

NOTICE: Varian, Inc. was acquired by Agilent Technologies in May 2010. This document is provided as a courtesy but is no longer kept current and thus will contain historical references to Varian. For more information, go to **www.agilent.com/chem**.



Agilent Technologies

Varian, Inc.
2700 Mitchell Drive
Walnut Creek, CA 94598-1675/USA

Star Chromatography Workstation

Version 6

Regulatory Compliance Manual



Trademark Acknowledgment

Microsoft, Windows, Windows 95, Windows 98, Windows NT, Windows 2000, and WindowsXP are registered trademarks of Microsoft Corporation.

Other brand and product names are trademarks or registered trademarks of their respective holders.

Table of Contents

Introduction.....	1
Scope and Purpose.....	1
Product Development Procedures	3
Outline of Development Procedures.....	3
The Star Workstation Test Plan: Scope and Objectives.....	4
Functional Testing.....	4
Operational Testing.....	5
Reliability Testing	8
New Product Checklist.....	8
Star Workstation Product Life Cycle Support	8
History of Star Workstation Development and Prior Releases	10
Validation of the Software Installation.....	15
Factory Installation and Checkout Procedure.....	15
Validation of Star Workstation Installation and Function	16
Validation of the Installed Star Workstation	16
Maintenance Schedule of the host computer	16
Data Documentation and Logs.....	19
Documentation of Data Collection: Operating Logs	19
Documentation of Data Collection: Validation and Verification	21
Documentation Provided by the User	22
Data Processing	25
Processing Raw Data: Data Collection.....	25
Peak Processing	25
Parameters Affecting Peak Detection.....	26
Elements of Peak Detection.....	28
Calculation of Peak Size	29
Peak Identification.....	32

Replicates and Calibration Levels.....	32
Weighted Regression.....	33
Calculations Performed by the Star Workstation	33
Determination of the Calibration Curve Equation	33
Solution of the Calibration Equation During Analysis	46
Coefficient of Determination.....	52
Relative Standard Deviation of Response Factors.....	53
Star Workstation Calculation Types.....	54

Introduction

Scope and Purpose

The purpose of this manual is to help you comply with regulations such as Good Laboratory Practice (GLP), Good Automated Laboratory Practice (GALP), and International Standards Organization (ISO) 9000. This section provides information and documentation that can be used to demonstrate that the Varian Star Chromatography Workstation software was designed, developed, and tested according to commonly accepted quality engineering practices. Our ISO 9001 certification ensures that the procedures and practices described in this manual have been carefully followed.

All hardware and software design and test procedures, along with source code and test results are available for customer review at Varian Chromatography Systems. This information may be reviewed by a Star Workstation user, or an agent of a user. Regulatory agencies may also confirm that an instrument was designed, developed, manufactured, and tested in accordance with the procedures described in this document.

Copies of the software source code and documentation for this product are also stored away from Varian Chromatography Systems to ensure that they will be accessible in the event of natural disaster or business failure. Varian cannot guarantee your compliance with GLP, ISO 9000, or the requirements of any other regulatory agency. This is strictly the responsibility of the instrument user. The information contained in this manual is intended to help you comply with these regulations.

Product Development Procedures

Outline of Development Procedures

The procedure Varian Chromatography Systems employs in product development is outlined in QPCS (Quality Procedures Chromatography Systems) number 070, which is part of our ISO 9000 documentation, and can be viewed at Varian Chromatography Systems. In brief, Varian's development of the Star Workstation moved through the following phases:

Definition of requirements for the Star Workstation product, based on information gathered from customers, the sales force, and from current research in the field of chromatography.

Development of the Star Workstation's Engineering Product Specifications/Software Requirements Specifications, listing operations to be performed by the Star Workstation, and their expected results.

Development of the Star Workstation's Software Design documentation based on its Software Requirements Specifications.

Formulation of a test plan to evaluate Star Workstation performance, including both boundary testing (maximum, minimum, invalid input), and parallel testing (independent verification of results).

Installation of a computer-based discrepancy tracking system to record all system discrepancies and change requests, as well as their final resolution, as determined by a change control board.

Development of the Star Workstation product in conformance with its Product Specifications.

Performance of acceptance testing to verify that the Star Workstation performed according to system specifications.

System evaluation by means of a pre-release program in which selected customers and Varian field representatives observed Star Workstation performance, rated reliability, and attempted to uncover any discrepancies not previously noted.

Analysis of Star Workstation performance data gathered throughout the Star Workstation development process, in order to document reliable long-term operation.

The Star Workstation Test Plan: Scope and Objectives

The scope of the Star Workstation test plan includes functional, operational, and reliability testing.

The objective of the test plan is to evaluate all aspects of the Star Workstation to verify that it meets both product and software specifications. Throughout the evaluation process, instrument communications were subjected to extra scrutiny, in order to discover any possible communication conflicts.

The Star Workstation Functional Evaluation test plan calls for checking every entry in every window for each of the applications that comprise the Star Workstation.

The Star Workstation Operational Evaluation test plan involves observing the actual performance and characteristics of the Star Workstation while it is performing routine tasks.

Functional Testing

Functional testing of the Star Workstation includes checks of invalid input handling, maximum and minimum accepted values, and default values for each field in each window of the Star Workstation software.

The test plan for Access Control and Audit Trail software is listed in the Access Control and Audit Trail Documentation and Setup Manual.

Operational Testing

The Star Workstation functions examined during operational testing included, but were not restricted to, the following:

- I. Concurrency
 - A. Star Applications with other Star Applications
 - B. Star Applications with non-Star Applications
- II. Calculations
 - A. Internal standard
 - B. External standard
 - C. Normalized percent
 - D. Area/Height percent
 - E. Peak Detection and quantitation
 - F. Re-integration/recalculation
 - 1. Automated recalculations
 - 2. Individual recalculations
- III. Printing
 - A. Batch printing
 - B. Individual printing
 - C. Chromatogram printout
 - D. Results printout
- IV. Reports
 - A. Calibration Reports
 - B. Calibration Block Reports
 - C. Verification Reports
 - D. Curve Printouts and display
 - E. Analysis Reports
 - F. Blank Baseline Reports

- G. Run Logs
- H. Error section
- I. Notes
- V. Error Handling
 - A. Instruments
 - B. Sequences, SampleLists
 - C. Data Handling
- VI. Help
 - A. Application specific
 - B. General
- VII. Advanced Applications
- VIII. PolyView

The organization of the Star Workstation Test Plan is shown in the following outline:

- I. Star Workstation Application Tests
 - A. Installation
 - B. Method Editor
 - 1. Functional Testing
 - a) LC Modules: 9012, 9010, 9001, 9002, 9050, 9065, 9100, 9300, 9200, Mistral Column Oven, AI-200
 - b) GC Modules: 3400, 3600, 3800, SPME II / 8200 Standalone
 - c) ADCB Control
 - d) Data Handling: 9050, 9065, ADCB, GC3800
 - 2. Operational Testing

- C. System Control
 - 1. Functional Testing
 - a) Configuration Window
 - b) Instrument Window
 - 2. Operational Testing
 - D. Report
 - 1. Functional Testing
 - 2. Operational Testing
 - E. Interactive Graphics
 - 1. Functional Testing
 - 2. Operational Testing
 - F. Concurrency with non-Star applications
 - G. Multipoint/Multilevel Data Handling
- II. Year 2000 Testing
- A. Roll over high risk dates: 12/31/98, 12/31/99, 2/28/00, 2/29/00
 - 1. Roll-over when power on
 - 2. Roll-over when power off
 - B. Handling of high risk dates: 9/9/99, 12/31/99, 1/1/00, 2/29/00
 - 1. Entry
 - 2. Display
 - 3. Print / Send

Reliability Testing

During the process of Star Workstation development, Star Workstation evaluators prepared reliability reports for every automated sequence they ran on their test systems. They compiled and evaluated data from these reports to make certain the Star Workstation met reliability standards, and to identify any problems with automation and operation.

In addition to the evaluation of reliability data, Varian Technical Service representatives conducted customer support evaluations, and reported their findings.

New Product Checklist

A New Product Checklist must be completed and signed-off by each department manager and by the General Manager of Varian Chromatography Systems prior to customer shipment of any new product.

Star Workstation Product Life Cycle Support

Product Introduction

Varian offers product support in all phases of the product life cycle. Prior to the introduction of each version of the Star Workstation product, Customer Support Representatives underwent extensive training in its use. Customer training and support are available for Star Workstation users after they acquire the product.

Product Support

Varian Chromatography Systems Quality Procedure QPCS 120 defines our support policy for software revisions and upgrades. According to this procedure, Varian Chromatography Systems guarantees to support a software version for 12 months from the date of the first shipment of the revision which replaces it or until the date of the first shipment of another revision. In other words,

at any one time Varian will support the two most recent versions of the software with the older version being supported for only the first 12 months of the newer version's lifetime. Varian software support means that Varian personnel will discuss application and operation questions about the software with customers and will assist in resolving service problems related to the software revision. In some cases the "fix" for the service problem with the old software version may be the purchase of the newer version of software.

Product Obsolescence

Varian Chromatography Systems Quality Procedure QPCS 120 states:

"When a product is discontinued, the final version of the product software is supported by Varian Chromatography Systems for five years from the discontinuation announcement."

This procedure is described in detail in Quality Procedure QPCS 115 "Product Discontinuance/ Obsolescence."

After the guaranteed five year support period for a discontinued product, the product is said to be "obsolete". Obsolete products are not supported. Varian may be able to supply repair parts or support for obsolete products, but this support is not guaranteed.

For software, product discontinuance occurs when the software is no longer being offered for general sale and there is not a more current revision or upgrade available. If Varian has a more current revision or upgrade of the software available, then the Product Support policy described in previous section applies. For hardware, product discontinuance occurs when the hardware is no longer offered for general sale and there is no direct replacement hardware that will work with the currently supported software revision.

Whenever a product is discontinued, Varian's sales and service representatives will be notified of this discontinuance at least six months before the product is discontinued.

History of Star Workstation Development and Prior Releases

The Star Chromatography Workstation is a continuously evolving product. The history of the Star Workstation is outlined below.

LC Star Workstation Rev. A, released 12/88, introduced PC Workstations to the Varian product line and included control of up to three modules at a time including either a 9065 PDA or a 9050 monochromator-based detector.

LC Star Workstation Rev. A.1, released 2/89, provided an update for Windows 2.1 operation.

LC Star Workstation Rev. B, released 8/89, introduced ADC Board control and allowed control of up to four modules including two detectors.

LC Star Workstation Rev. B.1, released 11/89, provided a update that allowed the LC Star Workstation to work on faster 80386 based PCs and included 9001 support.

Star PC Integrator, released 1/90, introduced a low cost integrator, based on the LC Star Workstation, which controlled a single ADC Board. The synchronization interface cable was especially adapted for use with the 3400 & 3600 GCs.

LC Star Workstation Rev. C, released 8/90, introduced a version of the LC Star Workstation that worked with Windows 3.0, thus providing true concurrent operation of the Star Workstation with other applications such as Excel. This version allowed use of higher memory and fixed the Printer Driver problem encountered with earlier Windows 286 based Star Workstations.

GC Star Workstation Rev. A, released 5/91, introduced dedicated control of up to two 3400 GCs and a new configuration procedure.

LC Star Workstation Rev. C2, released 6/92, fixed problems with using faster 80486 based PCs with the GPIB and introduced a new installation procedure.

GC Star Workstation Rev A2, released 6/92, introduced the control of the 3600 GC and made two-instrument operation standard.

Four Instrument Upgrade to GC Star Workstation Rev. A2, released 10/92, introduced the ability to control up to four 3X00 GCs with the GC Star Workstation..

GC Star Workstation Version 3.0, released 8/93, introduced automated integrated multilevel calibration, multichannel results, a new installation procedure, and tabular editor for SampleLists, Sequences, and Peak tables.

Report and Normalized Percent update, released 9/93, fixed problems with the Report application and the way the Normalized Percent calculation type calculated results.

Star Chromatography Workstation Version 4.0, released 6/94, introduced integrated control of both LC and GC modules in a single Workstation, able to control up to four instruments.

Report update, released 7/94, fixes an esoteric problem with Report, and gives full support for the LaserJet IV printer.

Star Chromatography Workstation Version 4.01, released 2/95, corrected integration problems for small peaks riding on high baselines, increased the reliability of automated applications, and added control capabilities to support new LC hardware.

Star Chromatography Workstation Version 4.02, released 9/95, added support for Windows '95 and corrected problems observed when using an 'Inhibit Integrate' event to end a peak.

Star Chromatography Workstation Version 4. 5, released 2/96, added new calibration options, direct printing of chromatogram overlays, many ease-of-use improvements in the post-run applications, and control of the Mistral Column Oven.

Star Chromatography Workstation Version 5.0, released 4/98, was the first version designed for a 32 bit platform and introduced many new features, with a focus on ease of use.

Star Chromatography Workstation Version 5.1, released 6/98, was a maintenance release adding control of the SD-200, SD-300 and UV-1 LC modules.

Star Chromatography Workstation Version 5.2, released 8/98 was a maintenance release addressing minor fixes and ease of use issues.

Star Chromatography Workstation Version 5.3, released 1/99, provided control for the new ProStar modules, including a new version of the ProStar Dynamax.

Star Chromatography Workstation Version 5.31, released 3/99, was a maintenance release addressing mostly the PolyView2000 software.

Star Chromatography Workstation Version 5.31 Update#1, released 5/99, added control of the new ProStar 430 AutoSampler, and fixed high absorbance issues in the 330 PDA software.

Star Chromatography Workstation Version 5.31 Update#2, released 8/99, added control of the ProStar 420 AutoSampler, 2002 and 2003 Micro-GCs, and addressed some problems related to the Microsoft DDEML library.

Star Chromatography Workstation Version 5.50, released 4/00, added control for the ProStar 701 fraction collector, HP/HPX/HPXL pumps, the standalone CIM module, and introduces new features in the ProStar 330 software (real-time isogram). Also includes improved control of the Micro-GCs.

Star Chromatography Workstation Version 5.52, released 7/01, added control of the 3900 GC, integrates the existing ProStar 363 driver and provides support for the Star 800 Module Interface Box. It also provides compatibility with Windows 2000.

Star Chromatography Workstation Version 6.0, released 4/02, adds control of the ProStar 500 CVM, the ProStar 701 Fraction Collector, the 4002 Micro-GC, and the Combi PAL. It incorporates control of the ProStar 310 with digital filtering, the ProStar 364, the ProStar 520, and the 3900 GC. It also provides a platform for the optional Access Control and Audit Trail software.

Star Chromatography Workstation Version 6.2, released 4/03, improves control of the ProStar 500 CVM, the ProStar 701 Fraction Collector, and the Combi PAL. All software changes incorporated into Star 6.2 related to Access Control and Audit Trail software are fixes of problems (Bug fixes). No additional features have been added to the Access Control and Audit Trail software in Star 6.2. Therefore relative to 21 CFR 11 compliance, Star 6.2 is considered a "service pack" or patch and does not need revalidation.

Star Chromatography Workstation Version 6.3, released 3/04, is a maintenance release derived from MSWS 6.30, addressing issues related to operating under Windows XP from non-Administrator accounts, and several minor operational and ease-of-use issues. New rights were added to the Access Control and Audit Trail option, including the right to delete methods and MS-specific rights.

Star Chromatography Workstation Version 6.41, released 11/04, is a maintenance release derived from MSWS 6.41, integrating drivers for the new ProStar 325 and 335 detectors, and several minor operational and ease-of-use issues. All software changes related to Access Control and Audit Trail software are fixes of problems (Bug fixes). No additional features have been added to the Access Control and Audit Trail software in Star 6.41. Therefore, relative to 21 CFR 11 compliance, Star 6.41 is considered a "service pack" or patch and does not need revalidation.

Appendix A contains a description of subsequent releases up to the version you licensed. To determine which version of the Star Workstation software is installed on your PC use the 'About' command in the Help menu of any Star Workstation component. This brings up a dialog box listing your Star Workstation version number.

To determine what the latest version of the Star Workstation is, contact your Sales Representative or - if you have Internet access - consult the Varian Chromatography Systems World Wide Web site by using the 'Product Support Web Site' command in the Help menu of any Star Workstation component.

You can upgrade to the latest version of the Star Workstation from any prior version. The upgrade may require a change of operating system (if you are still using a 16 bit version of windows), and possibly a hardware upgrade as well if you old computer does not meet the new hardware requirements.

Validation of the Software Installation

Factory Installation and Checkout Procedure

The Star Workstation software is installed on every computer purchased from Varian Chromatography Systems for use as a Star Chromatography Workstation system. The Star Workstation software is installed from a source CD-ROM. Optional software and hardware is installed unless the order is ambiguous (as may be the case if multiple computers are listed on the order) or there are explicit instructions to not install specific options.

The PC's hard disk is optimized (defragmented).

The operator subsequently boots the PC from its fixed disk and verifies that the workstation software is properly installed by:

- Executing the Star Workstation Validation program (see next section) and checking that it reports no error.
- Invoking System Control and making certain it starts correctly and recognizes any installed boards. This completes the installation process.

Validation of Star Workstation Installation and Function

You can validate the installation and function of your Star Chromatography Workstation after installation on your hard disk. Two levels of validation checks are available:

- To establish that the software you are running is what was tested at the factory, a validation utility compares the checksums of the Star Workstation executable files to checksums produced at the factory during the software release process.
- To establish that the data handling results obtained at the factory can be reproduced in your environment, reprocess the tutorials and examples included with the Star Workstation and compare your results with the original results (See [Appendix A](#)).

Validation of the Installed Star Workstation

Starting with Version 5.0, the 'Varian Star ...' menu (accessed from Start → Program Files) contains a 'Validate installed files' menu item.

This utility uses a file containing checksums produced at the factory for all the executable files comprising your Star Workstation software. The utility documents any error encountered and can print a report. The report documents in detail the file system tested, the success or failure of the test, the date printed & tested, the ID of the tester (if login IDs are used on this computer) and can optionally contain signature lines. For more details on this application, refer to the on-line help once the test program is launched.

Maintenance Schedule of the host computer

As part of a set of GLP procedures, we recommend you plan to perform a preventative maintenance session on the host computer on a weekly schedule. You should determine what

procedure and schedule is appropriate for your own environment. We recommend including the following steps:

- Run a Virus Detection program with recently updated virus data.
- Run a Disk Test program, such as the 'ScanDisk' utility included in Windows 95.
- Run a Disk Defragmentation program, such as the 'Disk Defragmenter' utility included in Windows 95.

Data Documentation and Logs

Documentation of Data Collection: Operating Logs

Records of the conditions under which your run data was collected can be found in System Control/Automation logs. In addition, you can add notes for each sample you inject, and for each data recalculation performed by the Star Workstation.

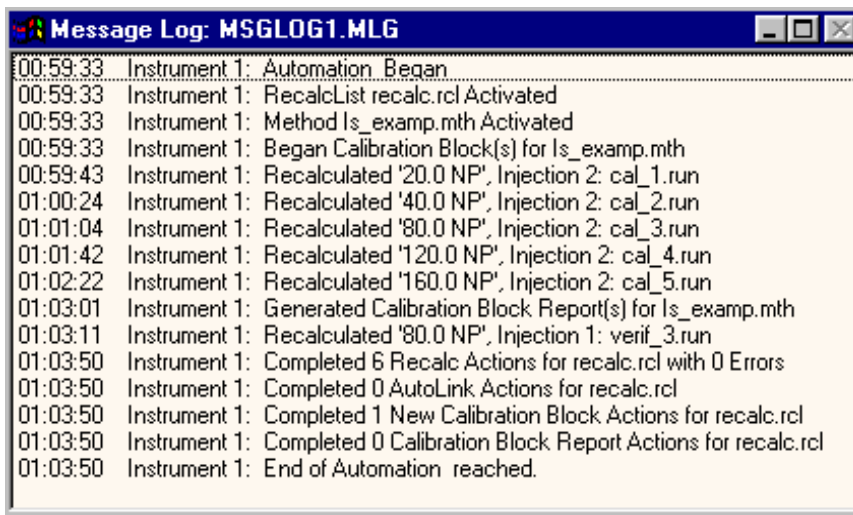
System Control/Automation maintains logs of operating conditions during data collection. Operating conditions are documented in the following places:

The **Configuration Log** contains messages generated for modules that are not assigned to an instrument. The Configuration Log is viewed by switching to the Configuration window (using the INSTRUMENT-> Configuration menu item), then selecting the Configuration Log window from the WINDOWS menu. You can also display the Configuration Log by double-clicking on the Configuration Log line at the bottom of the Configuration screen.

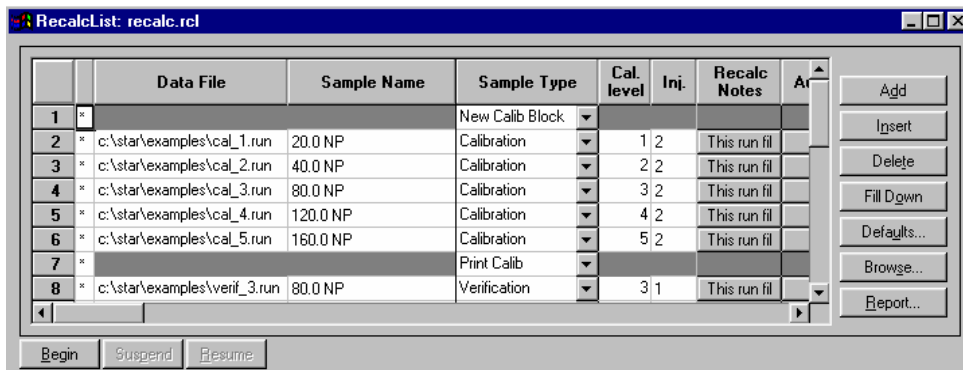
The **Message Log** contains messages generated during automation (execution of a Sequence List, a SampleList, a RecalcList, or a Single Injection). These messages include the names of Data Files generated, AutoLink applications run, Methods and SampleLists that were activated, Recalculation and Print Sequence actions, Pump Off/Lamps Off actions, and errors encountered, such as invalid methods and instrument faults. This information can be used to troubleshoot your system. There is a Message Log for each Instrument in System Control. The Message Log can be automatically printed during or at the end of the execution of a sequence.

As of Version 6.0, the **Message Log** is permanently stored and can be accessed at any time.

An example of a Message Log is displayed below:



The Star Workstation can optionally create a **RecalcList** when you run a SampleList. The RecalcList documents the names of the data files produced by the execution of the SampleList, along with the corresponding sample information. The RecalcList can be used later for batch printing and recalculations, either with the same method or with a different method. RecalcLists can also be created manually in System Control or in the Automation File Editor. An example of a RecalcList is shown below.



The **Run Log** is saved within the data file. It contains a listing of the methods and parameters used during the chromatographic run, including any changes that were made to the run conditions as data collection progressed. You can print the Run Log to verify conditions present during the sample run.

The **Instrument Log** contains the fault and cold start messages generated by the modules configured within an instrument. You can obtain a list of either recent faults or self-test faults.

The **Error Log** is saved as part of the data file, and contains a record of any errors which occurred during the data collection process. It can be printed to verify error-free data collection.

The **Revision Log**, also saved within the data file, contains a record of recalculations and data manipulations applied to each channel of data.

Documentation of Data Collection: Validation and Verification

Validation is the Star Chromatography Workstation's process of checking each Method, SampleList, or RecalcList to make certain it is valid. Validation is automatically performed when a sequence is run. Any invalid Method, SampleList, or RecalcList is indicated either by an error message box or by mention in the Message Log.

Verification is the process used to test whether the calibration curve is still accurate. The Star Workstation compares the calculated results for the standard to the amounts known to be in that standard. It calculates the difference between the known amount of each analyte and the result obtained from the current calibration curve. If the difference between the two results exceeds the Deviation Tolerance, a user-specified action takes place. The Deviation Tolerance and Out-of-Tolerance action are specified in the Calibration Verification dialog box in the Peak Table window of the Method.

Documentation Provided by the User

You can add notes about your sample and injection conditions before you start your sample run. You can also add notes about any recalculations you perform at a later time. These notes will be included in the data file report.

Injection Notes

If injections are being made automatically, you can add notes about each sample on your SampleList by clicking the on the button in the *Injection Notes* column in the **SampleList** window to open the **Notes** edit box. If you are making manual injections, you can do the same in the **Inject Single Sample** window. Notes inserted prior to the original data collection run are labeled *Original Notes* in your report. These notes become a permanent part of the data file, and cannot be changed later.

Recalculation Notes

Appended Notes can be added to your data file from your RecalcList when you perform recalculations on the data. These notes also become part of the data file report, but they change each time the file is reprocessed from a RecalcList.

Module Information Logs

Module Information Logs contain user-defined information about the hardware modules that comprise your chromatography systems. You can use this feature to keep track of any information relevant to your hardware components, from Serial Number to service information. These logs are maintained in the Star Workstation directory, as files named MODULEnn.MDF. At the end of the data acquisition, the module information logs for the appropriate components are permanently stored in the file along with the data. This information can be read from the data files, but not modified.

Data Files

The Data File (file with a .RUN extension) contains all of the raw chromatographic data, the most recent results calculated for each channel, the Run Log listing the instrument method and data handling sections, any notes entered in the Sample/RecalcList, errors, and the last recalculation parameters. Because the data file is a repository of all information about the data, its size will change every time it is reprocessed.

As of Star Version 6.0, if Access Control and Audit Trail software is installed, data files will not have results overwritten when a new recalculation is done. Instead, a version of the results and data handling method is added to the data file and older calculations can be accessed.

Method Files

The Method File (file with a .MTH extension) contains all of the instrument and data handling method sections required to process a set of samples. While most sections are only modified when edited by the user, a method may also contain sections that are modified under automation. Such sections contain 'historical' information that extends in scope beyond a single run, for example:

Blank Baseline:

The data from a designated sample is reduced to simplified baseline information and stored in the method, to be subtracted during the data handling from data files subsequently processed by the same method. The blank baseline information is incorporated into each data file as it is processed, along with the rest of the data handling method.

Calibration:

During a calibration, the method is updated based on new calibration data points, building the calibration curve one point at the time. Again, the calibration information is incorporated into data files as they are processed.

As a consequence of these automatic updates, methods can change size in the absence of any user edits.

Methods can be password-protected (the password is controlled by the security application, which is itself password-protected).

Methods can contain revision notes, and the system can be set-up (via the security application) to require a revision entry to be made when the method is changed.

As of Star Version 6.0, if Access Control and Audit Trail software is installed, when methods are modified, a new version of the method will be added to the file. Therefore, older versions of the method can be accessed.

Data Processing

Processing Raw Data: Data Collection

All the data generators (including the ADC Board) used with the Star Workstation are designed to acquire data based on their own precise time base and provide at least 20 seconds of buffering at their highest data rate. The service of the data generators (i.e. the transfer of data to the hard disk) is the highest priority task of the Star Workstation. As a result, data acquisition will be reliable even in a significantly overloaded system. If the computer became so busy that some of the data generators could not be serviced in a timely manner, the system would report error messages and shut down. No incorrect data can be produced as a result of an overload.

The data from a given data generator is saved into a single file, which can have one or more channels of data. After the data acquisition is completed, post-run processes are automatically performed, such as Peak Processing and the printing of Reports.

Peak Processing

Peak Processing transforms raw chromatographic data into useful results. It is controlled by the data handling section of the data file method, and is performed automatically at the end of every chromatographic run. Raw data can also be *recalculated* later in System Control or Interactive Graphics with a different method to adjust the processing parameters or focus on a different aspect of the data. The information used to recalculate the results is then saved in the data handling section of the data file method. The first section of this manual is devoted to a discussion of data handling fundamentals.

There are three basic peak processing functions:

- Peak detection and integration
- Peak area and height allocation
- Peak identification and results calculations

Parameters Affecting Peak Detection

Peaks are detected based on the slope of the chromatogram. The main method parameters affecting the peak detection are the Noise Options, the Peak Width settings, and the Signal-to-Noise ratio. Other parameters such as Inhibit Integrate (II), Solvent Reject (SR), Force Peak (FP), Split Peak (SP), Tangent Percent (TP) also affect peak detection.

Evaluation of the Chromatogram's Slope

The slope is evaluated by a seven point Savitzky-Golay convolution filter¹ which is equivalent to performing a least squares fit of a cubic polynomial on the seven points and determining the slope of the fitted curve at the center point.

Noise Options

The noise is monitored prior to the start of each run. The Monitor Length in the Detector Information Window determines the number of data points to be acquired, and indicates the total length of the monitoring period. At the end of the data acquisition, a noise value is computed and stored in the data file for each channel of data acquired. The noise is computed as the Root Mean Square (RMS) value of the slope.

The Data Handling Method gives you the options of using:

- The noise measured in the file.
- The noise measured in the first file processed by the method.

¹ E. Savitzky and M. Golay, 1627 Anal. Chem. 36, 1964

- A fixed value of your choice.

Note that it is possible for your instrument to report a zero noise reading. This may happen if you use a high bunch rate setting, or if a low gain is used on a detector connected to an ADCB (in which case, both signal and noise are decreased, and the noise may not be measurable).

Peak Width Setting

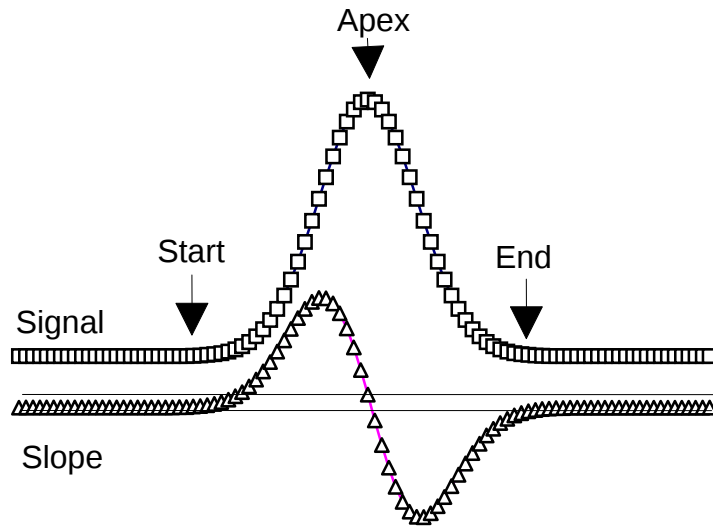
The width of chromatographic peaks vary greatly based on experimental conditions, and may also vary as a function of the retention time in a given experiment. To accommodate these variations, the data points stored in the data file may be further bunched (averaged) by the peak processing software so that the current peak width setting translates into a number of bunched points between 10 and 20 (the bunch factor is a power of 2). This bunching process affects the measured slope and it is important that the width setting be the right order of magnitude. If the Time Events table does not specify a peak width setting (WI event), the software will start with a peak width of 4 seconds and automatically update the peak width and the corresponding bunch rates based on the measured width of the peaks encountered. If peak width settings are specified, no automatic width update will occur. Instead, the software will use the specified time-programmed width values and deduce the corresponding bunch rates. In either case, width updates will not occur during a peak, and will only occur after the end of a peak.

Signal-to-Noise (S/N) Ratio

The Star Workstation compares the measured slope of the chromatogram to the noise multiplied by this S/N ratio. Exceeding this threshold indicates the onset of a peak.

Elements of Peak Detection

Peak sensing is a complex proprietary algorithm which cannot be fully described here. The following is a synopsis of the detection of an isolated peak.



Peak Start

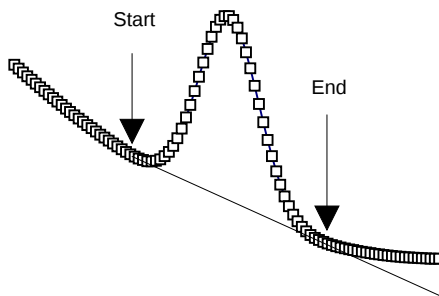
When the slope exceeds the threshold, a peak begins. The last 60 bunched points are examined backwards to find a local minimum, which is then taken as the Peak Start.

Peak Apex

A peak apex is found when the slope becomes negative. The time of the apex is determined by a linear interpolation between the last point with a positive slope and the first point with a negative slope. The amplitude of the apex is determined as the highest of the two corresponding bunched data points.

Peak End

A peak end is found when the slope reverts to zero after the apex, or (in the case of a negative drift) when the chromatogram rises above the lowest tangent drawn from the start point.



The Integration Baseline

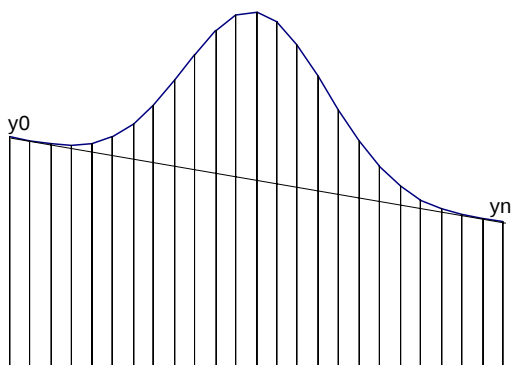
The line between the Start and End points is the **integration baseline**.

Calculation of Peak Size

The Star Workstation will calculate peak size in terms of peak area, peak height, or square root of peak height.

Calculation of Peak Area

Peak Areas are computed by adding the areas of trapezoids created by dropping perpendiculars from each data point to the zero value. Then, the area of the trapezoid comprised between the integration baseline and the zero line is subtracted.



If $y_0, \dots, y_n$ are the amplitudes of the $n+1$ points defining n small trapezoids, and Δt is the time interval between two consecutive points, in seconds, the areas can be computed as follows:

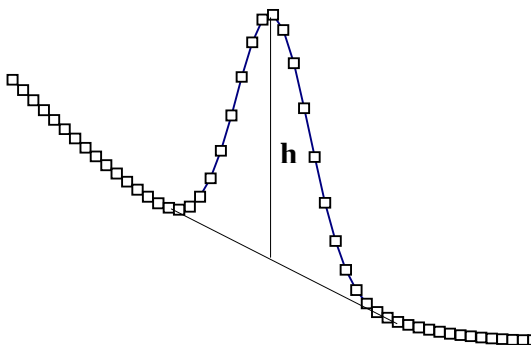
Total Area of small trapezoids: $\Delta t ((y_0 + y_n)/2 + y_2 + y_3 + \dots + y_{n-1})$

Area of Baseline trapezoid: $(y_0 + y_n) * n \Delta t / 2$

Peak Area: $\Delta t (y_1 + y_2 + \dots + y_{n-1} - (y_0 + y_n) (n-1)/2)$

Calculation of Peak Height

The Star Workstation measures the peak height as the vertical height of the apex above the integration baseline. You have the option of reporting the peak height or the square root of the peak height.



Verifying the Peak Size

The Area counts or Height counts in the Star Workstation report are based on an arbitrary characteristic of the data generator called the Least Significant Bit or LSB. The LSB represents the smallest data increment that can be measured.

Height counts are expressed in LSB

*Area counts are expressed in LSB*seconds*

To express the results in microunits (μ AU or μ Volts) or microunits*sec, area and height in counts must be multiplied by the proper LSB value. Starting with Version 6.0, the LSB value for the current channel of the current data file is shown in the Standard Report, after the peak results and next to the "Baseline Offset" field.

Using these conversion factors, you may independently verify that results are correctly computed by exporting the raw data points to a text file, using the "Convert Data to ASCII" program provided in the Advanced Applications Software, then using the previously indicated formulas to measure height or area using a spreadsheet program such as Microsoft Excel.

An acceptable error on the peak area should be equal to the peak to peak noise multiplied by the time difference between peak end and peak start.

A simple way to verify that Peak Areas are globally correct is to compare the peak area to the theoretical value of

$$1.06 \text{ height} * \text{width at half height}$$

which characterizes a Gaussian peak. This formula is not very accurate if the peak is lost in the noise or is asymmetrical.

Subtract Blank Baseline

If you choose, you can have the Star Workstation subtract a blank baseline from the chromatogram data. This allows you to correct for reproducible baseline drift over the chromatogram, which can result from temperature programming in GC analyses or elution gradient programming in LC analyses. The blank run is reduced to a set of 257 to 512 data points by a bunching process and is subtracted from the analysis run point-by-point after linear

interpolation, prior to peak detection. In the report, both the chromatogram and the results reflect this correction. In most cases, peak heights and areas should not be significantly affected by the use of a blank baseline.

Peak Identification



Peaks in a chromatogram are mapped to compound names by means of the peak table contained in the method. Identification of each compound is based on its retention time. You will find more information about peak identification in the Data Handling and Reports Operation Manual for Star Workstation Version 6.

Replicates and Calibration Levels

One or more concentration **levels** are used to define a calibration curve, and the Star Workstation supports the use of up to 10 levels. Within each level, multiple replicates can be used. The Star Workstation offers two options for the treatment of replicates.

Average Replicates: Each level is represented by one 'historical' value. Each new replicate is averaged in with the historical value with a method defined weight to produce a new historical value. For example, if the averaging factor is set to 30%:

$$\text{newHistoricalPeakSize} = 30\% \text{ PeakSize} + 70\% \text{ oldHistoricalPeakSize}$$

On the first replicate at each level in a new calibration, the historical peak size is set to the Peak Size.

Keep Replicates Separate: Each replicate within the level is kept as a separate data point. Up to ten replicates per level are allowed. The calibration curve and the calculations will include all replicates.

Weighted Regression

Regression weights can be attached to each data point based on how much confidence one can have in the validity of the data point. Statistical textbooks often use a weight of $1/\sigma_i^2$, i.e., inversely proportional to the variance of the data point. In chromatography, the most common options are to use no weights ($w=1$), the inverse of the amount or concentration ($w=1/\text{Amt}$), or the inverse of the square of the amount ($w=1/\text{Amt}^2$). The Star Workstation has combined these 3 options with a $1/n$ option (where n is the number of replicates for the level) which suppresses any bias due to a possibly different number of replicates in each level. The weight options are therefore 1, $1/n$, $1/\text{Amt}$, $1/(n*\text{Amt})$, $1/\text{Amt}^2$, $1/(n*\text{Amt}^2)$. The weight for point i will be noted w_i in the rest of this manual.

Calculations Performed by the Star Workstation

The formulas used in the Star Workstation's calculations are available for you to independently verify their results. The Star Workstation uses the **least squares** method to determine the coefficients of the equation which approximates the calibration curve.

You can calculate the coefficients for your calibration equation independently, using an Excel worksheet or by hand. If you use an Excel worksheet, bear in mind that Excel may display a negligible exponential value rather than zero when numbers are displayed in the 'General' format. If appropriate, you can solve this problem by selecting the 'Numeric' format.

Determination of the Calibration Curve Equation

The general forms of the calibration equations used by the Star Workstation are:

Linear	$Y = a_1X + a_0$
--------	------------------

Quadratic	$Y = a_2X^2 + a_1X + a_0$
-----------	---------------------------

Cubic	$Y = a_3X^3 + a_2X^2 + a_1X + a_0$
-------	------------------------------------

The determination of the calibration curve is also affected by the origin treatment, i.e. assumptions made regarding to the behavior of the calibration curve around $X=0$ and $Y=0$. The three options are to **Ignore**, **Include**, or **Force** the origin.

Include: Add the origin to the data set.

Force: Force the curve to go through the origin. This means that the coefficient a_0 is null.

Ignore: Do not make any assumption about the origin.

The number of distinct calibration levels determines the maximum degree of the calibration equation you can use. The least squares technique approximates your calibration curve equation by fitting a polynomial to a given data set $\{(x_i, y_i) | i=0, \dots, M\}$. The data set is composed of $M+1$ coordinate pairs (referred to as "datapoints" in what follows). The x coordinate is the concentration level of your sample, and the y coordinate is the measured size (height or area) of the compound's elution peak at the given level. The degree, n , of your selected polynomial must be less than the number of datapoints. The approximation requires $n+1$ equations in $n+1$ unknowns, a_j . As the number of datapoints must be greater than the degree of the derived polynomial, **at least** two pairs of coordinates (with different values of x , i.e., levels) are needed to generate a linear equation, three pairs of coordinates are needed to generate a quadratic equation, and four pairs of coordinates are needed to generate a cubic equation. If the Include or Force options are used, one fewer pair of real coordinates is needed, because there is one fewer coefficient to determine (*Force* option) or one extra pair of coordinates is artificially added (*Include* option).

The STAR\EXAMPLES directory for your Star Workstation contains data files CAL_1.RUN through CAL_5.RUN, providing data files with five different concentration levels, allowing you to examine linear, quadratic, and cubic calibration curves. We recommend, however, that you use your own set of data to do your own validation, so that you truly establish that the product operates satisfactorily in your own conditions.

The **Coefficient of Determination** measures how well the data conforms to the curve model chosen. The closer its value is to 1.00, the better the agreement with the model (see algorithm description after the numerical examples of least squares fit). However, the approximation is always better as the degree of the fit increases; therefore you may want to use the lowest degree that gives you a satisfactory fit. You can view the calibration curves stored in a method by pushing the **View Curves...** button in the **Calibration Setup** section of a Data Handling method. The Coefficient of Determination is displayed in the Calibration Curve window for each peak.

Note that if you use the minimum number of calibration levels required by a certain type of fit, for example, 2 levels for a linear fit, the correlation coefficient will be 1.00. This does not establish that the calibration curve perfectly follows the polynomial model, but rather that the number of levels used is insufficient to estimate a goodness of fit, and adding one intermediate level will certainly result in a less impressive fit. We suggest you always use a number of levels exceeding the minimum required by the model selected.

To derive the equations used to determine the polynomial fit, one must first express the total residual error U for the data set as the sum over all the points of the square of the difference between the actual y_i values and the value of the polynomial at $x = x_i$.

$$U = \sum_{i=0}^M w_i \left(y_i - a_0 - a_1 x_i^1 - a_2 x_i^2 - a_3 x_i^3 - \dots \right)^2$$

U is a function of $a_0, a_1, \dots$. The least squares solution is a set of values $a_0, a_1, \dots$ such that U shows a local minimum relative to all the a_k coefficients. This leads to $n+1$ equations for a n th degree polynomial, corresponding to:

$$\frac{dU}{da_0} = 0, \quad \frac{dU}{da_1} = 0, \quad \dots$$

Least Squares Linear Fit

The following two equations determine the first degree polynomial $y = a_1 x + a_0$ which best approximates the data:

$$a_0 \sum_{i=0}^M w_i x_i^0 + a_1 \sum_{i=0}^M w_i x_i^1 = \sum_{i=0}^M w_i x_i^0 y_i$$

$$a_0 \sum_{i=0}^M w_i x_i^1 + a_1 \sum_{i=0}^M w_i x_i^2 = \sum_{i=0}^M w_i x_i^1 y_i$$

Notes applying to the above formulas, and the formulas for the other degrees as well:

In the above formulas, x_i^0 can be replaced by 1; therefore,

$$\sum_{i=0}^M w_i x_i^0 \text{ can be replaced by } \sum_{i=0}^M w_i \text{ (which is } M+1 \text{ if all}$$

weights are taken as 1), and $\sum_{i=0}^M w_i x_i^0 y_i$ by $\sum_{i=0}^M w_i y_i$. The

x_i^0 notation was only used to better show the symmetry of the equations.

If the origin is forced, a_0 is null; all terms factored by a_0 disappear, and the first equation (which corresponds to the derivation of U relative to a_0) also disappears.

Least Squares Quadratic Fit

The following three equations determine the second degree polynomial $y=a_2x^2+a_1x+a_0$ which best approximates the data:

$$a_0 \sum_{i=0}^M w_i x_i^0 + a_1 \sum_{i=0}^M w_i x_i^1 + a_2 \sum_{i=0}^M w_i x_i^2 = \sum_{i=0}^M w_i x_i^0 y_i$$

$$a_0 \sum_{i=0}^M w_i x_i^1 + a_1 \sum_{i=0}^M w_i x_i^2 + a_2 \sum_{i=0}^M w_i x_i^3 = \sum_{i=0}^M w_i x_i^1 y_i$$

$$a_0 \sum_{i=0}^M w_i x_i^2 + a_1 \sum_{i=0}^M w_i x_i^3 + a_2 \sum_{i=0}^M w_i x_i^4 = \sum_{i=0}^M w_i x_i^2 y_i$$

Least Squares Cubic Fit

The following four equations determine the third degree polynomial $y=a_3x^3+a_2x^2+a_1x+a_0$ which best approximates the data:

$$a_0 \sum_{i=0}^M w_i x_i^0 + a_1 \sum_{i=0}^M w_i x_i^1 + a_2 \sum_{i=0}^M w_i x_i^2 + a_3 \sum_{i=0}^M w_i x_i^3 = \sum_{i=0}^M w_i x_i^0 y_i$$

$$a_0 \sum_{i=0}^M w_i x_i^1 + a_1 \sum_{i=0}^M w_i x_i^2 + a_2 \sum_{i=0}^M w_i x_i^3 + a_3 \sum_{i=0}^M w_i x_i^4 = \sum_{i=0}^M w_i x_i^1 y_i$$

$$a_0 \sum_{i=0}^M w_i x_i^2 + a_1 \sum_{i=0}^M w_i x_i^3 + a_2 \sum_{i=0}^M w_i x_i^4 + a_3 \sum_{i=0}^M w_i x_i^5 = \sum_{i=0}^M w_i x_i^2 y_i$$

$$a_0 \sum_{i=0}^M w_i x_i^3 + a_1 \sum_{i=0}^M w_i x_i^4 + a_2 \sum_{i=0}^M w_i x_i^5 + a_3 \sum_{i=0}^M w_i x_i^6 = \sum_{i=0}^M w_i x_i^3 y_i$$

The following pages provide numerical examples of such computations for an example of calibration including five datapoints. The data on which the calculations are based is summarized in two tables. To simplify the example, all weights were taken as 1. If you want to carry out the calculation with

weights, multiply each line of the table by the appropriate weight for this data point. As an example, if $x=12$ and you chose the $1/\text{Amt}^2$ weight option, all values in line 1 should be multiplied by $1/144$.

Table 2 contains the X values for all data points, the corresponding X^n values and their sums. The second column, headed X^1 , contains the amount of analyte for each run, as entered in the peak table. Line 0 is an optional line corresponding to the origin, in the event you include the origin in the computations. The only sum affected is the total weight, which is 5 for the Ignore and Force origin option, and 6 for the Include option.

i	X^0	X^1	X^2	X^3	X^4	X^5	X^6
0	1	0	0	0	0	0	0
1	1	12	144	1728	20736	248832	85984
2	1	40	1600	64000	2560000	102400000	4096000000
3	1	80	6400	512000	40960000	3276800000	262144000000
4	1	120	14400	1728000	207360000	24883200000	2985984000000
5	1	140	19600	2744000	384160000	53782400000	7529536000000
6							
7							
8							
9							
SUM	5 or 6	392	42144	5049728	635060736	82045048832	10781762985984
	$\sum_{i=0}^M X_i^0$	$\sum_{i=0}^M X_i^1$	$\sum_{i=0}^M X_i^2$	$\sum_{i=0}^M X_i^3$	$\sum_{i=0}^M X_i^4$	$\sum_{i=0}^M X_i^5$	$\sum_{i=0}^M X_i^6$

Table 1

Table 3 contains the Y values for all data points, the corresponding $Y \cdot X^n$ products and their sums. The second column, headed Y_i , is the measured response for each injection of analyte. Line 0 corresponds to the origin; none of the sums is affected by the origin option. Here again, to carry out a calculation using weights you would multiply each line by the weight for the corresponding data point.

i	Y_i	$Y_i X_i^0$	$Y_i X_i^1$	$Y_i X_i^2$	$Y_i X_i^3$
0	(0)	0	0	0	0
1	7187	7187	86244	1034928	12419136
2	12844	12844	513760	20550400	822016000
3	21436	21436	1714880	137190400	10975232000
4	27556	27556	3306720	396806400	47616768000
5	31440	31440	4401600	616224000	86271360000
6					
7					
8					
9					
SUM	100463	100463	10023204	1171806128	145697795136
	$\sum_{i=0}^M y_i x_i^0$	$\sum_{i=0}^M y_i x_i^0$	$\sum_{i=0}^M y_i x_i^1$	$\sum_{i=0}^M y_i x_i^2$	$\sum_{i=0}^M y_i x_i^3$

Table 2

You can now use these values to determine the least squares solution of the linear, quadratic, and cubic calibration curves.

Numerical example for Linear Curve Fit

After substitution of the summation values from Tables 2 and 3, the equations to be solved for the linear calibration equation become:

Ignore Option

$$\begin{array}{rcl} 5a_0 + & 392a_1 = & 100463 \\ 392a_0 + & 42144a_1 = & 10023204 \end{array}$$

Include Option

$$\begin{array}{rcl} 6a_0 + & 392a_1 = & 100463 \\ 392a_0 + & 42144a_1 = & 10023204 \end{array}$$

Force Option

$$42144a_1 = 10023204$$

You can either use your own tools to solve this system of equations, or use the general method described with the solution of the cubic fit, placing these equations in a matrix and using Excel to solve for a_0 and a_1 .

For the include option, the solution is:

$$y = 209.251x + 3072.760$$

Using the Star Workstation, the coefficients were:

$$y = 209.26x + 3072.99$$

Such differences are not significant, and stem from the fact that only the integer part of the area count is listed on the reports, omitting a (negligible) decimal portion used by the Star Workstation (For example 7187 vs. 7187.13).

Numerical example for Quadratic Curve Fit

After substitution of the summation values from Tables 2 and 3, the equations to be solved for the quadratic calibration equation become:

Ignore Option

$$\begin{array}{rclcl}
 5a_0 + & 392a_1 + & 42144a_2 = & 100463 \\
 392a_0 + & 42144a_1 + & 5049728a_2 = & 10023204 \\
 42144a_0 + & 5049728a_1 + & 635060736a_2 = & 1171806128
 \end{array}$$

Include Option

$$\begin{array}{rclcl}
 6a_0 + & 392a_1 + & 42144a_2 = & 100463 \\
 392a_0 + & 42144a_1 + & 5049728a_2 = & 10023204 \\
 42144a_0 + & 5049728a_1 + & 635060736a_2 = & 1171806128
 \end{array}$$

Force Option

$$\begin{array}{rclcl}
 & 42144a_1 + & 5049728a_2 = & 10023204 \\
 & 5049728a_1 + & 635060736a_2 = & 1171806128
 \end{array}$$

You can either use your own method to solve this system of equations, or use the general method described with the solution of the cubic fit, placing these equations in a matrix and using Excel to solve for a_0 and a_1 .

Numerical Example for Cubic Fit

After substitution of the summation values from Tables 2 and 3, the equations to be solved for the cubic calibration equation become:

Ignore Option

$$\begin{array}{rclcl}
 5a_0 + & 392a_1 + & 42144a_2 + & 5049728a_3 = & 100463 \\
 392a_0 + & 42144a_1 + & 5049728a_2 + & 635060736a_3 = & 10023204 \\
 42144a_0 + & 5049728a_1 + & 635060736a_2 + & 82045048832a_3 = & 1171806128 \\
 5049728a_0 + & 635060736a_1 + & 82045048832a_2 + & 10781762985984a_3 = & 145697795136
 \end{array}$$

Include Option

$$\begin{array}{rclcl}
 6a_0 + & 392a_1 + & 42144a_2 + & 5049728a_3 = & 100463 \\
 392a_0 + & 42144a_1 + & 5049728a_2 + & 635060736a_3 = & 10023204 \\
 42144a_0 + & 5049728a_1 + & 635060736a_2 + & 82045048832a_3 = & 1171806128 \\
 5049728a_0 + & 635060736a_1 + & 82045048832a_2 + & 10781762985984a_3 = & 145697795136
 \end{array}$$

Force Option

$$\begin{array}{rclcl}
 42144a_1 + & 5049728a_2 + & 635060736a_3 = & 10023204 \\
 5049728a_1 + & 635060736a_2 + & 82045048832a_3 = & 1171806128 \\
 635060736a_1 + & 82045048832a_2 + & 10781762985984a_3 = & 145697795136
 \end{array}$$

The above equations are simultaneous. They can be solved for a_0 , a_1 , a_2 , and a_3 using matrix algebra and following the rules for manipulation of simultaneous equations. You can use Excel to set up the matrices and find the value of these coefficients. The following example applies to the set of equations for the *Include Origin* option.

Set up the data from these equations in matrix form:

Row	a_0	a_1	a_2	a_3	
1	6	392	42144	5049728	100463
2	392	42144	5049728	635060736	10023204
3	42144	5049728	635060736	82045048832	1171806128
4	5049728	635060736	82045048832	10781762985984	145697795136

Table 3

Divide by 6 to make a_0 1 in row 1:

Row	a_0	a_1	a_2	a_3	
1	1	65.33	7024	841621.33	16743.83
2	392	42144.00	5049728	635060736	10023204
3	42144	5049728.00	635060736	82045048832	1171806128
4	5049728	635060736.00	82045048832	10781762985984	145697795136

Table 4

Eliminate a_0 in rows 2, 3, & 4 by multiplying row 1 by the value of a_0 in each row and subtracting it from each row:

Row	a_0	a_1	a_2	a_3	
1	1	65.33	7024	841621.33	16743.83
2	0	16533.33	2296320	305145173.33	3459621.33
3	0	2296320.00	339041280	46575759360.00	466154016.00
4	0	305145173.33	46575759360	6531804173653.33	61145991125.33

Table 5

Divide to reduce a_1 to 1 in row 2:

Row	a_0	a_1	a_2	a_3	
1	1	65.33	7024	841621.33	16743.8333
2	0	1.00	138.8903	18456.3613.00	209.2513
3	0	2296320.00	339041280	46575759360.00	46154016.00
4	0	305145173.30	46575759360	6531804173653.33	61145991125.33

Table 6

Eliminate a_1 in rows 1, 3, & 4 by multiplying row 2 by the value of a_1 in each row and subtracting:

Row	a_0	a_1	a_2	a_3	
1	1	0	-2050.17	-364194.27	3072.7.75
2	0	1	138.8903.00	18456.36	209.25
3	0	0	20104654.00	4194047801.81	-14353906.99
4	0	0	4194047801.81	899934608615.23	-2706030130.37

Table 7

Divide to reduce a_2 to 1 in row 3:

Row	a_0	a_1	a_2	a_3	
1	1	0	-2050.17	-364194	3072.749
2	0	1	138.89	18456.36	209.2513
3	0	0	1.00	208.61	-0.71396
4	0	0	4194047801.81	899934608615.23	-2706030130.37

Table 8

Eliminate a_2 in rows 1, 2, & 4 by multiplying row 3 by the value of a_2 in each row and subtracting:

Row	a_0	a_0	a_0	a_3	
1	1	0	0	63492.84	1609.013
2	0	1	0	- 10517.66	308.41
3	0	0	1	208.61	-0.71396
4	0	0	0	2510992983.21	288349714.28

Table 9

Divide to reduce a_3 to 1 in row 4:

Row	a_0	a_0	a_0	a_3	
1	1	0	0	63492.84	1609.013
2	0	1	0	- 10517.7	308.4133
3	0	0	1	208.6108	-0.71396
4	0	0	0	1	0.011529

Table 10

Eliminate a_3 in rows 1, 2, & 3 by multiplying row 4 by the value of a_3 in each row and subtracting:

Row	a_0	a_1	a_2	a_3	
1	1	0	0	0	877.0087
2	0	1	0	0	429.6706
3	0	0	1	0	-3.11902
4	0	0	0	1	0.011529

Table 11

Table 12 shows $a_0 = 877.0087$, $a_1 = 429.6706$, $a_2 = -3.11902$, and $a_3 = 0.011529$. The cubic equation $y = a_3x^3 + a_2x^2 + a_1x + a_0$ can now be written : $y = 0.011529x^3 - 3.11902x^2 + 429.6706x + 877.067$

As an example, when the Star Workstation performed the same calculation it derived the following results:

	x^3	x^2	x^1	Intercept
Excel	0.011529	-3.11902	429.6706	877.067
Star Workstation	0.011529	-3.11912	429.688	877.067

The above method is suitable for manual calculations. If you use Excel to solve the problem, you can use a faster matricial solution by computing ratios of determinants, which Excel can calculate with the MDETERM () function.

Solution of the Calibration Equation During Analysis

Once the Star Workstation software has determined the equation for the calibration curve, it can use the equation to estimate the amount of analyte in an analysis run.

Solution of Linear Calibration Equations

Solution of a linear equation, $Y = mX + b$, is straightforward:

$X = (Y-b)/m$, where m is the calibration coefficient, Y is the measured response of the sample, and b is the Y intercept.

Solution of Quadratic Calibration Equations

When you want to solve a quadratic equation, put it into the form

$ax^2 + bx + (c-y) = 0$ and solve the equation using the quadratic formula:

$$x = \frac{-b \pm \sqrt{b^2 - 4a(c - y)}}{2a}$$

A quadratic equation has zero to two real solutions for x . The Star Workstation picks the smallest positive root, or if none is found, the largest negative root. If no solution is found within the calibration range, or 2 solutions exist within the calibration range, error messages are reported.

Solution of Cubic Calibration Equations

To validate the Star Workstation's solution of a cubic equation, take the value of X derived by the Star Workstation, substitute it into the calibration curve equation, and compute Y . For example, if the calibration equation is

$$Y = 0.000632X^3 - 0.41593X^2 + 237.87X + 4280.22$$

and in an analysis report using this calibration curve, a peak with a reported size of 7187 area counts corresponds to a calibrated result of 12.489, you can validate the Star Workstation's solution by computing the polynomial for

$$X=12.489:$$

$$Y = 0.000632(12.489)^3 - 0.41593(12.489)^2 + 237.87(12.489) + 4280.22$$

$$Y = 7186.22 \quad (\text{Compare to the original peak size of 7187})$$

If the above approach is not adequate for your GLP requirements, the following process describes the solution of a cubic calibration equation, $Y = AX^3 + BX^2 + CX + k$. There may be one, two or three real roots. The Star Workstation looks for the smallest positive root, or the largest negative root if there is no positive root. It will give an error message if the root is not within the calibration range, or if there are multiple roots in the calibration range. To derive the value of X:

Subtract the value Y from both sides of the equation.

$$(1) 0 = AX^3 + BX^2 + CX + (k - Y)$$

Let $D = (k - Y)$

$$(2) AX^3 + BX^2 + CX + D = 0$$

Let $p = B/A$, $q = C/A$, and $r = D/A$

$$(3) X^3 + pX^2 + qX + r = 0$$

Let $X = (x - p/3)$, and substitute this value in equation 3.

Then it is possible to reduce the equation to the normal form:

$$(4) x^3 + ax + b = 0$$

where

$$a = (3q - p^2)/3, \quad b = (2p^3 - 9pq + 27r)/27$$

The three roots of a cubic equation, x_1, x_2, x_3 are:

$$x_1 = A + B, \quad x_2, x_3 = -\frac{1}{2}(A + B) \pm \frac{i\sqrt{3}}{2}(A - B),$$

where

$$i^2 = -1, \quad A = \sqrt[3]{\frac{-b}{2} + \sqrt{\frac{b^2}{4} + \frac{a^3}{27}}}, \quad B = \sqrt[3]{\frac{-b}{2} - \sqrt{\frac{b^2}{4} + \frac{a^3}{27}}}$$

If p, q, and r are real, and therefore a and b are real, the Star Workstation software examines the value of the expression

$$\frac{b^2}{4} + \frac{a^3}{27}$$

If (case 1) $\frac{b^2}{4} + \frac{a^3}{27} > 0$, there are one real, and two conjugate imaginary roots.

If (case 2) $\frac{b^2}{4} + \frac{a^3}{27} = 0$, there are three real roots, of which at least two are equal.

If (case 3) $\frac{b^2}{4} + \frac{a^3}{27} < 0$, there are three real and unequal roots.

If **case 1** is true, the value of this expression is greater than 0, and the value of a is greater than 0, the **real** root of the equation will be

$$(5) \ x = 2\sqrt{a/3} \operatorname{ctn} 2\Phi$$

where

$$(6) \ \operatorname{ctn} 2\psi = \mp \sqrt{\frac{b^2}{4} / \frac{a^3}{27}}$$

The sign of the cotangent in (6) is negative if b is positive and positive if b is negative. The value of Φ is derived from the relationship

$$(7) \ \tan \Phi = \sqrt[3]{\tan \psi}$$

If **case 2** is true, the roots are

$$x = \mp 2\sqrt{\frac{-a}{3}}, \pm\sqrt{\frac{-a}{3}}, \pm\sqrt{\frac{-a}{3}}$$

If this expression fits **case 3**, the **real** roots can be found more easily using the following formulas:

$$X_k = 2\sqrt{\frac{-a}{3}} \cos\left(\frac{\phi}{3} + 120^\circ k\right), k = 0,1,2$$

where

$$\cos\phi = \mp \sqrt{\frac{\frac{b^2}{4}}{\left(\frac{-a^3}{27}\right)}}$$

the upper sign is used if b is positive, and the lower sign is used if it is negative.

Example of the Solution of a Cubic Equation During Analysis

Using the previous example, the cubic calibration equation is

$$Y = 0.000632X^3 - 0.41593X^2 + 237.87X + 4280.22$$

and the peak size in the analysis run is 7187.

Place this equation in the form

$$AX^3 + BX^2 + CX + D = 0$$

by subtracting the value of Y from both sides of the equation.

$$0.000632X^3 - 0.41593X^2 + 237.87X - 2906.78 = 0$$

Divide each coefficient by the value of A .

$$X^3 - 65.8117X^2 + 376376.6X - 4599335 = 0$$

Let $p=B/A$, $q=C/A$, and $r=D/A$. Then

$$(8) \quad X^3 + pX^2 + qX + r = 0$$

Let $X = x - p/3$. Substitute this value in equation (8), then call the coefficient of x and the y intercept a and b respectively. The equation then becomes

$$(9) \quad x^3 + ax + b = 0$$

where

$$(10) \quad a = (3q - p^2)/3$$

$$(11) \quad b = (2p^3 - 9pq + 27r)/27$$

when these expressions are evaluated,

$$a = 374932.9$$

$$b = 3636212$$

The expression $b^2/4 + a^3/27$ has a value of $1.95538E+15$. The expression fits case 1. When this expression has a value greater than zero, and a is also greater than zero, equations (5), (6), and (7) can be used to derive the value of x . Excel can be used to compute this value:

$$x = 2\sqrt[3]{a/3} \operatorname{ctn} 2\Phi = 2(353.5217) \operatorname{ctn} 2\Phi$$

$$\operatorname{ctn} 2\psi = -\sqrt{3.30551E + 12 / 1.95538E + 15} = -0.04115$$

$$\tan \psi = \tan(\arctan(1/\operatorname{ctn} 2\psi)) = 0.9597$$

$$\tan \Phi = -1(\sqrt[3]{0.9597}) = -0.986381$$

$$\operatorname{ctn} 2\Phi = (1 - \tan^2 \Phi) / 2 \tan \Phi = -0.013713$$

$$x = 2(353.5217)(-0.01334) = -9.6959$$

The value $p/3$ is -21.9372 . Reverting to the original unit X in equations (1), (2), and (3) becomes

$$X = x - p/3 = -9.6959 - (-21.9372) = 12.2413$$

For the same data, the workstation reports a result of 12.489, which is in reasonable agreement. The difference between the result obtained by the Star Workstation and the result obtained using the formulas in Excel, can be attributed to rounding errors in acquiring the calibration curve coefficients.

Coefficient of Determination

For a calibration data set composed of points (x_i, y_i) , the fitted curve taking values $Y_i = f(x_i)$, the r^2 value is computed as follows:

$$r^2 = \frac{\left(N \sum_{i=1}^N y_i Y_i - \sum_{i=1}^N y_i \sum_{i=1}^N Y_i \right)^2}{\left(N \sum_{i=1}^N y_i^2 - \left(\sum_{i=1}^N y_i \right)^2 \right) \left(N \sum_{i=1}^N Y_i^2 - \left(\sum_{i=1}^N Y_i \right)^2 \right)}$$

The term “correlation coefficient” has often been misused in chromatography data systems. In statistical terms, r^2 is called the Coefficient of Determination, while its square root r is the Coefficient of Correlation. r^2 can vary from 0 to 1, so r^2 is always smaller or equal to r , and therefore if r^2 meets an arbitrary goal, r will exceed this goal by a greater margin.

The r^2 calculation measures the linearity of the relationship between the y_i and Y_i arrays. Any perfectly linear relationship will lead to a value of 1. The above equation is unaffected by changes in scale or offset of either array. If the coefficients result from a least squares fit, the fitting process has taken care of adjusting scale and offset between the arrays, and the r^2 value is a good indicator of fit. However, if the coefficients are manually edited, a visually imperfect fit may still result in a “good” r^2 value, as long as the two arrays can be scaled and centered to match each other.

When using the 'Force' origin option, the origin data point is added to the curve for the purposes of calculating r^2 . If this is not done, a perfectly linear curve not going through the origin (and therefore obviously not fitting the model) would lead to an r^2 value of 1.

The r^2 value does not take weighted regression schemes into account, but it will likely take different values for quadratic and cubic fits based on what weighting scheme is used.

If you wish to verify the r^2 values determined by the Star Workstation, the PEARSON () function in Excel implements the calculation of r based on two arrays of data.

Relative Standard Deviation of Response Factors

Another parameter reported on the calibration curve is the Relative Standard deviation of the response factors. This parameter is obtained by computing a response factor as $R = \text{Peak Size/Amount}$ for each level in the curve, then computing the unbiased standard deviation and expressing it as a percentage of the average:

$$\text{RespFactorRSD}\% = \frac{\sqrt{\frac{N \sum_{i=1}^{i=N} R_i^2 - \left(\sum_{i=1}^{i=N} R_i \right)^2}{N(N-1)}}}{\frac{\sum_{i=1}^{i=N} R_i}{N}} * 100$$

Peak size is either area or height depending on the calculation selected.

This parameter is used in some EPA methods as a test to determine whether or not it is acceptable to represent the data by a linear, forced through zero calibration model.

Note: Before version 5.0 of the Star Workstation, this formula was applied to all points in the curve. The new algorithm first reduces the data to a response factor for each level, then applies this calculation. This suppresses a possible bias due to different numbers of replicates per level, and eliminates the contribution of experimental noise. This factor is calculated as the method is calibrated, so if a method is upgraded but not recalibrated with 5.0, the value shown will still have been obtained with the old algorithm.

Note: In Version 5.5 this algorithm was modified so that if a single point calibration (with replicates) is used, the RFRSD% is calculated on all the points, as it was prior to 5.0.

Star Workstation Calculation Types

The type of calculations performed with the chromatographic data is selected in the Integration Parameters section of the method. They include:

- Area/Height percent
- Internal standard
- External standard
- Normalized percent

Area/Height Percentage Calculations

Choose the Percentage (%) Calculation Type in the Integration Parameters window of your method when you are making initial analysis runs to build your peak table. As this type of calculation yields the area percentage of each peak relative to the total area of all peaks, no calibration runs are required. The percentage result for each peak is calculated according to the formula

$$(12) \text{ Result}_i = \frac{\text{Area}_i}{\sum_{i=1}^n \text{Area}_i} * 100 * \text{Multiplier} / \text{Divisor}$$

Note: The Multiplier and Divisor are specified on a sample by sample basis in the sample list. Their default value is 1.

Internal Standard Calculations

The following subscripts are used to identify areas and amounts referred to in the results calculations contained in this section.

- CA** - compound present in the analysis run
- CC** - compound present in the calibration run

StdA - Standard compound present in the analysis run

StdC - Standard compound present in the calibration run

Internal Standard Calibrations

The calibration curve for the internal standard is fitted to points defined by:

$$x = \frac{Amt_{(CC)}}{Amt_{(StdC)}} \quad y = \frac{Area_{(CC)}}{Area_{(StdC)}}$$

$Amt_{(CC)}$ is the amount of the analyte, recorded in the Peak Table.

$Amt_{(StdC)}$ is the amount of the standard, recorded in the Peak Table.

$Area_{(CC)}$ and $Area_{(StdC)}$ are taken from the results for the calibration chromatogram.

Internal Standard Analysis

Internal Standard analysis calculation is used to correct for variations in injection volume. It is the most accurate method of calibrated analysis. This analysis requires that a standard of known concentration be added to both the calibration standards and the analysis samples. Amounts of other components are adjusted using the ratio of the area of the peak of the injected amount of Internal Standard to the area of the peak of the calibrated amount of the internal standard.

$$(13) \text{ Result}_{(IS)} = Amt_{(CA)} * \frac{\text{Multiplier}}{\text{Divisor}}$$

where Multiplier and Divisor are scaling constants which you can use to convert from one unit of measure to another. Their default value is 1.0, which leaves your results unchanged. $Amt_{(CA)}$ is derived from the equation for the calibration curve.

$$x = \frac{Amt_{(CA)}}{Amt_{(StdA)}} \quad y = \frac{Area_{(CA)}}{Area_{(StdA)}}$$

Solve the equation for x, and then multiply by the $Amt_{(StdA)}$ to derive $Amt_{(CA)}$.

For example, if the peak of interest shows 31441 area counts, the internal standard peak shows 6086 area counts, the Amount Standard is 1.000, and the calibration curve for is:

$$y = 0.030111 x + 0.98603$$

The equation can then be rewritten as

$$31441/6086 = 0.03011 x + 0.98603$$

When solved, this equation yields a value of 138.827 for x. Multiplication by $Amt_{(StdA)}$, which was 1.000, yields a value for $Amt_{(CA)}$ of 138.827.

External Standard Calculations

External Standard Calibration

The calibration curve is fitted to points defined by

$$y = Amt_{(CC)}, x = Area_{(CC)}$$

The examples given earlier in this chapter for the determination of the calibration curve were using an external standard calculation.

External Standard Analysis

External Standard analysis requires standards containing known amounts for each peak. A calibration curve is developed from the injection of different concentrations of the standards. The response y of the calibration curve equation is $Area_{(CA)}$. The $Amt_{(CA)}$ of each compound is found by solving the calibration curve equation for its designated peak in the peak table. The examples given earlier in this chapter for the solution of the calibration equation during analysis were using an external standard calculation.

Normalized Percent Calculations

Normalized percentage analysis gives the percentage amount of each peak relative to the total amount of all peaks. This type of calculation requires data from a previous calibration run using an internal standard to determine the relative response for each identified peak. Calibration curve coefficients, derived from the calibration run, are stored in the method peak table. During a Normalized Percentage Analysis run, they are used to adjust results by comparing the response for the analyte to the response for a known amount, and adjusting the amount for each peak accordingly.

Normalized Percent Calibration

The calibration curve for Normalized Percent is fitted to the points defined by

$$x = \frac{Amt_{(CC)}}{Amt_{(StdC)}} \quad y = \frac{Area_{(CC)}}{Area_{(StdC)}}$$

Normalized Percent Analysis

Result_{N%} is the percentage amount of the peak relative to the total amount of all peaks. It is calculated from the following equation:

$$(14) \text{ Result}_{N\%} = \frac{Amt_{(CA)}}{Amt_{(Total)}} * 100$$

Calculation of Amt_{CA} for Identified Compounds

The amount of the compound, Amt_{CA}, for peaks listed in the method peak table, is derived by solving the compound's calibration equation for x.

In Normalized Percent analysis, the values of x and y in the calibration equation are:

$$x = Amt_{(CA)} \quad y = \frac{Area_{(CA)}}{Area_{(StdA)}}$$

Calculation of Amt_{CA} for Unidentified Compounds

The Amt_{CA} for unidentified peaks is derived from the equation

$$(15) \text{ Amt}_{(CA)} = \frac{\text{Area}_{(CA)}}{\text{Area}_{(StdA)} * \text{UnID'dPkFactor}}$$

UnID'dPkFactor, the Unidentified Peak Factor, is placed in the Reintegration List in Interactive Graphics, or in the Sample Log if you are doing your recalculations in System Control.

Calculation of Amt_{Total}

Amt_{Total} is the total of the amounts for each peak in the peak table, as derived from their calibration equation plus the total amount of any unidentified peaks.

(16)

$$\text{Amt}_{(Total)} = \sum_{i=1}^n \frac{\text{Area}_{(CA_i)}}{\text{Area}_{(StdA)} * \text{UnID'dPkFactor}} + \sum_{i=1}^n \text{Amt}_{(CA_i)}$$

In the following example, only the first three peaks are defined in the peak table. The 'Report Unidentified Peaks' option was checked in the Integration Parameters section of the method.

Peak No	Peak Name	Result	Ret Time	Time Offset	Area (Counts)	Sep Code	Width 1/2
1	X	1.183	4.624	0.003	5716	BB	1.5
2	Y	49.553	5.671	0.003	12844	BB	1.5
3	Z	48.342	6.975	0.006	3486	BB	1.9
4		0.159	7.506	0.000	7703	BB	1.5
5		0.170	8.472	0.000	8194	BB	1.6
6		0.233	8.995	0.000	11252	BB	1.8
7		0.179	9.322	0.000	8660	BB	1.8
8		0.180	10.853	0.000	8703	BB	1.8
Totals:		99.999	0.000		66558		
Total Unidentified Counts : 44512 counts							
Detected Peaks: 10		Rejected Peaks: 2		Identified Peaks: 3			
Amount Standard: 1.000000		Multiplier: 1.000000		Divisor: 1.000000			

Calculation of Amt_{CA} for Identified Peaks

Three peaks are listed in the Peak Table. Amt_{CA} is derived from the equations for their calibration curves. Peak 1 is the Internal Standard.

Peak	Equation	Amt _(CA)
1	5716/5716 = x Internal Std.	1.00
2	12844/5716=0.030111x + 0.986027	41.88
3	3486/5716 = 0.011249x + 0.150232	40.86
	Total	83.74

Calculation of Amt_{CA} for Unidentified Peaks

The Amt_{CA} for each unidentified peak is calculated using equation (15). This equation requires the value of the Unidentified Peak Factor. For ease of calculation, the value 10 was selected.

Peak	Equation	Amt _(CA)
4	7703/((5716)(10))	0.134762
5	8194/((5716)(10))	0.143352
6	11252/(5716)(10)	0.196851
7	8660/((5716)(10))	0.151505
8	8703/((5716)(10))	0.152257
	Total	0.778727

Calculation of Amt_{Total}

Use equation (16) to calculate Amt_{Total}:

$$\text{Amt}_{\text{Total}} = 0.778727 + 83.74 = 84.518727$$

Calculation of Result_{N%}

Now that you have Amt_{Total} and Amt_{CA}, you can compute Result_{N%} for each peak using equation (14):

Peak	Equation	Result _{N%}
1	$(1/84.518727)(100)$	1.183
2	$(41.88/84.518727)(100)$	49.553
3	$(40.86/84.518727)(100)$	48.342
4	$(0.1347621/84.518727)(100)$	0.159
5	$(0.1433520/84.518727)(100)$	0.170
6	$(0.1968509/84.518727)(100)$	0.233
7	$(0.1515045/84.518727)(100)$	0.179
8	$(0.1522568/84.518727)(100)$	0.180
	Total	99.999