this enables the subsequent files to be

named automatically with an incrementing number. It is recommended that both files are named “Gel 1”. Click **OK**.

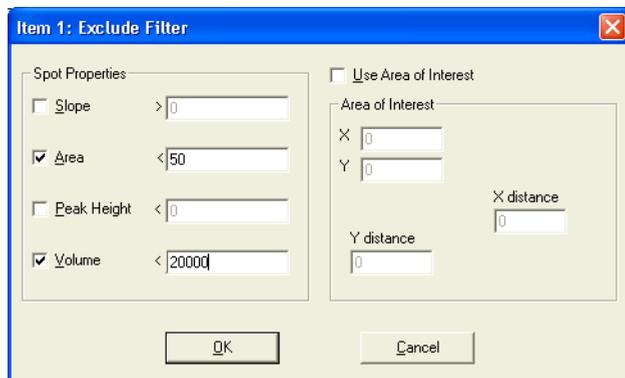
***Note:** If the name does not end in a “1”, then you will be given the opportunity to name the DIA and XML filenames for each batch item.*

The DIA results are generated in two different formats:

- A DIA file, which contains the images and raw data information for each pair analysis.
- An XML file, which contains the raw data file that is subsequently exported to the BVA (Biological Variance Analysis) module.

- 5 Enter the value **2500** for the estimated number of spots and ensure that the **Autodetected Pick Reference** markers check box is deselected and the **Include in BVA batch list** check box is selected. Click **OK**.

- 6 The filter parameters generated previously in DIA can now be entered in the **Exclude filter** dialog box. Place a tick in each of the boxes on the left-hand side of the different filter properties to activate the filter, then enter the filter parameters generated in DIA in their respective boxes. Click **OK**.



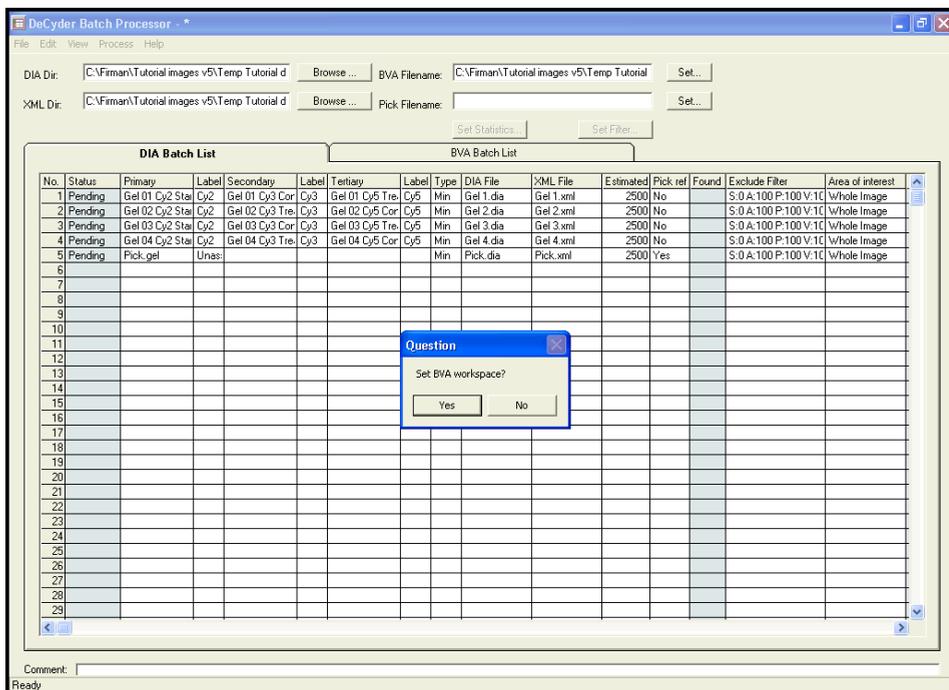
The second set of images associated with the second gel now have to be entered in the batch processor.

- 7 In the dialog box which now appears enter the primary, secondary and tertiary gel images associated with the second gel in an identical manner to the first gel. Click **OK**. The remaining gel processing parameters (such as spot detection and the exclude filter variables entered for the first gel) is automatically applied to the second gel.

Continue loading all the images in this sequence until all the images from all four analytical gels have been loaded.

- 8 The preparative gel used for generating the pick list and subsequent spot excision is conventionally loaded last. Load the “Pick.gel” as a primary image then click **OK**.
- 9 Since the file name contains the text “pick”, the software prompts the user to set this gel as the “pick”, i.e. to use this gel for generating the co-ordinates for the pick list. Select **Yes** in this dialog box.

- 10 When all the images have been loaded click **Cancel**. The batch information is then entered into the DIA batch list and the user is prompted to set the BVA workspace. Click **Yes**.

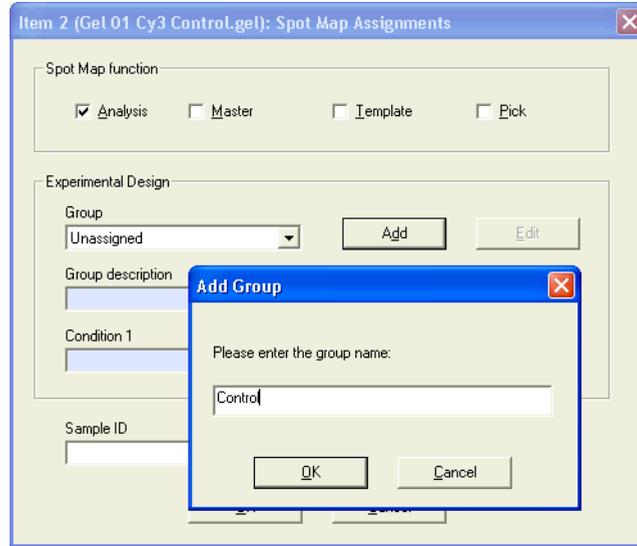


13.5.2 Assignment of spot map attributes

To generate statistical data for the proteins that are expressed differentially between the control and treated groups, it is necessary to identify the three different spot map groups (control, treated and standard) within the software, and to assign every image to one of the three groups. The group names have to be created so that the appropriate images can be assigned to them.

- 1 The **Spot Map Assignment** dialog box, associated with the first spot map, automatically appears after the previous step. The standard spot maps will be automatically assigned to the standard group if the text **Standard** or **Std** is in the image name. It is therefore always important to include the name of the group in the image name. Click **OK** to confirm that the first spot map is an internal spot map.
- 2 The **Spot Map Assignments** dialog box for the second spot map then appears. The second spot map will be derived from either a control or treated sample, denoted in the gel image file name (apparent in the dialog box title bar).

- To assign a spot map to either a control or treated group, the two group names must first be created. Click on the **Add** button in the Experimental Design section of the dialog box and type the name **Control**. Click **OK**. Click the **Add** button again then type the name **Treated**, click **OK**.

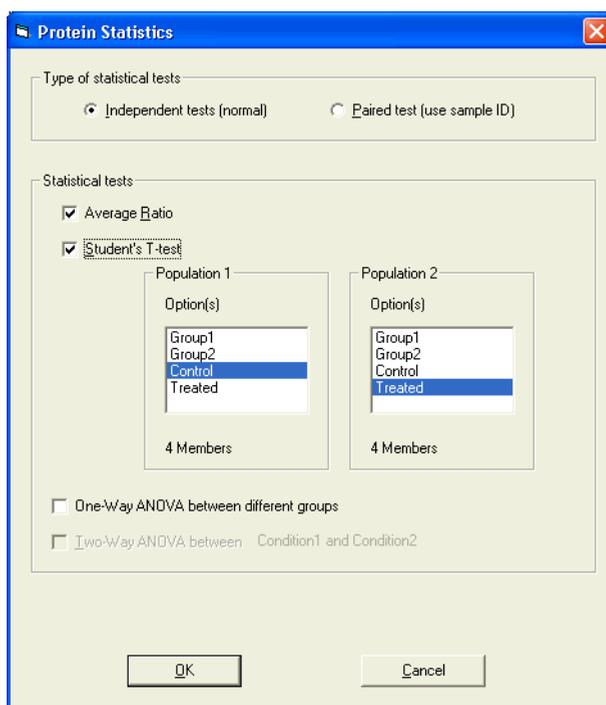


- Once the two experimental groups have been created the remaining analytical spot maps can be assigned to their group, sequentially using the **Group** pull down menu in the **Spot Assignment** dialog box followed by clicking **OK**. The last spot map is the pick gel (item 13). This gel does not belong to an experimental group and can therefore be left **Unassigned**. Deselect the **Analysis** check box in the **Spot Map Function** part of the dialog box. Click **OK** to complete the spot map assignment process.

13.5.3 Protein statistics

Immediately after setting the spot map attributes the **Protein Statistics** dialog box appears automatically. The level of statistical significance in protein differences between the two experimental groups is calculated using the T-test.

In the box entitled **Population 1**, select **Control**, and select **Treated** in the box entitled **Population 2**. Ensure that the **Student's T-test** and **Average Ratio** check boxes are selected and the **Independent tests** option button is selected. Click on **OK**.



13.5.4 Protein Filter

Immediately after setting the protein statistics parameters the **Protein Filter** dialog box automatically appears. The Protein Filter assigns a **Protein of Interest** and/or **Pick status** (dependent on which **Filter Action** check box is selected) to protein spots based on user defined criteria such as Student's T-test value, average ratio or volume.

- 1 Select the **Assign proteins of interest** and **Assign Pick status** check boxes in the **Filter Action** section of the dialog box.

For the purpose of this tutorial, select to assign protein spots that are on 75% of spot maps (i.e. restricted to proteins present in ≥ 9 spot maps) with a T-test score less than 0.01 and with an average ratio greater than or equal to 1.5 or less than or equal to -1.5 and a pick gel volume of 1×10^5 - 1×10^8 the pick gel (setting the parameters to those identical to the figure below). Click **OK**.

The screenshot shows the 'Protein Filter' dialog box with the following settings:

- Filter action:**
 - Assign proteins of Interest
 - Assign Pick status
- General filter settings:**
 - Select all
 - Restrict to proteins present in \geq no of spot maps.
- Select proteins with:**
 - Student's T-test value \leq
 - Average Ratio \geq
 - or
 - Average Ratio \leq
 - Average Ratio \geq and \leq
 - One way ANOVA value \geq and \leq
 - Two way ANOVA - Condition1 value \geq and \leq
 - Two way ANOVA - Condition2 value \geq and \leq
 - Two way ANOVA - Interaction value \geq and \leq
- Properties for proteins in:**
 - Volume \geq and \leq
 - X co-ordinate \geq and \leq
 - Y co-ordinate \geq and \leq

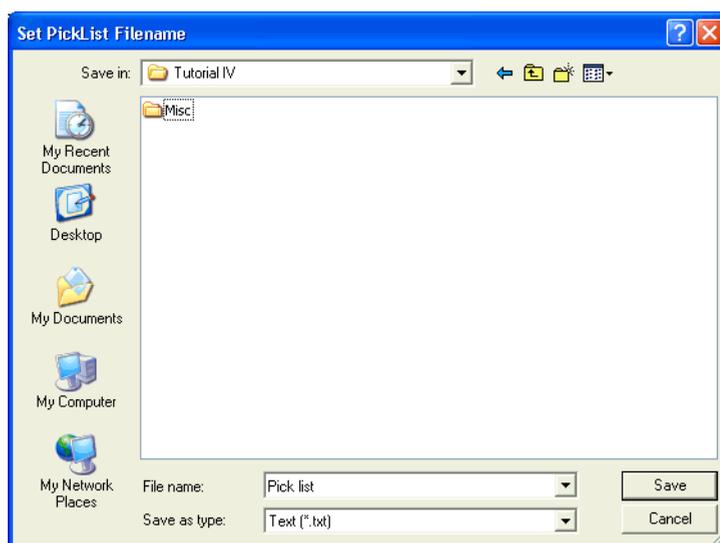
Buttons:

13.5.5 Saving the Pick List

Immediately after setting the protein filter parameters the **Set Pick List File Name** dialog box automatically appears.

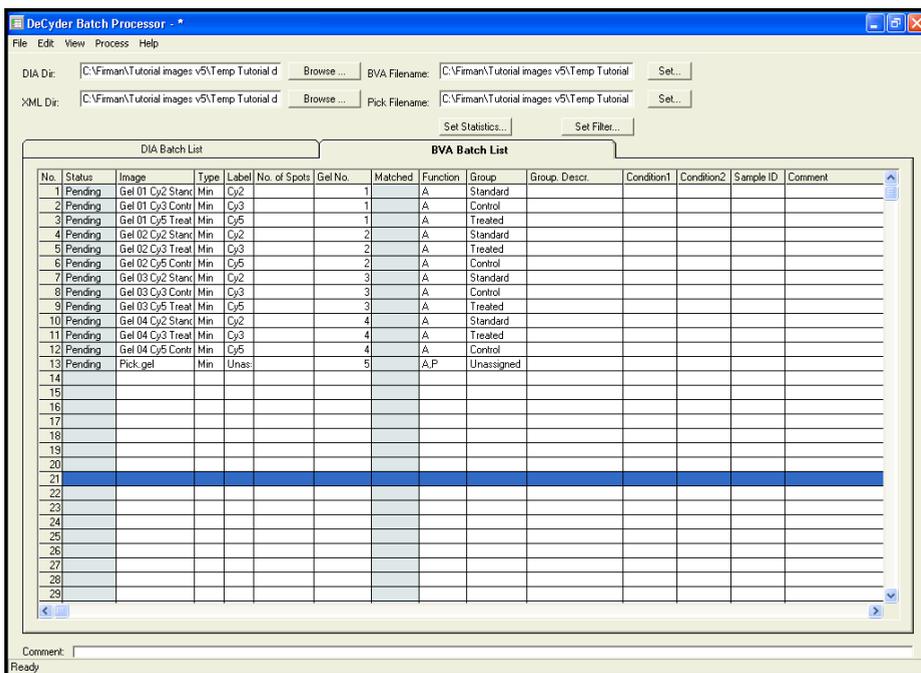
- 1 Browse to select the folder in which to save the pick list (for convenience it may be placed in the folder **Tutorial IV**).
- 2 Enter an appropriate name of the pick list file.

Using the **Save as type** pull-down menu the file extension for the pick list can either be a text or XML format for use with either Ettan Spot Picker or Ettan Spot Handling Workstation, respectively. Click **Save**.



13.5.6 Saving the batch workspace

- 1 To save the batch workspace, select **File:Save As** and enter a name for the workspace. It is useful to enter a name that relates to the experiment. Click **Save**.
- 2 Set the BVA filename by clicking on the button marked **Set** on the side of the BVA filename box (top right of screen). Enter a filename that relates to the experiment (e.g. *E. coli* control-treated). This creates an empty file into which the BVA analysis output will subsequently be entered.



The Batch Processor workspace is now completed and ready to be run

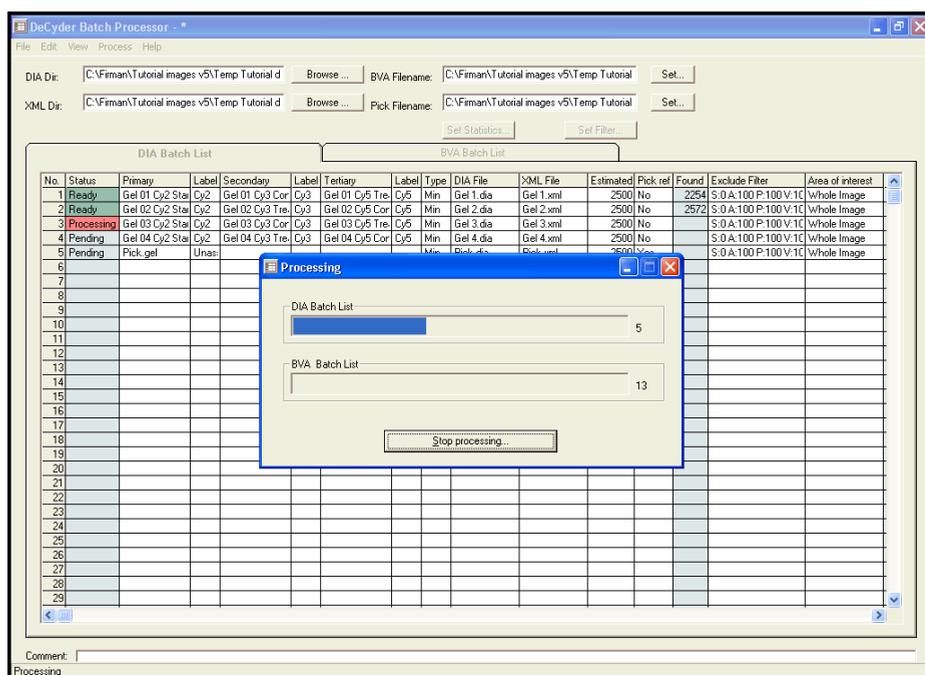
Note: All items in the Batch Processor workspace can be edited by double clicking the relevant cell in the DIA and BVA batch list or selecting the relevant dialog window button on the top of the screen.

13.5.7 Running the batch processor

Select **Process:Run batch**. A prompt appears to request automatic setting of the Master (the spot map that all gels are matched with). Select **Yes** and the gel with largest number of detected spots will be assigned as the Master spot map. Click **OK** in the next dialog box, if you wish to start processing immediately. Alternatively, a user defined delay can be introduced, however the batch processor must be left open during the delay period.

*Note: The Master can also be set when assigning the spot map attributes by selecting the **Master** check box for the desired master spot map.*

The following screen is displayed while the Batch Processor is working.



Upon completion of the batch run the DIA, XML, BVA and pick list files are automatically generated in the requested folders. The preparative gel and pick list can then be taken to the appropriate Ettan spot picking instrument for automated spot excision.

However, it is recommended that the BVA workspace is reviewed prior to generating the pick list. A workflow of using the Batch Processor to assign proteins of interest only is therefore recommended. This involves following the above tutorial but only selecting the **Protein of Interest** check box (without selecting **Pick** status) in the **Protein Filter** dialog box,

then avoiding generating a pick list by cancelling the **Set Pick List Filename** dialog box. In this way the subsequent BVA file will have protein differences highlighted as proteins of interest, which can be manually reviewed then confirmed, before assigning the proteins of interest with a Pick status (see chapter 12).

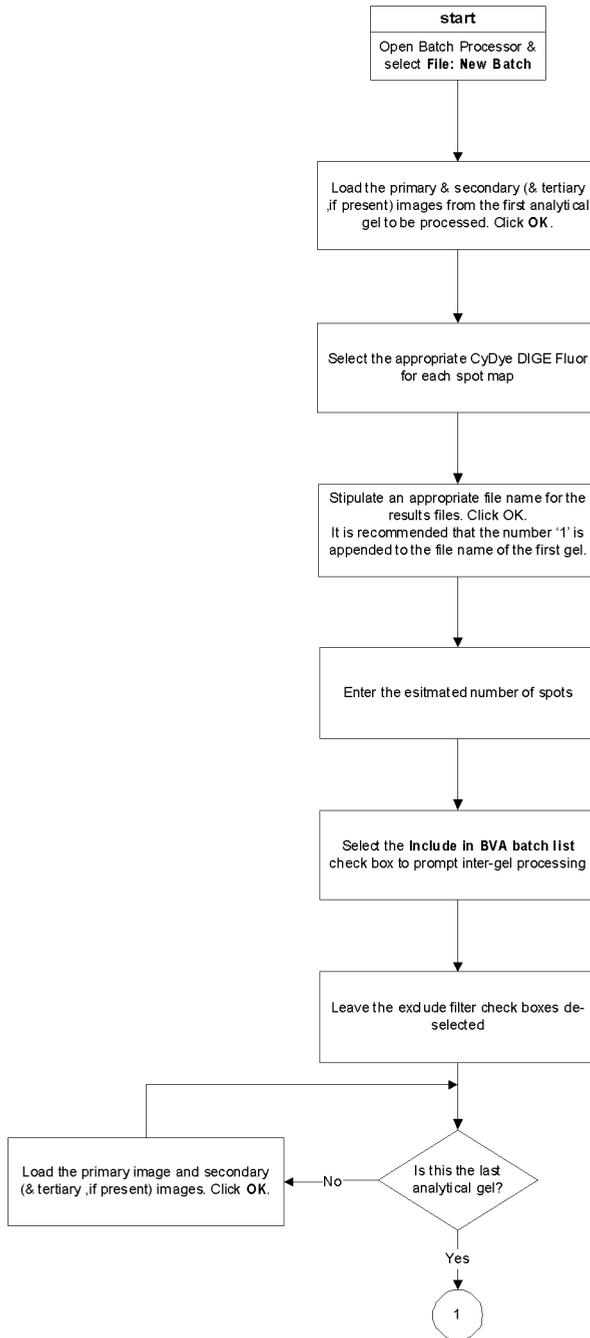
Appendix A Recommended workflow for analysis of multiple gels

The following workflows describe the various recommended steps for processing experiments that employ an internal standard, including data analysis and pick list generation.

The workflow is split into three separate components which represent three phases of the analysis process:

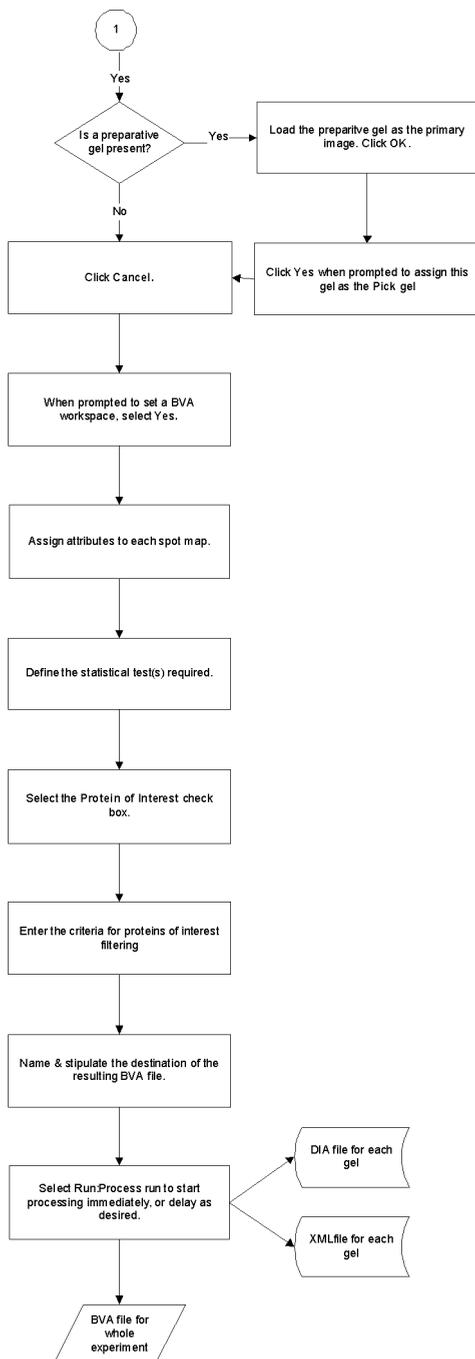
- Flow diagram 1 describes the steps in preparing and processing the gels in the Batch Processor.
- Flow diagram 2 describes the steps in reviewing the proteins of interest in the BVA module.
- Flow diagram 3 describes the steps involved in processing a preparative gel then generating a pick list.

A.1 Flow Diagram 1

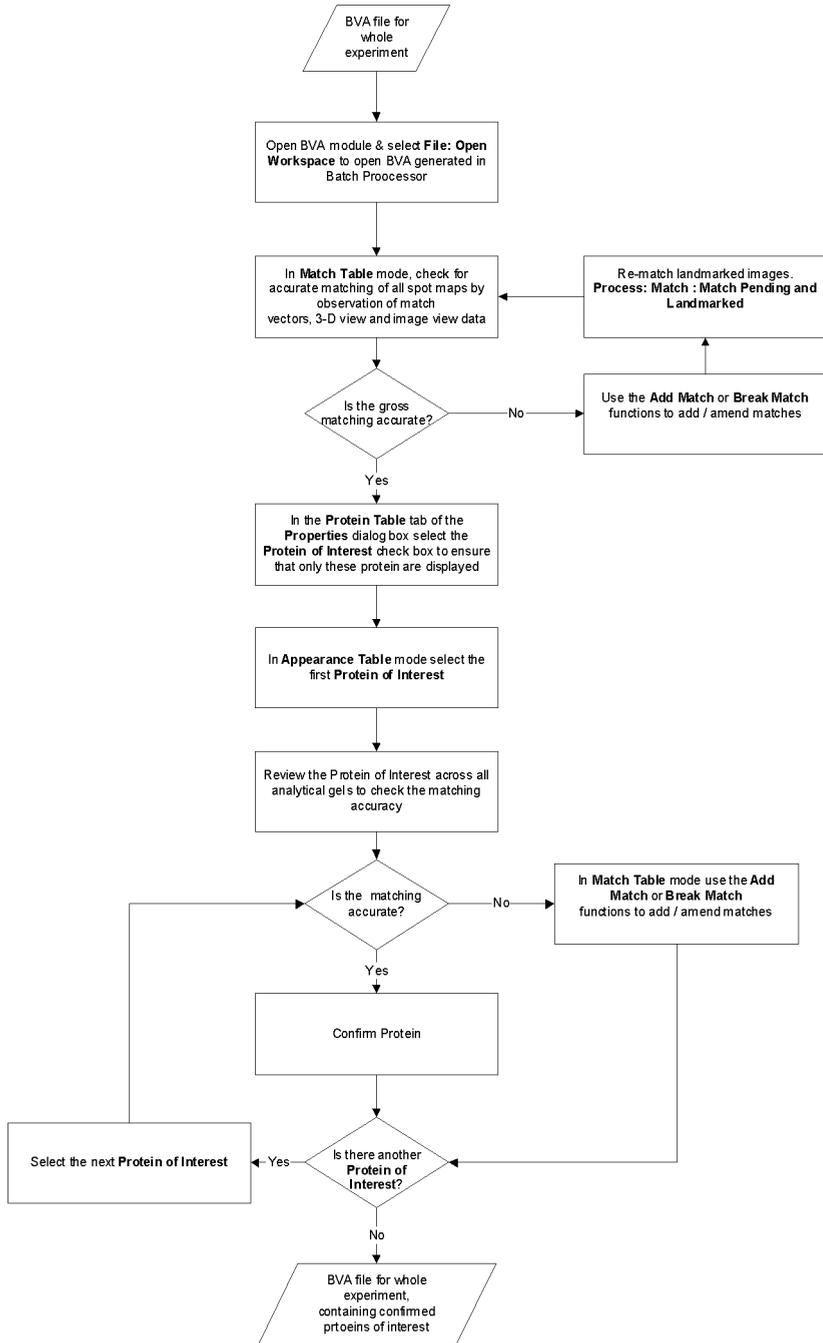


Continued on next page.

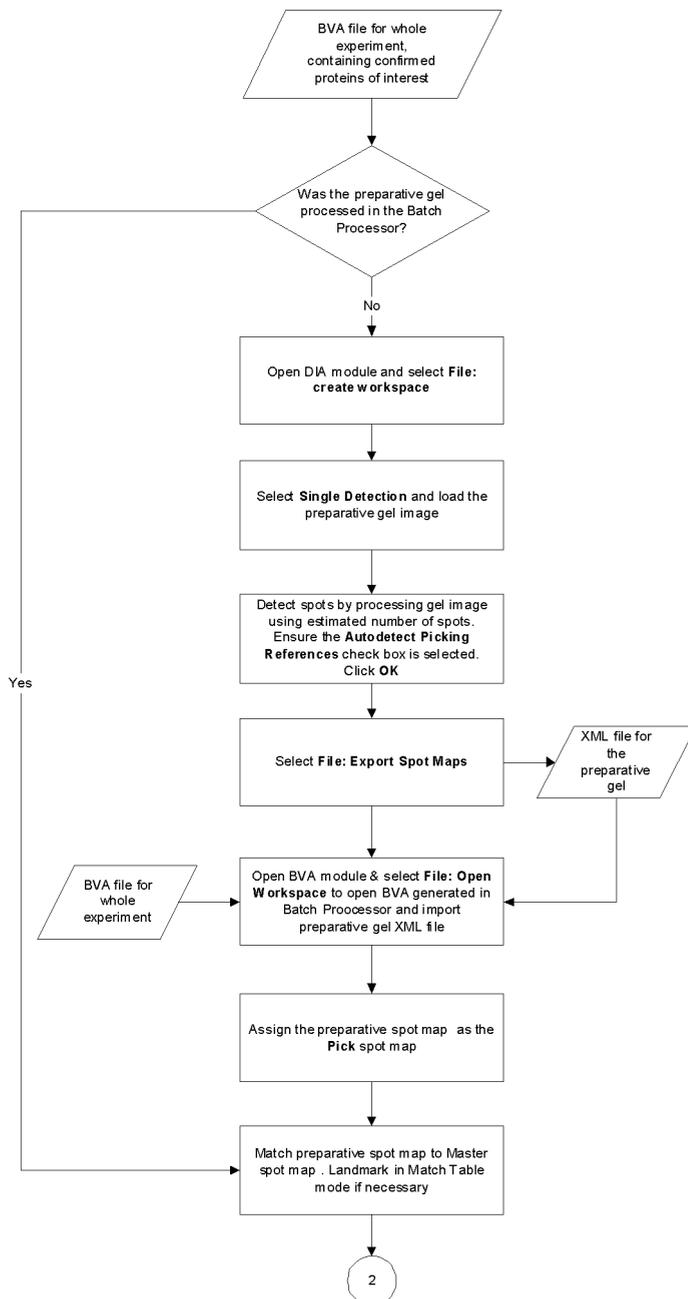
Continued from last page



A.2 Flow Diagram 2

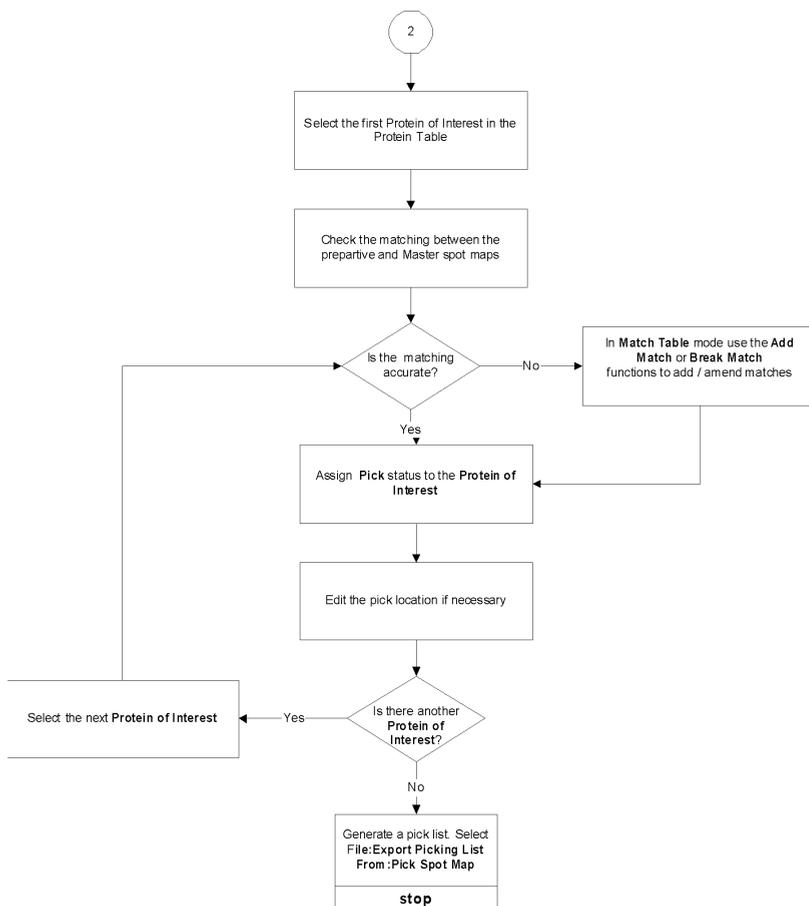


A.3 Flow Diagram 3



Continued on next page.

Continued from last page:



Appendix B Understanding the digital image

This Appendix deals with the acquisition and definition of the digital image. It explains some of the more common terms used in association with digital images and gives a brief insight into the GEL file format, which is the Typhoon imaging systems default format.

B.1 Image acquisition

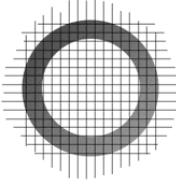
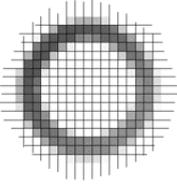
In order for any digital computer processing to be carried out on an image, it must first be acquired and stored within the computer in a suitable form. Image acquisition is a critical step and all primary data should ideally be stored exactly as it is recorded by the imaging device without significant data compression, as this may affect the accuracy of the recording. It should also be of the highest quality required for image resolution in the particular application. Acquisition can be achieved using a variety of scanners or digital imagers. These are usually based on PhotoMultiplier Tubes (PMTs) or Charge-Coupled Devices (CCD).

A PMT is an electro-optic device that converts light energy into electrical current and amplifies the current, whilst a CCD is a silicon-based integrated circuit consisting of a dense matrix of photodiodes that operate by converting light energy in the form of photons into an electric charge.

B.2 The digital image

The most practical way of storing an image as digital data is to divide the image into a grid of very small regions called “picture elements,” or “pixels”. In the computer this digital grid or “bitmap” represents the image. Each pixel is identified by its position in the grid, as referenced by its row (x) and column (y) number. Each pixel has a different color or gray scale value and together they form a representation of the image.

The images below show the digitalization procedure:

			
Original image.	Image is partitioned into a two-dimensional array of square sections. Each division will be used to form a pixel.	The scanning device then collects a single value to represent the entire square, which it then transmits to the computer.	After the computer has received values for each section, the entire image can be reconstructed using the pixel (x, y) values.

B.3 Formatting graphic files

Once an image has been acquired, it is converted to a particular file format for storage. There are a number of widely used image formats, such as TIFF, GIF etc. and knowing which format to use is a key issue. Some formats are proprietary and only useful in the program used to modify or acquire the image. Others are useful for exchange between programs and computer platforms or for presentation in web pages.

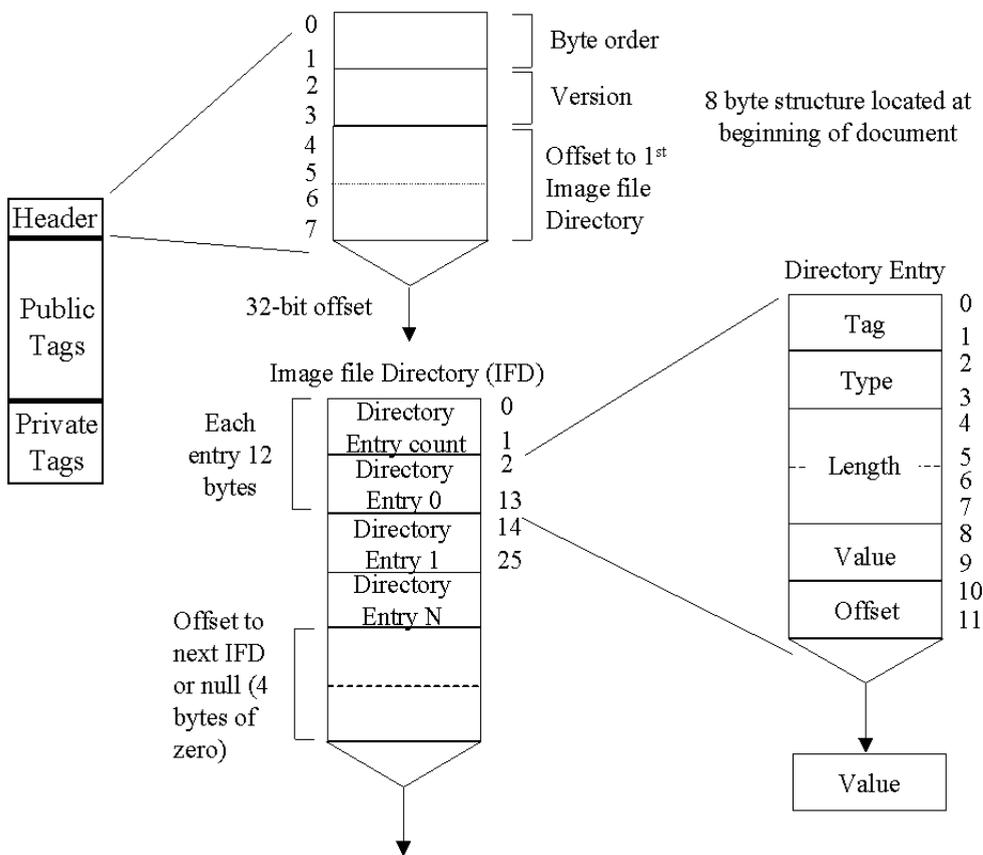
Image files include a certain amount of technical information, which is stored in an area called the image header or tag. The image header may be of use in displaying the image (e.g. length and width in pixels), identifying the image (e.g. name or source), or identifying the owner.

Scientific images should be acquired and archived using an information preserving or non-destructive format. The compression algorithms associated with some file formats actually distort some of the pixel information when the file is saved, while others do not. Loss of data may be unacceptable for use in quantitative analysis and as a consequence, DeCyder Differential Analysis Software only supports specific compression formats.

• **Tagged Image File Format (TIFF) images**

TIFF, an extremely complex and flexible image format is used to exchange files between platforms and software applications. The TIFF file consists of a number of labels (tags) that describe certain properties of the file (such as gray levels, color table, byte format, compression size etc.). After the initial tags, comes the data, which may be interrupted by more descriptive tags.

Although the TIFF file is an industry standard it has many variants. DeCyder Differential Analysis Software has been validated for file formats generated by Amersham Biosciences imaging devices recommended for Ettan DIGE system applications. The 16-bit TIFF format has $2^{16} = 65536$ levels of signal resolution and is the most commonly used file format for images. The image below shows a schematic diagram of a TIFF file:



- **Gel Image File Format**

Images scanned on the Typhoon are stored as GEL files, which are a variation of the 16 bit TIFF and are purely gray scale files with one bit for each pixel. These images have very wide dynamic ranges; typically images with 10^5 or 1-100,000 levels of signal resolution are possible. The GEL format uses a square root algorithm to compress the possible 100,000 levels of an image into the 65536 levels available; this is located in the TIFF private tag domain. The square root compression also provides higher signal resolution at the low end where small changes in signal are more critical. This is important when discriminating between two small values. The TIFF file will assign the same number to two values that differ by a small amount, whereas the GEL file accurately represents the data. The recommended instrument for scanning Ettan DIGE system gels for DeCyder Differential Analysis Software analysis is the Typhoon Variable Mode Imager.

B.4 Elements of the digital image

- **Pixel intensity value**

Each of the pixels that represent a stored image has a pixel intensity value, which describes how bright that pixel is. This intensity value represents a measured physical property i.e. emitted light. This value is the average for the whole area covered by the pixel. A pixel's address is denoted by its row and column co-ordinates in the two dimensional image.

- **Pixel or bit depth**

Pixel or bit depth (referred to as bit depth from now on) refers to the amount of information allocated to each pixel in a graphic image. Pixels have different bit depths, which determine how many grayscales are available to the image. Most image file formats store grayscale information in the header section of the file. An 8-bit image, 2^8 or 256 grayscale values, a 16-bit image 2^{16} or 65 536 grayscale values.

Due to variations in how graphics programs and conversion algorithms work, it is best to start off with as much bit depth as possible. Large bit depth also increases the dynamic range available when collecting images of, for example, gels with protein or nucleic acid bands that are going to be quantified by pixel intensity of the bands. When working with images that have high bit depths, remember that to actually see this information on the computer, the display must also be set to a high bit depth (thousands or millions of colors).

For black and white or binary images, pixels need only two bits of information (black or white), and hence the bit depth is 4.

Another type of image is the grayscale image, which is one in which the only colors are shades of gray. The number of grayscales used can vary but 16-bit systems have 65, 536 shades of gray, 0 being black and, 65 535 being white.

Table B-2. Bit depth and their corresponding number of different combinations available.

<i>Depth</i>	<i>Example</i>	<i>Combinations</i>
1 bit	0	2
2 bit	10	4
4 bit	0110	6
8 bit	01010011	256
12 bit	011010011101	4096
16 bit	1011000100101101	65536

• Pixel frequency histogram

The frequency histogram refers to the frequency representation of different shades of gray or color in the image. A frequency histogram displays the number of pixels representing each grayscale or color value. Frequency histograms most commonly represent grayscale images and have many uses: the histogram may reveal an under-or overexposed image (too many pixels with values close to 0, or too many with values close to 255 respectively), and the histogram can be manipulated to change the image: frequency values can be deleted, and upper and lower thresholds can be set.

B.5 Image dimensions

• Resolution

Image resolution is often confused with pixel dimension and refers to the number of pixels displayed per unit length of an image. Pixel resolution is the fineness of the divisions into which the scanner partitions the image. When the divisions are extremely small, the scanner is said to have high resolution; when the divisions are course, the scanner has low resolution. DeCyder Differential Analysis Software requires images which have a pixel resolution of 100 μm . Anything less does not contain the required amount of information, whilst anything more adds no further advantage to the image analysis procedure.

Resolution determines the area occupied by the images in conjunction with the pixel dimension and is a measurement of clarity, or detail. It can also refer either to an image file or the device, such as a monitor, used to display it. The relationship between number of pixels and area

is commonly expressed by number of pixels per inch (ppi) the more pixels per inch the better the resolution.

Output (print or display) resolution is more commonly expressed in terms of dots per inch (dpi). The resolution of a scanned image is also expressed in dpi where the number of dpi is equal to the number of ppi i.e. an image scanned in at 300 dpi will give you an image resolution of 300 ppi. Image-file resolution and output resolution combine to influence the apparent clarity of a digital image when it is viewed. The display monitor used will also influence apparent image quality.

The amount of resolution is often ruled by practical considerations e.g. the higher the dpi number the more information in the file, and the greater the ability to enlarge a detail from that image. If the principal life of an image is on screen e.g. an image for a web page, as opposed to being printed out, and if details for enlargement are not required from it, a resolution of 100 dpi will be sufficient. So, just as with image type, resolution needs to be matched to the purpose of the scan. The Table below shows the size of an uncompressed 1" x 1" image in different types and resolutions:

<i>Resolution (dpi)</i>	<i>2-bit Black and white (Kb)</i>	<i>8-bit grayscale (Kb)</i>	<i>24-bit color (Kb)</i>
400x400	20	158	475
300x300	11	89	256
200x200	5	39	118
100x100	1	9	29

• **Dynamic range**

Dynamic range is the ability of an imaging system to quantitatively detect very dim and very bright features within a single image and is related to bit-depth. It is a measurement of the number of bits used to represent each pixel in an image and hence determines the number of colors or shades of gray (grayscale) that can be represented in a digital image.

The dynamic range of a system is a function of the analogue-to-digital converter, the purity of the illuminating light, colored filters, and any system noise. It is measured on scale from 0.0 (perfect white) to 4.0 (perfect black), and the single number given for a particular imaging system tells how much of that range the unit can distinguish.

Variations in dynamic range of a system impact the quality of the digitized image more than simple resolution does. High-end imaging systems are more sensitive to the range of colors in the spectrum and can record minor differences between two almost identical colors.

Several variables determine a system's dynamic range: pixel depth (number of bits per pixel per color), sensitivity of the image capture device i.e. CCD / PMT, accuracy of the focusing optics, and precision of the measurement of the black and white points.

B.6 Image quality

Visual image quality is the cumulative result of the scanning resolution, the dynamic range of the scanned image and the scanning device or technique used. Image quality is often expressed in terms of resolution, but other factors also affect the quality of an image file. Images are often stored at much higher quality than they are displayed on a monitor because most printing devices are capable of a much higher resolution than screen displays.

A key trade-off in defining an appropriate level of image quality is the balancing of file size and resulting storage requirements with quality needs. Since pixel dimensions and color depth of a graphic image are directly proportional to the file size of the image, the higher the quality of an image, the more storage space it will occupy. High quality images also require more system resources e.g. higher bandwidth, networks, increased memory requirements and increased time and cost of the scanning process. Effective image compression provides a key to maintaining quality while using less storage space and system resources. However, it is highly recommended that images be archived onto CD-ROM to preserve storage space, particularly if using a network file server.

• Background & noise

Background is defined as undesired signal often resulting from autofluorescence or light scatter from a matrix or sample support. It can be minimized by the selection of appropriate matrix or sample support e.g. low fluorescence glass. Noise is defined as the statistical uncertainty inherent in a measurement, such as the standard deviation associated with measured background counts, e.g. with 2-D gels noise can be attributed to contaminants with fluorescent properties similar to the specific fluor being used. By ensuring that only high quality reagents are used and recommended procedures are followed noise can be minimized. The sensitivity of the imaging system can be adjusted by changing exposure times for CCD based systems or voltage settings for PMT based systems. The specific signal can be optimized to give the highest signal-to-noise ratio thus ensuring the maximum amount of information is obtained from the image.

Appendix C Spot processing algorithms

C.1 DIA normalization

The DIA normalization algorithm represents a series of steps that allows spot quantitation to be expressed as a function of the primary image, where an unchanged protein is represented as zero. Therefore when using an experimental design employing an internal standard, all the analysis gels are expressed as a function of the internal standard allowing accurate quantitative inter-gel comparative analysis.

To clarify some of the procedures involved in the DIA normalization, a step by step algorithm description is provided.

1 Calculate spot volumes.

Based on the result of the spot detection algorithm, the volumes of the detected spots are calculated and compensated for the appropriate background level. Background is subtracted on a spot-specific basis, by excluding the lowest 10th percentile pixel value on the spot boundary.

2 Calculate spot ratios.

All the spot volume pairs are combined to create the set of ratio values for the entire spot map. The ratios are calculated according to:

$$R_i = \log_{10}(V_{2i}/V_{1i}), \quad (i)$$

Where V_{1i} is the volume of spot i in the left, or primary gel image and V_{2i} is the volume of spot i in the right, or secondary gel image. The index i runs over all spots that are included in the analysis. Spots that have status “excluded” are thus not part of the normalization. The ratio R_i is also limited to the range $[-6, 6]$ to avoid infinite ratios for zero volumes.

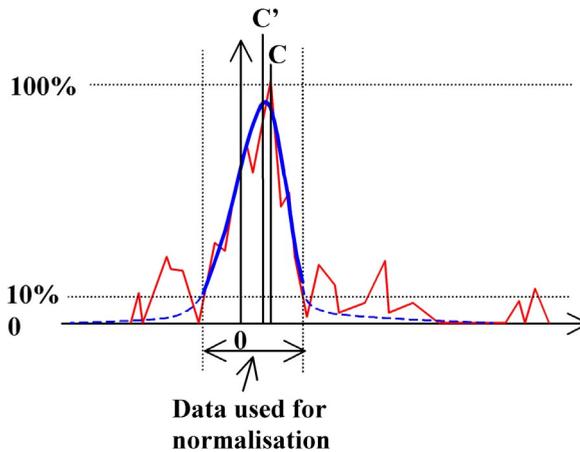
3 Calculate data histogram.

The ratio values are then combined to generate a histogram, plotting spot frequency against R_i (i.e. log volume ratio). The resolution of the histogram is 0.02.

4 Optimize a model histogram curve.

A normal distribution is fitted to the main peak of the frequency histogram. The tallest peak of the histogram C is used as the starting center position of the model curve, and all histogram data extending

from the tallest peak in both directions is included until the histogram value reaches below 10% of the main peak height. At this point any data outside is excluded from the model curve fitting procedure. This procedure is illustrated in the figure below. The model curve parameters are then optimized to the selected histogram data using a standard Least Means Square gradient descent algorithm. When the optimization is terminated, the center of the model curve is denoted C' .



5 Normalize the primary spot map

An assumption is made that the majority of the spots do not change, hence the modal peak of the model histogram is normalized to occur at zero on the log volume ratio axis.

For this purpose the spot volumes for the primary gel are multiplied $10^{C'}$ resulting in the mode volume for the primary and secondary being equivalent. i.e. the spot volumes in the primary spot map are normalized using the following procedure.

$$V_{1i}' = V_{1i} * 10^{C'} \quad (\text{ii})$$

Where V_{1i}' is the resulting normalized volume of spot in the left, or primary gel image.

The actual calculated volume of the detected spot is however never altered. This modification of the volume is local to the normalization procedure and ensures that the model peak is shifted so that the majority of spots expressed as unchanged (i.e. log volume ratio of 0). The modified spot ratios are calculated in this step according to:

$$R_i' = \log_{10}(V_{2i}/V_{1i}'), \quad (\text{iii})$$

This parameter is termed standardized log abundance in the BVA module and is used for the statistical analysis.

6 Re-calculate data histogram.

The modified ratio values are then combined to form a new histogram of the same resolution as the previous one. The normalized model histogram and data histogram are viewable in the histogram view of the DIA module, represented by the blue and red lines, respectively.

7 Calculate data standard deviation.

The standard deviation of all the spot volume ratios according to (iii) is calculated. This gives a rough indication of the spread of the data set.

Additional comments on spot normalization

For this normalization procedure to be accurate, the number of spots needs to be quite large. It is recommended that more than 100 spots are present in the normalization. However, the more valid spots that are present the better. If the number of spots to normalize is lower than 50, a model curve is not calculated. In this case the position of the tallest histogram peak is used as C' for normalization in (ii).

When spots in the workspace have been manually excluded, the workspace should be re-normalized based on the new set of included spots. This becomes exceedingly important if a large number of spots have been manually excluded.

In the DeCyder Differential Analysis Software-DIA environment, the expression ratios are displayed differently in the table than in the histogram. The histogram ratio representation R_i' is illustrated above, while the expression ratio, E , presented in the table is calculated according to:

$$E = 10^{R_i'} \text{ for } (R_i' \geq 0)$$

$$E = -1/10^{R_i'} \text{ for } (R_i' < 0) \quad (\text{iv})$$

This results in a value less than -1 for under-expressed ratios and a value larger than 1 for over-expressed ratios. Equal ratios are expressed as a ratio of 1 .

C.2 BVA Normalization

If an internal standard is not used in the experimental approach DeCyder Differential Analysis Software can normalize spot maps facilitating the comparison of spot volumes on different gels. The normalized volumes for a set of spots for a protein are then expressed as a ratio of the weakest spot. This parameter is referred to as the abundance.

This approach is not recommended, as inter-gel quantitative comparisons are not as accurate as the quantitation using an internal standard.

The steps involved in the BVA normalization process are described below.

1 Calculate spot volumes.

Based on the result of the spot detection algorithm, the volumes of the detected spots are calculated and compensated for the appropriate background level. Background is subtracted on a spot specific basis, by excluding the lowest 10th percentile pixel value on the spot boundary.

2 Calculate data histogram.

A histogram is plotted for spot frequency against \log_{10} spot volume for each spot map.

3 Calculate center of volume.

A Gaussian distribution is fitted to the main peak of the frequency histogram. The model curve parameters are optimized to the histogram data using a standard Least Means Square gradient descent algorithm. When the optimization is terminated, the center of the model curve is denoted as center of volume (CoV). The CoV therefore represents the central tendency of the spot volume population (i.e. a “typical” spot volume on specific spot map).

4 Calculate normalized spot volumes

The normalized spot volume is calculated for each spot on every gel.

$$V_{\text{norm}_g} = V_g / \text{CoV}_g \quad (\text{i})$$

Where V_{norm_g} is the normalized volume of a spot, V_g is the volume of the same spot and CoV_g is the center of volume for the spot map.

5 Calculate normalized spot ratios

The ratio of the normalized spot volumes relative to the smallest normalized spot volume for a group of matched spots (i.e. of the same protein) is calculated, according to:

$$R_{\text{norm}} = V_{\text{norm}_g} / V_{\text{norm}_{\text{min}}} \quad (\text{ii})$$

Where R_{norm} is the normalized ratio, V_{norm_g} is the normalized volume for a specific spot and $V_{\text{norm}_{\text{min}}}$ is the smallest normalized volume associated with the spot.

This ratio allows some degree of comparison between spots of the same protein. This parameter is referred to as the abundance in the BVA module.

Appendix D Experimental design and set up examples

In this appendix there are examples of complex experimental designs and statistical analyses. There are descriptions of both a paired testing and Two-Way ANOVA, which include details of the experimental design, setting up the statistical analyses and interpretation of example proteins.

D.1 Example of a paired experiment

Experimental objective

An experiment is designed to investigate changes in protein expression caused by a drug treatment after 24 hours in human subjects.

Experimental design

Blood samples are taken prior to drug treatment from five volunteers. Blood protein was individually extracted and stored. The drug was then administered to each individual, blood samples were taken 24 hours later and protein was extracted. Ten aliquot samples derived from the five individuals were taken and pooled to act as the internal standard. The pooled standard was labelled with CyDye DIGE Fluor Cy2 minimal dye. Pre-treated and treated blood samples from each volunteer were independently labelled with either CyDye DIGE Fluor Cy3 or Cy5 minimal dye. Equivalent amounts of labelled standard, volunteer 1 pre-treated, volunteer 1 post-treated were mixed and subjected to 2-D gel electrophoresis. The samples from the other 4 volunteers were similarly subjected to 2-D gel electrophoresis on separate gels (as described in the table overleaf – standards on individual gel not shown).

Assigning groups and individuals in DeCyder Differential Analysis Software

Spot data can be assigned into two experimental groups: pre- and post-drug treatment represented by 1 and 2, respectively.

The same 5 volunteers are present in both groups, therefore the data can be paired (i.e. assignment of individual volunteers within groups). The CyDye DIGE Fluor Cy2 minimal dye labelled internal standard samples are omitted from the table below.

Number	Gel	CyDye	Individual	Group
1	1	Cy3	1	1
2	1	Cy5	1	2
3	2	Cy3	2	2
4	2	Cy5	2	1
5	3	Cy3	3	1
6	3	Cy5	3	2
7	4	Cy3	4	2
8	4	Cy5	4	1
9	5	Cy3	5	1
10	5	Cy5	5	2

The above parameters can be designated either in the DeCyder Differential Analysis Software Batch Processor or from DeCyder Differential Analysis Software BVA module. The resultant spot map table is illustrated.

Spot Map Table

No.	Status	Image	Type	Label	No. of Spots	Matched	Function	Group	Group Description	Sample ID
13	Matched	Gel 01 Cy2 Standard.gel	DIGE Min	Cy2	2254	1677	A	Standard		
14	Matched	Gel 01 Cy3.gel	DIGE Min	Cy3	2254	1677	A	1	Pre-treated	1
15	Matched	Gel 01 Cy5.gel	DIGE Min	Cy5	2254	1677	A	2	Post-treated	1
4	Master	Gel 02 Cy2 Standard.gel	DIGE Min	Cy2	2572	2572	M, A	Standard		
5	Matched	Gel 02 Cy3.gel	DIGE Min	Cy3	2572	2572	A	2	Post-treated	2
6	Matched	Gel 02 Cy5.gel	DIGE Min	Cy5	2572	2572	A	1	Pre-treated	2
7	Matched	Gel 03 Cy2 Standard.gel	DIGE Min	Cy2	2507	1901	A	Standard		
8	Matched	Gel 03 Cy3.gel	DIGE Min	Cy3	2507	1901	A	1	Pre-treated	3
9	Matched	Gel 03 Cy5.gel	DIGE Min	Cy5	2507	1901	A	2	Post-treated	3
10	Matched	Gel 04 Cy2 Standard.gel	DIGE Min	Cy2	2353	1948	A	Standard		
11	Matched	Gel 04 Cy3.gel	DIGE Min	Cy3	2353	1948	A	2	Post-treated	4
12	Matched	Gel 04 Cy5.gel	DIGE Min	Cy5	2353	1948	A	1	Pre-treated	4
1	Matched	Gel 05 Cy2 Standard.gel	DIGE Min	Cy2	2298	1705	A	Standard		
2	Matched	Gel 05 Cy3.gel	DIGE Min	Cy3	2298	1705	A	1	Pre-treated	5
3	Matched	Gel 05 Cy5.gel	DIGE Min	Cy5	2298	1705	A	2	Post-treated	5

The experiment assesses the difference in protein expression between two experimental groups of paired data, therefore a paired Student's T-test is applicable.

The **Student's T-test** and the **Paired Tests** check boxes must therefore be selected in the statistics dialog box. The **Average Ratio** check box can also be selected to calculate the magnitude of any protein expression change between groups.

Example proteins

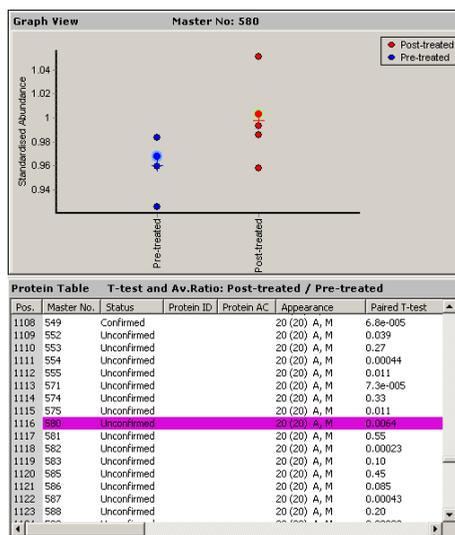


Fig D-1. Statistical data displayed for protein 580.

The independent T-test p value indicates that there is no significant difference in the mean expression of protein 580 from pre- to post-drug treatment groups. However, when data pairing is accounted for, the Paired T-test p value indicates that there is a significant change. Protein 580 expression levels appear to consistently increase regardless of the initial protein abundance.

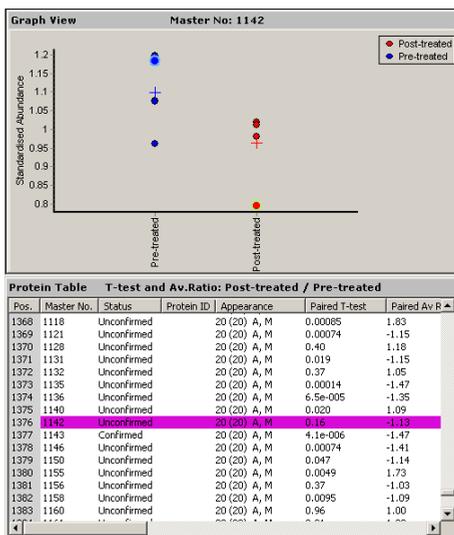


Fig D-2. Statistical data displayed for protein 1142.

The independent and paired T-test p values indicate that there is no significant difference in expression of protein 1142 from pre- to post-drug treatment groups. The observed decrease in protein expression between experimental groups is not due to a consistent abundance change of this protein between individuals. The independent and paired T-test p values indicate that there is no significant difference in expression of protein 1142 from pre- to post-drug treatment groups ($p=0.066$ and $p=0.160$, respectively). The observed decrease in protein expression between experimental groups is not due to a consistent abundance change between individuals.

D.2 Example of a two condition experiment

Experimental objective

An experiment is designed to investigate changes in protein expression in two *E.coli* bacterial strains incubated at 37 °C over a 90 minute time period.

Experimental design

Bacterial cultures derived from each strain were incubated at 37 °C. Cell sample aliquots were taken immediately prior to incubation then 30, 60 and 90 minutes after commencing incubation.

The cell samples were lysed, labelled with CyDye DIGE Fluor minimal dye, then equivalent amounts of protein were subjected to 2-D gel electrophoresis (as described in the table on the following page – standards on individual gel not shown). Triplicate gels were run and analyzed to account for experimental variation. Each gel contained a CyDye DIGE Fluor Cy2 minimal dye labelled internal pooled standard with CyDye DIGE Fluor Cy3 and Cy5 minimal dye labelled test samples.

Assigning groups and conditions

There are two conditions present in this experiment:

- Condition 1: two bacterial strains represented by 1 and 2.
- Condition 2: Four sampling time points: Pre-incubation 30, 60, and 90 minutes incubations are represented by 1, 2, 3, and 4, respectively.

The table below shows the designation of groups, conditions and individuals for the spot maps generated from the 12 experimental gels. The Cy2 dye labelled internal standard samples are omitted from the table below.

<i>Item</i>	<i>Gel</i>	<i>CyDye DIGE Fluor minimal dye</i>	<i>Strain (Condition 1)</i>	<i>Time (Condition 2)</i>	<i>Group</i>
1	1	Cy3	1	1	1
2	2	Cy3	1	1	1
3	3	Cy3	1	1	1
4	1	Cy5	1	2	2
5	2	Cy5	1	2	2
6	3	Cy5	1	2	2
7	4	Cy3	1	3	3
8	5	Cy3	1	3	3
9	6	Cy3	1	3	3
10	4	Cy5	1	4	4
11	5	Cy5	1	4	4
12	6	Cy5	1	4	4
13	7	Cy3	2	1	5
14	8	Cy3	2	1	5
15	9	Cy3	2	1	5
16	7	Cy5	2	2	6
17	8	Cy5	2	2	6
18	9	Cy5	2	2	6
19	10	Cy3	2	3	7
20	11	Cy3	2	3	7
21	12	Cy3	2	3	7
22	10	Cy5	2	4	8
23	11	Cy5	2	4	8
24	12	Cy5	2	4	8

The above parameters can be designated either in the DeCyder Differential Analysis Software Batch Processor or from DeCyder Differential Analysis Software BVA. The resultant Spot Map Table is illustrated on the following page.

Spot Map Table											
No.	Status	Image	Label	No. of Spots	Matched	Function	Group	Group Description	Condition1	Condition2	
2	Matched	Gel01 Cy3.gel	Cy3	2304	956	A	1	Strain 1, 0 mins	1	1	
6	Matched	Gel02 Cy3.gel	Cy3	2163	891	A	1	Strain 1, 0 mins	1	1	
10	Matched	Gel03 Cy3.gel	Cy3	2359	932	A	1	Strain 1, 0 mins	1	1	
4	Matched	Gel01 Cy5.gel	Cy5	2182	943	A	2	Strain 1, 30 mins	1	2	
8	Matched	Gel02 Cy5.gel	Cy5	2097	871	A	2	Strain 1, 30 mins	1	2	
12	Matched	Gel03 Cy5.gel	Cy5	2318	992	A	2	Strain 1, 30 mins	1	2	
14	Matched	Gel04 Cy3.gel	Cy3	2219	1281	A	3	Strain 1, 60 mins	1	3	
18	Matched	Gel05 Cy3.gel	Cy3	2085	2085	A	3	Strain 1, 60 mins	1	3	
22	Matched	Gel06 Cy3.gel	Cy3	2244	1260	A	3	Strain 1, 60 mins	1	3	
16	Matched	Gel04 Cy5.gel	Cy5	2288	1284	A	4	Strain 1, 90 mins	1	4	
20	Matched	Gel05 Cy5.gel	Cy5	2256	1795	A	4	Strain 1, 90 mins	1	4	
24	Matched	Gel06 Cy5.gel	Cy5	2303	1295	A	4	Strain 1, 90 mins	1	4	
26	Matched	Gel07 Cy3.gel	Cy3	2322	1243	A	5	Strain 2, 0 mins	2	1	
30	Matched	Gel08 Cy3.gel	Cy3	2086	1074	A	5	Strain 2, 0 mins	2	1	
34	Matched	Gel09 Cy3.gel	Cy3	2061	1097	A	5	Strain 2, 0 mins	2	1	
28	Matched	Gel07 Cy5.gel	Cy5	2183	1270	A	6	Strain 2, 30 mins	2	2	
32	Matched	Gel08 Cy5.gel	Cy5	2418	1289	A	6	Strain 2, 30 mins	2	2	
36	Matched	Gel09 Cy5.gel	Cy5	2349	1183	A	6	Strain 2, 30 mins	2	2	
38	Matched	Gel10 Cy3.gel	Cy3	2272	1155	A	7	Strain 2, 60 mins	2	3	
42	Matched	Gel11 Cy3.gel	Cy3	2255	936	A	7	Strain 2, 60 mins	2	3	
46	Matched	Gel12 Cy3.gel	Cy3	2111	1116	A	7	Strain 2, 60 mins	2	3	
40	Matched	Gel10 Cy5.gel	Cy5	2142	1208	A	8	Strain 2, 90 mins	2	4	
44	Matched	Gel11 Cy5.gel	Cy5	2176	892	A	8	Strain 2, 90 mins	2	4	
48	Matched	Gel12 Cy5.gel	Cy5	2459	1214	A	8	Strain 2, 90 mins	2	4	
1	Matched	Gel01 Standard Cy2.gel	Standard Cy2	2304	956	A	Standard				
3	Matched	Gel01 Standard Cy2.gel	Standard Cy2	2182	943	A	Standard				
5	Matched	Gel02 Standard Cy2.gel	Standard Cy2	2163	891	A	Standard				
7	Matched	Gel02 Standard Cy2.gel	Standard Cy2	2097	871	A	Standard				
9	Matched	Gel03 Standard Cy2.gel	Standard Cy2	2359	932	A	Standard				
11	Matched	Gel03 Standard Cy2.gel	Standard Cy2	2318	992	A	Standard				
13	Matched	Gel04 Standard Cy2.gel	Standard Cy2	2219	1281	A	Standard				
15	Matched	Gel04 Standard Cy2.gel	Standard Cy2	2288	1284	A	Standard				
17	Master	Gel05 Standard Cy2.gel	Standard Cy2	2085	2085	M, A	Standard				
19	Matched	Gel05 Standard Cy2.gel	Standard Cy2	2256	1795	A	Standard				
21	Matched	Gel06 Standard Cy2.gel	Standard Cy2	2244	1260	A	Standard				
23	Matched	Gel06 Standard Cy2.gel	Standard Cy2	2303	1295	A	Standard				
25	Matched	Gel07 Standard Cy2.gel	Standard Cy2	2322	1243	A	Standard				
27	Matched	Gel07 Standard Cy2.gel	Standard Cy2	2183	1270	A	Standard				
29	Matched	Gel08 Standard Cy2.gel	Standard Cy2	2086	1074	A	Standard				

The experiment assesses two conditions (strain and time), consequently a Two-Way ANOVA analysis is applicable.

The Two-Way ANOVA check box must therefore be selected in the Protein Statistic dialog box to perform the statistical test.

Examples of the statistical outcome for selected spots are illustrated on the following pages.

Example proteins

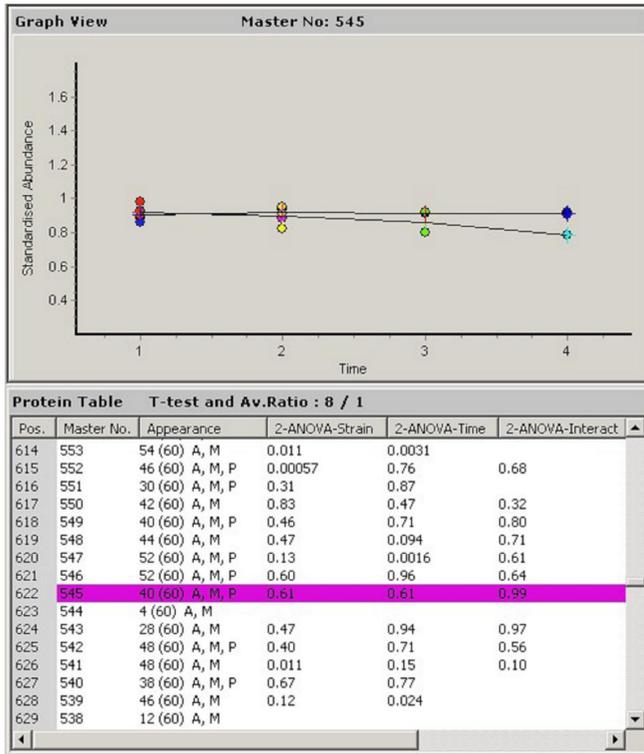


Fig D-3. Statistical outcome for protein 545.

The 2-ANOVA-Strain and 2-ANOVA-time values are not significant for protein 545. Therefore there is no strain-to-strain or duration of incubation effects on the expression of this protein.

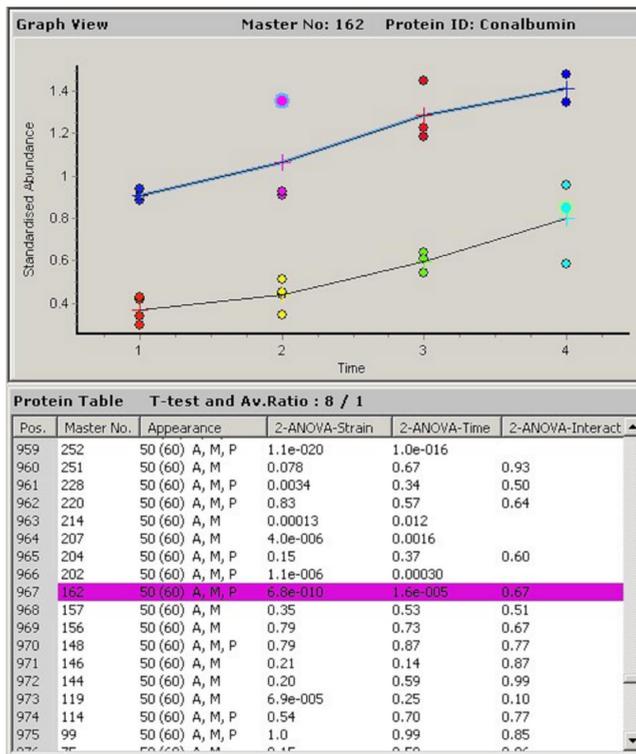


Fig D-4. Statistical outcome for protein 162.

The 2-ANOVA-Strain and 2-ANOVA-time values are statistically significant for protein 162. Therefore there are strain specific changes in expression of this protein. (i.e. one strain consistently has higher expression of this protein). Furthermore, the expression of this protein increases significantly in both strains over time. However, there is no significant interaction between strain and time of incubation, since the 2-ANOVA-interact is not significant.

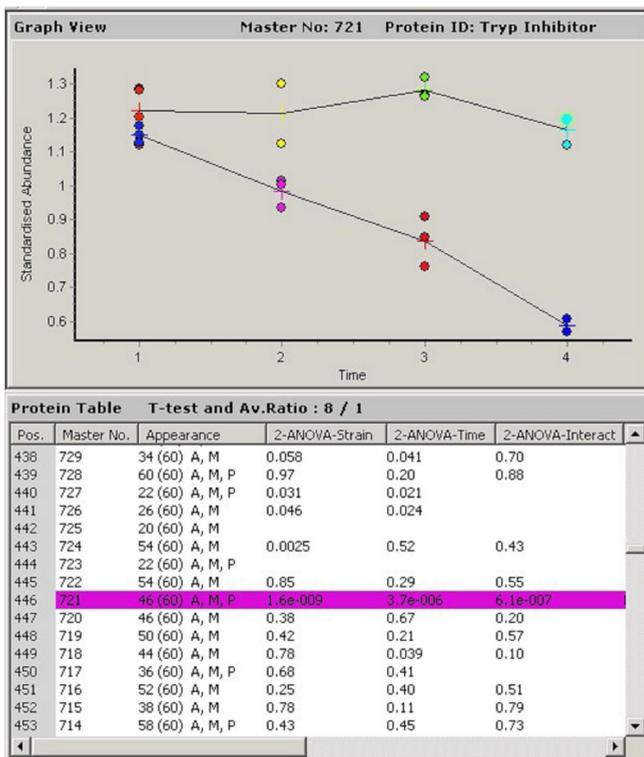


Fig D-5. Statistical outcome for protein 721.

The 2-ANOVA-Strain and 2-ANOVA-time values are statistically significant for protein 721. Therefore there are strain specific changes in expression of this protein. The expression of this protein also changes significantly in at least one timepoint. Moreover, there is a significant interaction (2-ANOVA-Interaction <0.01) between strain and time, whereby there is a decrease in expression of protein 721 over time that is specific to one strain.

Appendix E DeCyder Differential Analysis Software keyboard shortcuts

E.1 DIA module keyboard shortcuts

<i>Spot Controls</i>	
Alt + C	Confirm spot
Alt + P	Previous
Alt + N	Next
Alt + I	Include
Alt + X	Exclude
Alt + K	Pick
Alt + S	Spot ID
Alt + M	Comment
<i>Workspace</i>	
Alt + F	Open File menu
Alt + E	Open Edit menu
Alt + V	Open View menu
Alt + R	Open Process menu
Alt + H	Open Help menu
<i>File menu</i>	
Ctrl + N	Create new workspace
Ctrl + O	Open workspace
Ctrl + S	Save workspace
Ctrl + P	Show print dialog
<i>Edit menu</i>	
Ctrl + T	Add Spot to Table
Ctrl + C	Copy
<i>View menu</i>	
F7	Zoom In
F8	Zoom Out
F9	Fit to window
Ctrl + B	Contrast - Brightness
Ctrl + A	View Area in 3-D
Ctrl + R	Rotate 3-D
<i>Help menu</i>	
Shift + F1	Help What's This?

E.2 BVA module keyboard shortcuts

Workspace	
Alt + F	Open File menu
Alt + E	Open Edit menu
Alt + V	Open View menu
Alt + R	Open Process menu
Alt + H	Open Help menu
File menu	
Ctrl + N	Create new workspace
Ctrl + O	Open workspace
Ctrl + S	Save workspace
Ctrl + P	Print
View menu	
F7	Zoom In
F8	Zoom Out
F9	Fit to window
Ctrl + B	Contrast - Brightness
Ctrl + A	View Area in 3-D
Ctrl + R	Rotate 3-D
Alt + Enter	Properties
Modes	
Alt + S	Spot Map Table
Alt + M	Match Table
Alt + P	Protein Table
Alt + A	Appearance Table
Help menu	
Shift + F1	Help What's This?
Protein Related functions	
Alt + ↓	Select next row in current table
Alt + ↑	Select previous row in current table
Alt + Page Down	Scroll down current table
Alt + Page Up	Scroll up current table
Alt + →	Select next Spot Map/protein
Alt + ←	Select previous Spot Map/ protein
Alt + X	Set focus to Controls
Alt + C	Confirm/Unconfirm protein or match
Alt + D	Add/Break match
Alt + K	Pick protein
Alt + I	Protein of Interest
Ctrl + Q	Primary image scroll down menu
Alt + Q	Secondary image scroll down menu

E.3 Batch Processor keyboard shortcuts

<i>File menu</i>	
Ctrl + N	New Batch
Ctrl + O	Open
Ctrl + S	Save
Ctrl + P	Print
Ctrl + B	Add DIA Batch Item
Ctrl + M	Remove DIA Batch Item
Ctrl + D	Add Spot Map(s)
Ctrl + T	Remove Spot Map(s)
<i>Edit Menu</i>	
Ctrl + E	Edit Item
<i>View Menu</i>	
Ctrl + I	DIA Batch List
Ctrl + V	BVA Batch List
<i>Process Menu</i>	
Ctrl + R	Process

Appendix F Related products and documentation

F.1 Related products

<i>Item</i>	<i>Product Code</i>
DeCyder Differential Analysis Software pre-installed on a PC and a single user licence	18-1163-05
DeCyder Differential Analysis Software including single user licence	18-1156-17
DeCyder Differential Analysis Software additional user licence	18-1150-45
Ettan Spot Picker	18-1145-28
Ettan Spotter	For product information please inquire with your local Amersham Biosciences sales office.
Ettan Spot Handling Workstation	18-1164-05
Ettan Spot Handling Workstation - LWS for integration with Ettan LWS	18-1164-06
Ettan Digester 100-120V	18-1152-59
Ettan Digester 220-240V	18-1142-68
Ettan MALDI-ToF Pro 120 V	18-1156-54
Ettan MALDI-ToF Pro 240 V	18-1156-53
2-D Electrophoresis, Principles and Methods	80-6484-89
Ettan DIGE Training CD	18-1164-39
Typhoon Variable Mode Imager	For product information please inquire with your local Amersham Biosciences sales office
Table continued on next page	

Item	Product Code
CyDye DIGE Fluor Cy2 minimal dye (25 nmol)	RPK 0272
CyDye DIGE Fluor Cy3 minimal dye (25 nmol)	RPK 0273
CyDye DIGE Fluor Cy5 minimal dye (25 nmol)	RPK 0275
CyDye DIGE Fluor Cy2 minimal dye (10 nmol)	25-8008-60
CyDye DIGE Fluor Cy3 minimal dye (10 nmol)	25-8008-61
CyDye DIGE Fluor Cy5 minimal dye (10 nmol)	25-8008-62
CyDye DIGE Fluor Labelling Kit for Scarce Samples	25-8009-83
CyDye DIGE Fluor Labelling Kit for Scarce Samples plus Preparative Gel Labelling	25-8009-84
PlusOne Urea	17-1319-01
PlusOne CHAPS	17-1314-01
PlusOne DTT 5 g	17-1318-02
Ettan 2-D Quant kit	80-6483-56
Ettan 2-D Clean-Up kit	80-6484-51
PlusOne Bromophenol Blue	17-1329-01
PlusOne Tris	17-1321-01
PlusOne SDS powder	17-1313-01
PlusOne Glycerol (87%)	17-1325-01
PlusOne ReadySol IEF (40% T, 3% C)	17-1310-01
PlusOne TEMED	17-1312-01
PlusOne Ammonium Persulphate	17-1311-01
PlusOne Glycine	17-1323-01
Ultrapure DMF	US14862
DeStreak Rehydration Solution	17-6003-19

F.2 Related documentation

<i>Document</i>	<i>Code no.</i>
Ettan DIGE system User Manual	18-1173-17
Ettan DIGE Quick Protocols	18-1164-41
Typhoon™ Instrument Guide	63002831
Ettan 2D-MS and LWS software Laboratory Guide	18-1167-27
Ettan 2D-MS and LWS software Laboratory Guide Appendices	18-1170-99
Ettan Spot Handling Workstation 2.01 - LWS 1.0, User Manual	18-1163-89
Ettan MALDI ToF Pro, User Manual	18-1144-01

Appendix G Glossary

G.1 DeCyder Differential Analysis Software glossary

2SD value. 2 standard deviation of the spot ratio distribution, 95% of the spots lie within this ratio for normally distributed data.

Abundance. The relative volume among all spots representing a particular protein in a BVA data set. The weakest spot is taken as 1.00 and the others are displayed relative to this.

ANOVA. ANalysis Of VAriance is a family of methods used to perform statistical analysis on experimental results.

ANOVA 1 way. One-Way ANOVA test (assigns statistical significance to differences in standardized protein abundance between experimental groups).

ANOVA 2 way. Two-Way ANOVA test, (assigns statistical significance to both separate and mutual effects of two experimental conditions on standardized protein abundance).

Appearance Table. Region within a BVA workspace where users can track all information on a particular spot from all the gel images.

Area of Interest. User defined region outside of which any detected protein spots are excluded from analysis.

Artifact Rejection. Filtering out of non proteinaceous signal.

Auto Level. The stage in the algorithm in which spots are matched.

Bandwidth. The transmission capacity of a communications channel, usually expressed in bits or bytes per second.

Batch Processor. DeCyder Differential Analysis Software module capable of performing fully automated co-detection, quantification and matching of multiple spot maps.

BVA. Biological Variance Analysis module of DeCyder Differential Analysis Software.

BVA Batch list. The spreadsheet in the batch processor depicting which spot maps are to be matched.

Centre of Volume (CoV). Central tendency of the spot volumes on a spot map.

Co-detection. Simultaneous detection of labelled protein spots from two in-gel images.

Comment. User defined text string that can be linked to specific protein or spot map.

DeCyder Differential Analysis Software XML. XML format used to transfer data between the various DeCyder Differential Analysis Software modules.

DIA. Differential In-gel Analysis module of DeCyder Differential Analysis Software capable of the fully automated co-detection, quantification and matching of an in-gel image pair.

DIA Batch List. Spreadsheet in the batch processor depicting which image pairs are to be detected and quantified.

Exclude Filter. Filter used to remove non proteinaceous artifacts such as dust.

Groups. The collection of spot maps relating to image pairs whose sample images have undergone exactly the same experimental conditions (e.g. 3 spot maps of samples from disease patients treated with drug A at a specific dose is a single group).

Histogram Selections. User adjustable parameters that alter the appearance of the histogram in the DIA module.

Independent Data. Data sets which have no effect on one another.

Label. Indicates the fluor used to label the protein.

Landmarking. Process of manually matching spots to the master image before automated matching to aid the matching algorithm.

Master Image. Spot map to which all others are matched to.

Match confirmation. Manual verification by the user of matched spots.

Matching. Automatic linking of spots on selected images to the corresponding spot on the master image.

Match table. Area within a BVA workspace where spot match data is displayed.

Maximum Peak height. Pixel value at the X,Y position of the spot. This is the actual detected peak height value compensated for background level.

Maximum Volume. The highest volume from two co-detected spots.

Normalization. Process that allows the direct comparison of spot data derived from different gels.

Null hypothesis. Posits that there is no difference between variables being tested. To reject it is to infer “statistical significance”.

Paired Data. Data sets with an association between data point in every group.

Pick Filter. Function that enables the user to designate spots for picking based on user defined parametric values.

Primary image. Refers to the left hand gel image present in the image view of the DIA module and the spot map table, protein table and appearance table modes of the BVA module. In the Match Table mode of BVA module, the primary image is the right hand image as the left hand image is taken by the master gel.

Protein ID. Unique protein identifier that can be used to search.

Protein Table. Area within a BVA workspace where all the statistical data from the analytical gels is located.

Protein Statistics. Function in the BVA module that applies statistical tools to the protein data (providing that an experimental design with an internal standard was used).

Repeated Measures. ANOVA Statistical ANOVA test applied to paired data.

Scatter Parameter. The spot data type used to display the spot ratios in the histogram. This is shown in the right y-axis on the histogram and can be maximum slope, volume, peak height or area.

Secondary image. Refers to the right hand gel image present in the image view in the DIA module and the spot map table, protein table and appearance table mode of the BVA module.

Slope. Maximum gradient associated with the 3 dimensional attributes of a spot map pair.

Spot Controls. Allows user to control, assign or input data to the workspace or individual protein spots.

Spot Map Table. Area within a BVA workspace where all the spot map images and their associated assignments are located.

Spot Number. Unique identifier for a spot in DIA or a spot set in BVA.

Standard Image. Image relating to the internal standard from each gel.

Standardization. Process of quantifying spot map data in relation to the corresponding standard spot map data.

Standardized Abundance. Abundance relative to the standard image displayed as a ratio.

Student's T-test. Statistical test that assesses differences between two populations.

Threshold Mode. Value above or below which spots are classed as being differentially expressed.

TIFF Image. Flexible image format used to exchange files between platforms and software applications.

Volume. Sum of the detected pixel values above background within a spot boundary.

Volume Ratio. Refers to the ratio of the normalized volumes of a pair of spots from a spot map pair. A value of 2.0 represents a two-fold increase while -2.0 represents a two-fold decrease, whilst a value of 1.00 represents an unchanged spot.

Workspace. Environment where all the experimental data is stored.

XML (Extended Markup Language). Structured universal tagged language.

XML Toolbox. Shell, housing modules used for the extraction of data from DeCyder Differential Analysis Software XML files in the form of tab separated text or web tables.

Index

Numerics

3-D View 28, 36, 56

A

Abundance 255
Add attribute 142
Add databases 95
Addresses 3
Analysis image 58
Annotation 55, 102
ANOVA 81
 one-way 75, 83, 98
 two-way 75, 85, 98
 two-way, example 86, 261
Appearance 98
Appearance Table 101
Area of interest 43
Artifact 44
Assign picking references 108
Attributes 142
Average ratio 78

B

Background 249
Batch list 121
Batch processor 121, 157, 226
BVA 51, 156
BVA Module 51

C

Calculate Mw and pl 92
Center of volume 254
Co-detection 31
Computer requirements 23
Conditions 87
Consumables 272
Contrast and Brightness 34
Create Workspace 29, 53
Customizing display colors 48, 103
CyDye DIGE Fluor minimal dyes 12

D

Data Control Panel 28
Database linking 94

Database query in DeCyder BVA	151
DeCyder XML	129
DeCyder XML toolbox	140
LWS Pick List tool	146
Protein Identification tool	140, 148
Detected proteins in gel	149
DIA	27, 155, 159
DIA Graphical User Interface	28
DIA Module	27
DIA Processes	27
Digital image	243
Display Mw and pI values	93
Dust particles	44
Dynamic range	248

E

Enter Mw and pI	91
Enter samples into Ettan LWS	140
as a gel	144
as a tube	142
overview picture	141
Estimated Number of Spots	32
Ettan DALT electrophoresis unit	32
Ettan DIGE system	12
Ettan Spot Handling Workstation	27
Ettan Spot Picker	27
Evaluation file	144, 145, 148
Exclude Filter	44, 203
Exporting data	49, 50, 106
Extract data	130

F

Frequency distribution	40
------------------------------	----

G

Gel artifact removal	224
Generate pick list	119
Generate standard report "Detected proteins in gel"	149
Graph view	71
Graphical user interface	51
Groups	59
defining	74

H

Histogram	40
Histogram View	28, 40

I	
Image acquisition	243
Image file	144, 148
Image View	28, 34, 54
Immobiline DryStrip	32
Import data	114
Importing protein data templates into the BVA module	151
Independent analyses	76
Installation	23
Isoelectric point calculation	91
K	
Keyboard shortcuts	267
L	
Landmarking	63
Load XML files	130
M	
Manual spot exclusion	46
Master image	58, 63
Match confirmation	65
Match Table	70
Matching	63–65
Max peak height	38
Max volume	38
Model curve	40
Molecular weight calculation	91
Multiple images	226
MW value	148
N	
NCBI	151
Noise	249
Non-protein spots	44
Normalization	
BVA	254
DIA	251
O	
One-Way ANOVA	75, 83, 98
Open Workspace	30
Original DeCyder BVA workspace	148, 151
P	
Paired analyses	76
example	257

Paired data	76, 258
Patents and Licences	3
Peak height	37
PI value	149
Pick filter	110
Picking image	58
Picking reference	56, 108
Pixel	246
Post-translational modification	48, 94
Preparative gel	107
Primary image	29
Print	49
Process multiple images	226
PROSITE	95
Protein AC	98
Protein confirmation	94
Protein data	151
Protein Filter	90, 112
Protein ID	98, 148
Protein Name	148
Protein of interest	94
Protein Table	98
PTM	48, 94

Q

Quality score	67
Quantitation	33, 53

R

Reagents	272
Recommended workflow	237
Flow Diagram 1	238
Flow Diagram 2	240
Flow Diagram 3	241
Reference manual	11
Reference markers	108
Related documentation	273
Related products	271
Repeated measures	83
Resolution	247
Run batch processor	235

S

Sample Attribute Templates	142
Sample Definition	
gel	145
tube	143
Save	49, 104

Scatter parameter	40
Sierra LWS	143, 145
Sierra web	142, 144
Secondary image	36
Signature	55
Slope	39
Spot area	37
Spot confirmation	46
Spot detection	31
single, double, triple	29
Spot exclusion	43
Spot Map attribute	58
Spot Map function	58
Spot Map Table mode	58
Spot merging	68
Spot number	37
Spot position	37
Spot ratio	33
Spot statistics	41
Spot volume	33, 37
Standard report "Detected proteins in gel"	149
Statistics	73
Student's T-test	75, 78
SwissProt	95, 97

T

Table View	28, 38
parameters	39
Tagged Image File Format (TIFF)	245
Template Spot Map	58
Threshold	41
Trademarks	3
TrEMBL	95
Tutorials	11
Two-Way ANOVA	75, 85, 98
Two-Way ANOVA, example	261

U

User manual	11
User-defined protein labelling	94

V

Volume	33
--------------	----

W

Workspace	
create	29
open	30

properties 31

X

XML 129

 Tag definitions 132

XML Toolbox 129

Z

Zoom 34

