Mars Science Laboratory
Software Interface Specification
Chemistry Mineralogy (CheMin) Reduced Data Record (RDR)
Initial Release

Version 0.4
JPL D-38121
SIS-SCI018-MSL

February 15, 2013

Prepared by:

_____
David Vaniman, PSI

.

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

SIGNATURES

Prepared by:

Custodian:

David Vaniman, PSI
MSL Project — Date

Concurrence:

MSL Science Manager:

Joy Crisp — Date

Instrument Scientist:

Albert Yen — Date

Principal Investigator:

David Blake — Date

MSL OPGS Cognizant Engineer:

Maher Hanna — Date

PDS Data Engineer:

Betty Sword — Date

Geosciences Node Manager:

Raymond Arvidson — Date

PDS Program Manager:

Thomas H. Morgan — Date

CHANGE LOG

DATE	SECTIONS CHANGED	REASON FOR CHANGE	REVISION
08/31/11	All	First draft	Draft v0.1
05/31/12	All	Second draft, following review by S. Slavney and J. Crisp	Draft v0.2
06/8/12 accepted 06/12/12	Page header, acronyms, 2.2, 2.3, 2.4.2, 2.4.4, 3.2, 3.3, 5, all appendices	Revision by S. Slavney	Draft v0.3
08/11/12	All	Edited in response to external review comments	Draft v0.4
02/15/13	Fig. 2 update, minor edits	Internal review prior to first CheMin data release	Initial Release v0.4

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS AND ABBREVIATIONS

ASCII	American Standard Code for Information Interchange
APSS	Activity Planning and Sequencing Subsystem
CHEMIN	Chemistry and Mineralogy Instrument
CODMAC	Committee on Data Management and Computation
CSV	Comma Separated Value
CWA	Current Working Area
DN	Digital Number
EDH	Energy Dispersive Histogram
EDR	Experiment Data Record
FEI	File Exchange Interface
ICD	Interface Control Document
ISO	International Standards Organization
JPL	Jet Propulsion Laboratory
Kbyte	Kilobytes
LSB	Least Significant Byte
MB	Mega Bytes
MIPL	Multimission Image Processing Laboratory
MPCS	Mission data Processing and Control Subsystem
MSB	Most Significant Byte
MSL	Mars Science Laboratory
NASA	National Aeronautics and Space Administration
ODL	Object Description Language
ODS	Operations Data Store
OPGS	Operations Product Generation Subsystem
PDL	Payload Downlink Lead
PDS	Planetary Data System
PEL	Payload Element Lead
PPPCS	Pointing, Positioning, Phasing & Coordinate Systems
PUL	Payload Uplink Lead
RA	Robotic Arm
RAM	Random Access Memory
RAMP	Rover Avionics Mounting Platform
RCE	Rover Compute Element
RDR	Reduced Data Record
RMC	Rover Motion Counter

RSVP	Rover Sequence and Visualization Program
SA-SPaH	Sample Acquistion – Sample Processing and Handling
SCM	Spacecraft Configuration Manager
SFDU	Standard Formatted Data Unit
SIS	Software Interface Specification
SOAS	Science Operations Analysis Subsystem
SOWG	Science Operations Working Group
TBD	To Be Determined
TC	Temperature Compensation
TDS	Telemetry Delivery Subsystem
URL	Universal Resource Locator
WEB	Warm Electronics Box
XRD	X-ray Diffraction
XRF	X-ray Fluorescence

GLOSSARY

TERM	DEFINITION
Meta-Data	Selected or summary information about data. PDS catalog objects and data product labels are forms of meta-data for summarizing important aspects of data sets and data products.

1. INTRODUCTION

1.1 Purpose and Scope

The purpose of this data product Software Interface Specification (SIS) is to provide users of the Chemistry and Mineralogy (CheMin) Reduced Data Records (RDRs) with a detailed description of the products and how they were generated. The CheMin instrument determines the mineralogy and elemental composition of powdered samples through the combined application of X-ray diffraction (XRD, producing mineral identification and quantification) and X-ray fluorescence (chemical analysis based on Energy Dispersive Histograms, EDH).

This SIS is intended to provide enough information to enable users to understand the CheMin RDR data products. The users for whom this SIS is intended are software users or developers of programs used in generating the RDR products and scientists who will analyze the data, including those associated with the Mars Science Laboratory (MSL) Project and those in the general planetary science community.

1.2 Contents

This data product SIS describes how the MSL CheMin instrument acquires its data and how the data are processed, formatted, labeled, and uniquely identified. The document discusses standards used in generating the product and software that may be used to access the product. The data product structure and organization is described in sufficient detail to enable a user to read the product. Examples of four product labels are provided (XRD data filtered by primary X-ray energy, XRD "film" data, energy-dispersive histograms, EDH, and a mineral abundance table). Differences between these data types are described below in sections 2.3 and 2.4.

1.3 Applicable Documents and Constraints

This data product SIS is responsive to the following MSL documents:

1. Mars Exploration Program Data Management Plan, R. E. Arvidson, S. Slavney and J. Ward, Rev. 4, June 15, 2011.
2. Mars Exploration Rover Project Archive Generation, Validation and Transfer Plan, R. E. Arvidson and S. Slavney, JPL D-19658, March 22, 2002.
3. CheMin Functional Design Document, MSL 375-1232, JPL D-34222, July 27, 2010.
4. Pointing, Positioning, Phasing & Coordinate Systems (3PCS), Volume 1, Santi Udomkesmalee, MSL-376-1297, JPL D-34642, May 29, 2007.
5. MSL CheMin EDR Software Interface Specification, H. Mortensen, Version 1.0, JPL D-69260, MSL-576-3503, September 16, 2011.

This SIS is also consistent with the following Planetary Data System documents:

6. Planetary Data System Archive Preparation Guide, Version 1.4, JPL D-31224, April 1, 2010.
7. Planetary Data System Standards Reference, February 27, 2009, Version 3.8, JPL D-7669, Part 2.
8. MSL CheMin Science Team and PDS Geosciences Node Interface Control Document (ICD), Version 2.0, May 14, 2007.

9. MSL Project Experiment Data Record Archive Volume Software Interface Specification (SIS), Version 1.0, JPL D-64995, May 5, 2011.

10. Description of the CheMin instrument:

Blake, D. et al., Characterization and Calibration of the CheMin Mineralogical Instrument on Mars Science Laboratory, Space Science Reviews, v. 170, p. 341-399. DOI 10.1007/s11214-012-9905-1, 2012.

Finally, this SIS is meant to be consistent with the contract negotiated between the MSL Project and the CheMin Principal Investigator (PI) in which experiment data records and documentation are explicitly defined as deliverable products.

1.4 Relationships with Other Interfaces

Changes to this CheMin RDR SIS document affect the products, software, and/or documents listed in Table 1.

Table 1: Product and Software Interfaces to this SIS

Name	Type P=product S=software D=document	Owner
CheMin EDRs	P	OPGS/MIPL
MSLEdrGen	S	MIPL
MIPL database schema	P	MIPL
CheMin RDRs	P	CheMin Science Team
CheMin RDR SIS	D	CheMin Science Team
Other CheMin Programs/Products/Documents	P/S/D	CheMin Science Team

2. DATA PRODUCT CHARACTERISTICS AND ENVIRONMENT

2.1 Instrument Overview

A detailed description of the MSL Chemistry & Mineralogy (CheMin) instrument can be found in Blake et. al, 2012). The CheMin instrument performs mineralogical and elemental analyses of samples. Mounted inside the rover (Figure 1), CheMin consists of a sample funnel, analysis cells, X-ray source, detector assembly, and electronics. It has 27 reusable cells for analyses of powered rock and regolith. In addition, it has 5 sample cells containing standard materials (Figure 2). CheMin exposes the sample to a collimated beam of unfiltered cobalt X-rays and the fluoresced and diffracted X-rays are detected by a CCD. The X-ray diffraction patterns are used to identify and quantify minerals in samples and the fluoresced X-ray energies are used to establish the chemical composition. The nominal duration of a single experiment, sufficient to quantitatively assess simple mixtures of crystalline phases, is less than 4 hours. Complex assemblages such as basalt with 8 or more phases may require up to 10 hours of data. This data need not be taken contiguously.

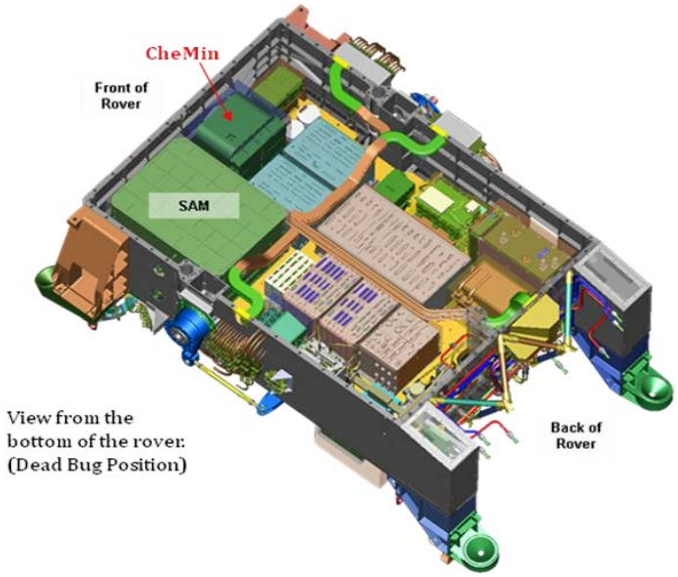

Figure 1: Location of CheMin within the rover body of Curiosity.

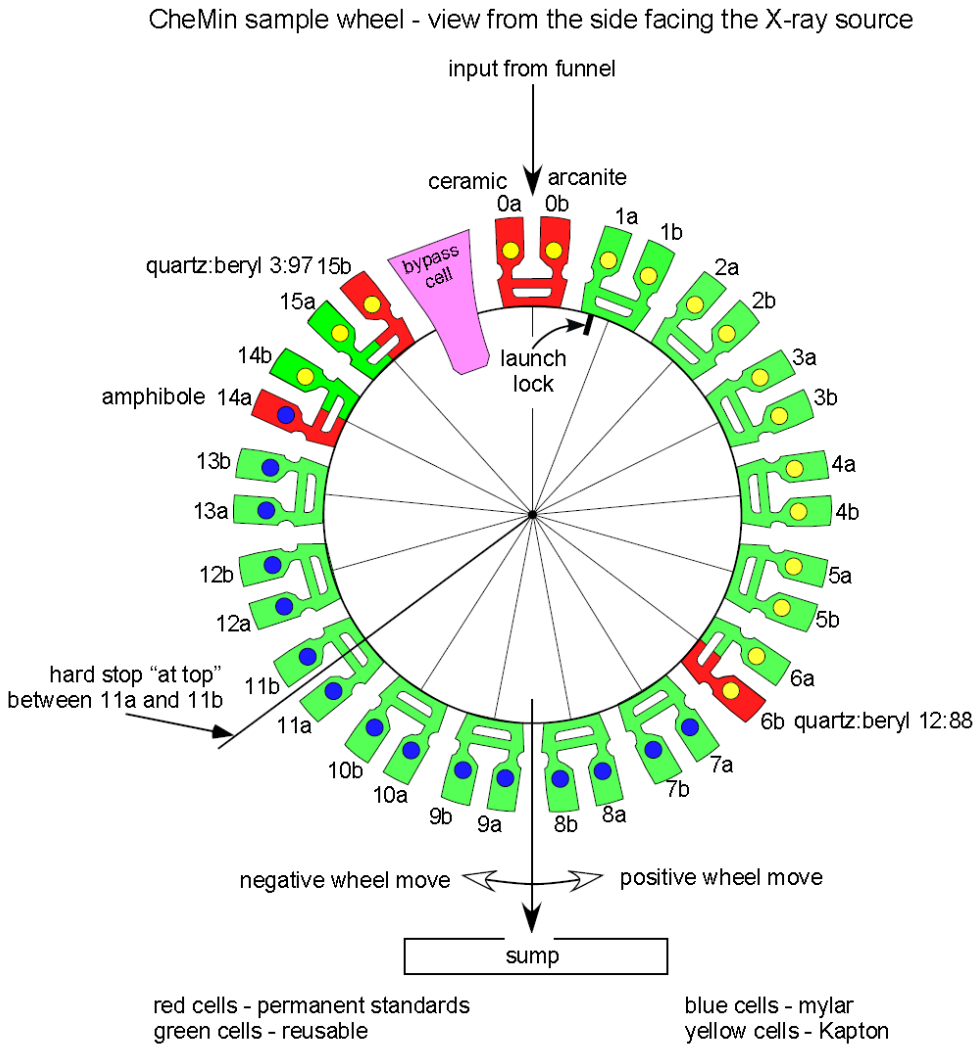

Figure 2: Schematic image of the CheMin sample wheel.

2.1.1 Standards Carried by CheMin

The five standard materials carried by CheMin are amphibole, arcanite, a synthetic ceramic material spiked with many elements, and two mixtures of beryl and quartz in different ratios. Standards were chosen to provide coverage of both XRD and EDH calibration. The amphibole standard provides diffraction over a wide two-theta range and has a complex composition suitable for EDH calibration. The arcanite standard provides XRD calibration at mid to high two-theta and is the primary standard for sulfur calibration (along with high potassium content). The ceramic material is complex and partly amorphous in XRD; it is used for calibration of many elements, particularly those not provided by the amphibole or arcanite standards. The beryl in the beryl and quartz mixtures provides a wide range of diffraction for XRD calibration, particularly to low two-theta. The beryl and quartz mixtures are also used to evaluate XRD detection limits (beryl:quartz in 97:3 ratio) and accuracy (beryl:quartz in 88:12 ratio).

Table 2 summarizes the chemical compositions of the five CheMin standards and illustrates the relative utility for EDH calibration.

Table 2: Compositions of Standards Carried by CheMin

	1	2	3	4	5
Standard:	Amphibole	Beryl with 3% Quartz	Beryl with 12% Quartz	Arcanite	Ceramic
*H_2O	*1.41*				
*Li					*0.84*
*B					*2.63*
*BeO		*13.82*	*12.54*		
*Na_2O	*2.45*				*0.06*
*MgO	*13.21*				*0.12*
Al_2O_3	15.32	16.30	14.78		1.39
SiO_2	42.84	69.40	72.24		25.10
P					2.11
SO_3				45.95	10.31
Cl					5.83
K_2O	0.70			54.05	1.57
CaO	9.80				7.98
TiO_2	0.98				7.92
Cr		0.48	0.44		2.05
MnO	0.08				2.80
Fe_{tot} as FeO	12.01				15.23
Ni					2.06
Zn					2.19
Br					2.65
Rb					2.80
Sr					2.69

*Components in red italics are not effectively detected by the CheMin CCD.

2.1.2 CheMin Sample Handling

CheMin sample cells have thin windows of low-Z material that allow X-ray transmission. Kapton windows are tough and resist acid attack, but have a small diffraction peak at low 2-theta that may interfere with detection of small amounts of smectite. Mylar windows have no diffraction interference at low 2-theta (albeit a very small contribution at ~19.5 degrees) but are not as rugged as Kapton. The choice of which window to use will depend on expected sample properties. Figure 3 illustrates the differences in diffraction background between Mylar and Kapton windows. The small amount of background diffraction is known and of minor concern unless a sample may have a very small amount of smectite of interest, in which case a Mylar cell is preferable in order to avoid the small background diffraction rise in Kapton that overlaps with the major smectite 001 diffraction at low two-theta.

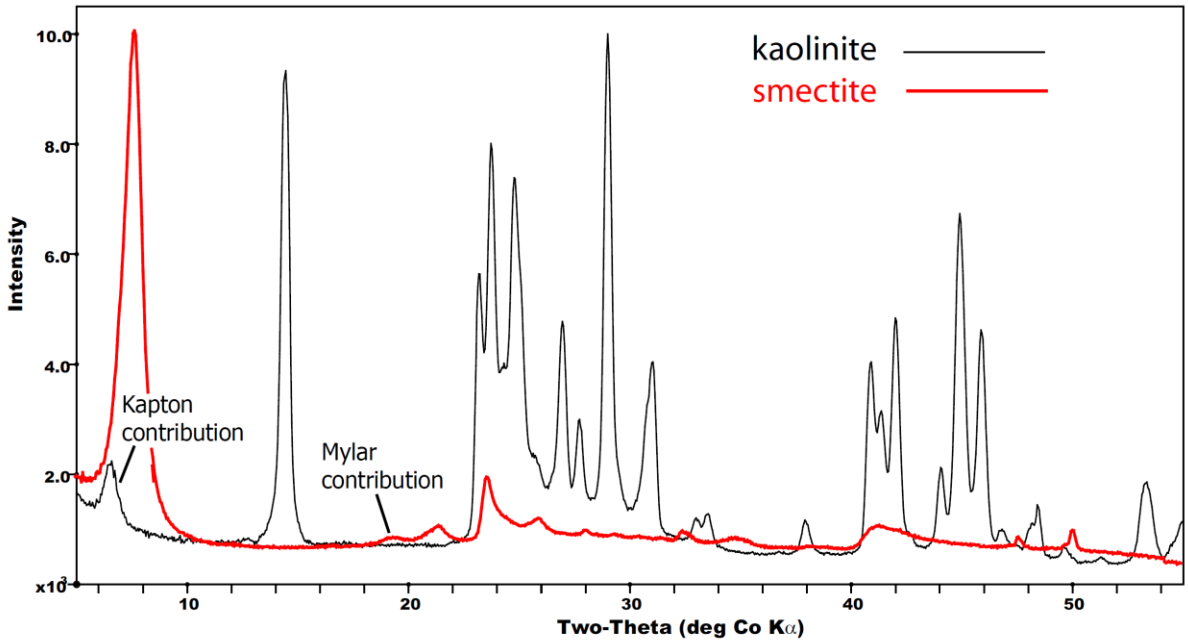

Figure 3: Diffraction background contributions in the two different CheMin sample cell window materials (Kapton and Mylar). These effects are illustrated using kaolinite in a Kapton cell and smectite in a Mylar cell (data from prototype CheMin IV instrument).

2.1.3 CheMin Data Products

In collecting data, the CheMin CCD detector is exposed, read and erased many times. CCD exposure times can be varied from 5-30 seconds and many hundreds of frames can be collected. Ideally, the exposure time will be short enough that each CCD pixel receives only one photon of either diffracted or fluoresced energy (a single-pixel event). The CCD acquires a 582 x 600 pixel image of diffracted (Co Kα and Co Kβ) and fluoresced X-rays (see the frame images in Figure 4). The energy difference between Co Kα and Kβ is 724 eV, easily separated with the 250 eV resolution of the CCD. Incident X-rays can be binned by energy to distinguish primary Co X-rays from those fluoresced from elements in the sample. Only diffracted X-rays produce organized Debye rings as shown in Figure 4; fluoresced X-rays are distributed at random. In

order to distinguish primary Co Kα X-rays from fluoresced Fe Kα X-rays, the CCD should have an energy resolution of 250 eV or better, obtained by cooling. When 2-D Debye diffraction rings are converted to 1-D patterns of energy versus intensity, the effective 2-theta range of the pattern is ~5 to 55 degrees (based on the Bragg relationship for Co radiation).

The ideal case in CheMin operation is to collect CCD frames in which each of the CCD pixels contains energy from a single photon, whether diffracted or fluoresced. In practice, photon events may 'split' between pixels but these can be recovered in CheMin software. If adjacent, edge-sharing pixels (the four pixels immediately above, below, right and left of the x,y pixel of interest) have a summed energy that falls within a ground-specified energy window, the 2-D array is incremented by one at that x,y location. Split pixels can be recovered for primary diffracted cobalt Kα and Kβ energies as well as for fluoresced energies. Note however that fluoresced photons are rare enough compared to diffracted that a good diffraction pattern can also be obtained by readout of all events, without single-pixel or split-pixel energy discrimination. This type of diffraction pattern is referred to as "film mode," somewhat like leaving a camera shutter open and recording visible photons of all energies (colors) as a grayscale film image.

A complete analysis can comprise up to 10 hours of data collected over one or more sols, and is made up of 10-20 "minor frames" that are each stand-alone analyses. Minor frames are recombined by the CheMin science team to obtain a major frame analysis (Figure 4).

Each minor frame includes: (1) X-ray energy histogram (all events, which will include single pixel hits, split pixel hits whether confined to edge-sharing pixels or not, and multiple photon hits within pixels), (2) X-ray energy histogram (single pixel events), (3) X-ray energy histogram (single pixel events plus split pixel events), (4) 2-D diffraction pattern (all events), (5) 2-D diffraction pattern (single pixel events) and (6) 2-D diffraction pattern (single pixel events plus split pixel events).

"Quick look" data products include (1) X-ray Energy-Dispersive histogram, and (2) 2-D diffraction pattern (Figure 5). The "quick look" products provide sufficient information for the PDL to provide the SOWG with a chemical and mineralogical summary.

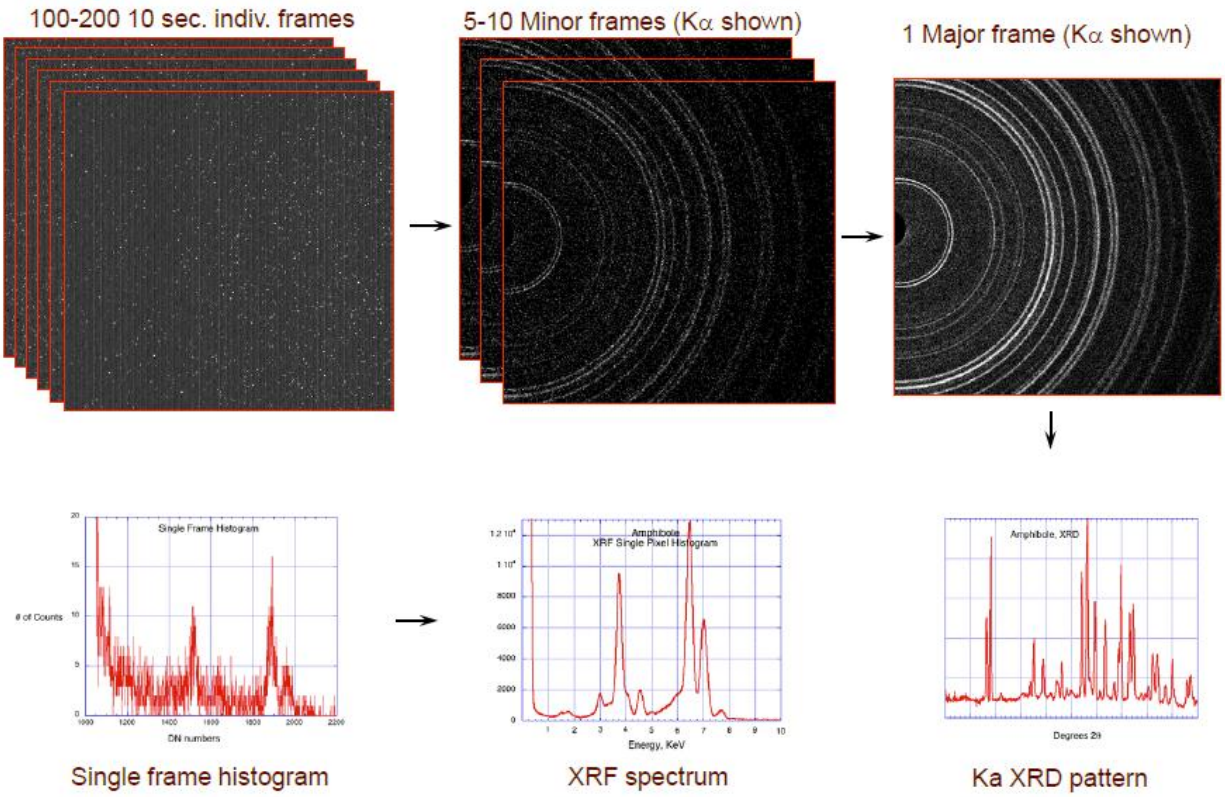

Figure 4: Progression from individual CCD frames to minor frames, a major frame (single-pixel mode), and data products.

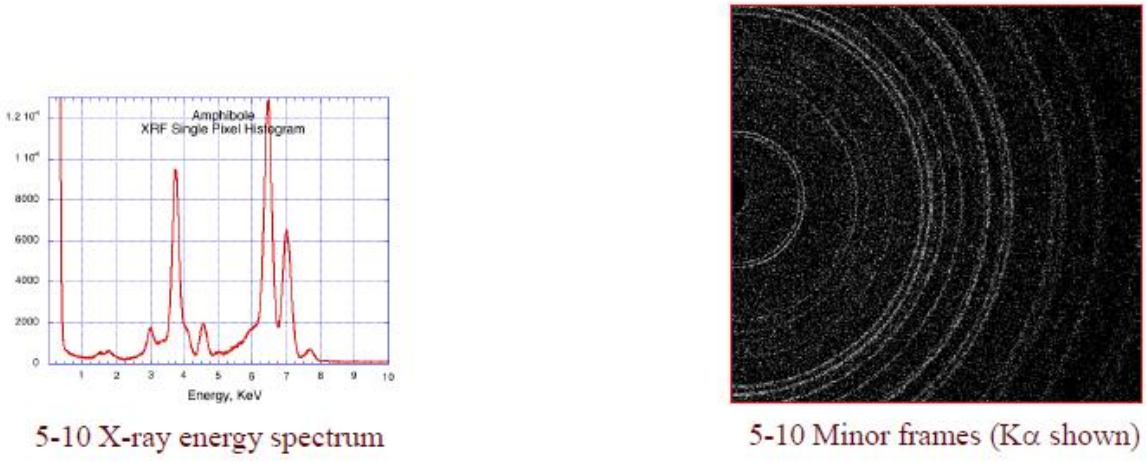

Figure 5: CheMin "Quick-look" PDL products (the minor frames may also be processed by the PDL into a preliminary 1D diffraction pattern for mineral identification).

2.2 Data Product Overview

Each CheMin RDR will consist of two files. The first file is an ASCII formatted detached PDS label. The second file is an ASCII file typically containing housekeeping data and either a range of intensities in histogram format (intensity versus two-theta or intensity versus energy) or a table with rows of mineral identifications versus columns with mineral abundances in weight percent and estimated errors. Tabular data are stored in comma-separated-value (CSV) format: fields in a row are separated by commas, and the lengths of rows may vary within a single file.

2.3 RDR Product Descriptions

The RDR products that may be provided by CheMin are listed in **Error! Reference source not found.**, aligned with the EDR products used to prepare the RDRs. The three-character product type appears in the product's file name as described in section 2.4.4. There are two primary level 1B RDR product types, diffraction and energy histograms. In practice, although multiple EDRs are collected, the best analytical results are generally obtained using "diffraction all" (EDA) and "energy single" (EE1), resulting in equivalent RDR files labeled RDA for diffraction and RE1 for energy (Table 3). These and the other RDR products less commonly generated are described in more detail in section 2.4.2. The level 2 RDRs (tables of mineral abundance) can be developed from any diffraction RDR.

Examples of RDR data file labels, format files, and product tables are provided in Appendix A (Level 1B Co Kα XRD data), Appendix B (Level 1B film XRD data from stacked raw frames), and Appendix C (Level 1B EDH data); these Level 1B data products will be delivered to PDS with data set identification as MSL-M-CHEMIN-4-RDR-V1.0. Tables with Level 2 mineral identifications and abundances will be delivered separately, with data set identification as MSL-M-CHEMIN-5-RDR-V1.0; an example is in Appendix D (mineral identification and abundance data). The separate delivery of mineral identification and abundance data allows analysis and revision of these data independently of the processed level 1 source data.

Table 3: Relationships between CheMin EDRs and RDRs

EDR Product Type Description	EDR Value	Daughter Level 1B RDR Type	Level 1B RDR Value	Level 2 RDR
CCD Frame	"**ECC**"	n/a	n/a	n/a
Housekeeping N	**"EHK"**	n/a	n/a	n/a
Diffraction Single	"**ED1**"	Diffraction Single, windowed K-alpha (rarely K-beta)	"**RD1**"	"**MIN**"
Diffraction Split	"**EDS**"	Diffraction Split, windowed K-alpha (rarely K-beta)	"**RDS**"	"**MIN**"
Diffraction All	"**EDA**"	Diffraction All, windowed K-alpha (rarely K-beta)	"**RDA**"	"**MIN**"
Transmit Raw	"**ETR**"	Diffraction All, Raw	"**RTR**"	"**MIN**"
Film	"**EFM**"	Diffraction, Film Mode	"**RDF**"	"**MIN**"
Energy All	"**EEA**"	Energy All	"**REA**"	n/a
Energy Single	"**EE1**"	Energy Single	"**RE1**"	n/a
Energy Split	"**EES**"	Energy Split	"**RES**"	n/a

2.4 Data Processing

2.4.1 Data Processing Level

This SIS uses the Committee on Data Management and Computation (CODMAC) data level numbering system to describe the processing level of the RDR data product. CheMin RDR data products are considered CODMAC "Level 4" or "Resampled Data" (equivalent to NASA level 1-B) products. The RDR data files are generated from CODMAC "Level 2" or "Edited Data", which are based on the EDR products as provided by OPGS. Tables of mineral abundances are NASA Level 2 products, derived from Level 1B RDR files of XRD data. At present there are no equivalent NASA Level 2 products for quantitative XRF based on EDH data; derivation of quantitative XRF data from CheMin is a research effort on the part of the CheMin team and any future creation of NASA Level 2 XRF data will be described in a revision to this document. Refer to Table 3 for a breakdown of the CODMAC and NASA data processing levels.

Table 3: Processing Levels for Science Data Sets

NASA	CODMAC	Description
Packet data	Raw – Level 1	Telemetry data stream as received at the ground station, with science and engineering data embedded.
Level-0	Edited – Level 2	Instrument science data (e.g., raw voltages, counts) at full resolution, time ordered, with duplicates and transmission errors removed.
Level 1-A	Calibrated - Level 3	Level 0 data that have been located in space and may have been transformed (e.g., calibrated, rearranged) in a reversible manner and packaged with needed ancillary and auxiliary data (e.g., radiances with the calibration equations applied).
Level 1-B	Resampled - Level 4	Irreversibly transformed (e.g., resampled, remapped, calibrated) values of the instrument measurements (e.g., radiances, magnetic field strength).
Level 2	Derived - Level 5	Geophysical parameters, generally derived from Level 1 data, and located in space and time commensurate with instrument location, pointing, and sampling.
Level 3	Derived - Level 5	Geophysical parameters mapped onto uniform space-time grids.

2.4.2 Data Product Generation

CheMin RDR data products will be generated by the CheMin science team, either at JPL for the first 90 days of surface operations or at JPL and remote institutions after the initial 90 days of surface operations. The RDR data products will be prepared from calibrated and transformed EDR data products (see below) and formatted according to this RDR SIS. Meta-data acquired from the telemetry data headers will be used to populate the PDS label. There may be multiple versions of a CheMin RDR product. If telemetry packets are missing during the initial downlink from the rover memory and initial EDR product preparation, partial data sets will be created and the missing data will be filled with zeroes. The data will be reprocessed after all data are received and the original version will be overwritten.

Calibration and transformation of the RDR data will vary by source EDR file type. The calibration and transformation process is described below for the different XRD and EDH file types. Primary file types are those that will be generally produced. The primary EDR for generating diffraction RDRs is "diffraction all" (EDA), which is processed into the equivalent RDR product labeled RDA. Special diffraction products are those that will typically not be prepared unless there are problems or uncertainties in the analysis of the Co Kα diffraction pattern. For instance, analysis of the Co Kβ as well as the Co Kα pattern may help resolve phase identification issues in complex samples with many peaks. Raw frames may become more critical for XRD analysis as the mission progresses and CCD energy resolution is degraded. For energy dispersive histograms (EDH), the primary product will be the single-pixel energy histogram (EE1), which is processed into the equivalent RDR product labeled RE1. The energy split-pixel and total-energy histograms are ancillary. The EDH conversion from digital number (DN) of an EDR to energy (keV) of an RDR relies on a conversion factor determined using the known energy of the Co Kα peak (6.925 keV).

The reference paper by Blake et al., (2012), "Characterization and Calibration of the CheMin Mineralogical Instrument on Mars Science Laboratory," provides descriptions of the several EDR diffraction and energy products with discussion of when products other than diffraction all (EDA) or single-pixel energy (EE1) might be selected for producing archived level 1B data products.

The relations between EDR files and RDR products are summarized below.

From EDR files of single-pixel XRD data (from ED1 to product RD1): This diffraction pattern is derived from an EDR "diffraction single" file (ED1, Table 3). The ED1 file is made of multiple frames (e.g., 120) that are combined into a single 2D array with the same dimensions as the original images. Specific low and high energy thresholds are defined, generally to select Kα energy photons (in special cases Kβ may be selected). Only those pixels are counted that are within the energy window and surrounded on all edges by pixels with energy below a defined threshold (removing background). To process an ED1 file into an RD1 product, Debye diffraction rings in these 2D images are summed along 2-theta radii and the rings are adjusted for arc length to produce tables of intensity versus 2-theta and standard 1D diffraction patterns of intensity versus 2-theta. Occasionally hot pixels are observed in the EDR image; these are artifacts and they are replaced manually by local background before the 2D EDR is processed into a 1D RDR. This can be done using any image processing software (e.g., ImageJ). The Kα single-pixel XRD pattern is a special diffraction product that will typically not be prepared.

From EDR files of split-pixel XRD data (from EDS to product RDS): This diffraction pattern is derived from an EDR "diffraction split" file (EDS, Table 3). The "diffraction split" EDS file includes all pixels binned within a specified energy widow (similar to the ED1 file) but includes those pixels where the energy sum of two edge-sharing pixels is also within the energy window. This allows for the relatively common case in which a photon's energy is split between edge-sharing pixels. Generally the Kα photons are selected (in special cases Kβ may be selected). As with the ED1 file, multiple frames (e.g., 120) are combined into a single array with the same dimensions as the original images. The RDS product is prepared from the EDS file by summing the Debye diffraction rings in the 2D images along 2-theta radii and adjusting them for arc length to produce tables of intensity versus 2-theta and standard 1D diffraction patterns of intensity versus 2-theta. Occasionally hot pixels are observed in the EDR image; these are artifacts and they are replaced manually by local background before the 2D EDR is processed into a 1D RDR. This can be done using any image processing software (e.g., ImageJ). The split-pixel XRD pattern is a special diffraction product that will typically not be prepared.

From EDR files of XRD data for all pixels within an energy window (from EDA to product RDA): This diffraction pattern is derived from an EDR "diffraction all" file (EDA, Table 3). The "diffraction all" EDA file uses a selected energy window (generally Kα, although Kβ may be selected in special cases). The EDA file is similar to an ED1 file but places no restrictions on the energy state of adjacent pixels. The EDA file often has the highest accumulated signal because diffraction events tend to cluster together along Debye diffraction arcs, producing more instances of adjacent pixels with Co X-ray hits that would be rejected in an ED1 file. However, multiple or split fluorescence events may also fall within the energy window and be counted as primary Co X-rays, increasing the background. The RDA product is prepared from the EDA file by summing the Debye diffraction rings in these 2D images along 2-theta radii and adjusting them for arc length to produce tables of intensity versus 2-theta and standard 1D diffraction patterns of intensity versus 2-theta. Occasionally hot pixels are observed in the EDR image; these are artifacts and they are replaced manually by local background before the 2D EDR is processed into a 1D RDR. This can be done using any image processing software (e.g., ImageJ). The diffraction-all XRD pattern is generally the strongest diffraction product with best signal strength and good separation of diffraction peaks. ***In most cases, RDA will be the primary diffraction RDR.***

From EDR files of raw XRD frames (from ETR to product RTR): This diffraction pattern is derived from an EDR "transmit raw" file (ETR, Table 3). Raw frames contain all pixel energy information without energy filtering. Ideally, all raw frames would be transmitted to Earth for processing because this is the most complete record of all diffraction and fluorescence events within a CheMin analysis. However, the file size requirements for sending down all raw frames are excessive for MSL operations. In some cases a minor frame may be small enough in file size to be returned in this mode. Another alternative is to process the analysis onboard to return a "modified raw" major frame that includes all individual frames but discards all pixels with energy below a selected energy value and run-length-encodes the data, allowing frame reconstruction on receipt. This process works well with strong diffraction lines but may lose minor lines of major phases and all lines from minor phases, so this is a special XRD product for specific instances where it may be desirable to retrieve all frames. Occasionally hot pixels are observed in the EDR image; these are artifacts and they are replaced manually by local background before the 2D EDR is processed into a 1D RDR. This can be done using any image processing software (e.g., ImageJ). The raw-frame XRD pattern is a special diffraction product that will typically not be prepared.

From EDR files of XRD "film" data (from EFM to product RDF): This diffraction pattern is derived from an EDR "film" file (EFM, Table 3). An EFM file consists of summed raw frames with no energy discrimination and no background removal. In such a product the CCD is treated much like a piece of black-and-white film in a camera with no capability to resolve or filter color. All fluorescence events are recorded as well as diffraction. Diffraction events will generally far exceed fluorescence, so a Debye pattern can usually be resolved, but the peak-to-background ratio is lower than in other diffraction products, Similar to other diffraction processing, the RDF file is prepared from an EFM file by summing Debye diffraction rings in the 2D images along 2-theta radii and adjusting for arc length to produce tables of intensity versus 2-theta and standard 1D diffraction patterns of intensity versus 2-theta. Flagging of hot pixels is applied. This is a special XRD product from a CheMin analysis, generally used only if issues are identified in onboard processing.

From a single-pixel Energy Dispersive Histogram (EDH) of counts versus energy (from EE1 to product RE1): This energy histogram product is prepared from an "energy single" EDR

file (EE1, Table 3). The EE1 file is a histogram that includes only those X-ray photon energies where the pixel that absorbed the photon is surrounded on all edges by pixels with background values. The summed DN values from multiple frames are used. To prepare a reduced RE1 product from as EE1 file, DN values are converted to energy in keV using a conversion factor specified in the file metadata. The processed data are presented in a single table that can be plotted as a histogram representing counts versus energy. ***This is the primary EDH product from a CheMin analysis.***

From a split-pixel Energy Dispersive Histogram (EDH) of counts versus energy (from EES to product RES): This energy histogram product is prepared from a "split pixel" EDR file (EES, Table 3). The EES file histogram is made up of single pixel photons (as in an EE1 file) plus the summed energies of split pixels. The summed single-pixel and reconstructed split-pixel DN values from multiple frames are used. To prepare a reduced RES product from as EES file, DN values are converted to energy in keV using a conversion factor specified in the file metadata. The processed data are presented in a single table that can be plotted as a histogram representing counts versus energy. This is an ancillary processed EDH product from a CheMin analysis, generally not produced.

From the total Energy Dispersive Histogram (EDH) of counts versus energy (from EEA to product REA): This energy histogram product is prepared from an "energy all" EDR file (EEA, Table 3). The EEA file sums all X-ray photon energies deposited in all pixels (single pixels, split pixels, and multiple photon hits). The summed DN values from multiple frames are used. To prepare a reduced REA product from an EEA file, DN values are converted to energy in keV using a conversion factor specified in the file metadata. The processed data are presented in a single table that can be plotted as a histogram representing counts versus energy. This is an ancillary processed EDH product from a CheMin analysis and is generally not produced.

In addition to these Level 1B products directly tied to EDR files, the level 2 product of mineral identifications and abundances can be derived from any CheMin RDR with a "D" or "T" in the product name (RD1, RDS, RDA, RTR, and RDF).

Level 2 Mineral identifications and abundances (product MIN): A list of minerals present in a CheMin analysis is compiled from the NASA Level 1 processed XRD data using one or more mineral identification programs (see section 4.1); the identification and quantification programs(s) used are specified in the file metadata. The databases used for mineral identification are standard; some such as the American Mineralogist (Mineralogical Society of America) Crystal Structure Database (AMCSD) are free whereas the International Centre for Diffraction Data (ICDD) must be purchased. Identifications are scrutinized for reliability against detection limits and potential complications (e.g., peak overlaps). Mineral identifications are further evaluated against information from EDH element identifications and other instrument datasets (particularly ChemCam, APXS, and SAM) to evaluate whether minor phases near detection limits or phases problematic in XRD analysis (e.g., poorly crystalline or amorphous) need to be considered. Rietveld or other pattern-fitting methods will be used to generate a table of mineral abundances in weight percent along with estimated errors.

2.4.3 Data Flow

The CheMin RDR data products generated by the CheMin science team during operations are created using NASA Level 0 data products from the OPGS (MIPL) pipeline so that CheMin data can be reduced and reported to the SOWG in a timely manner to support both tactical and

strategic operations. The generation of "quick-look" 2-D plots from Level 0 data products and "quick-look" 1-D plots from Level 1 data products are required for tactical operations. These "quick-look" products will be available to the SOWG through the ODS. A JPL server will distribute the CheMin Level 0 and 1 data products and the CheMin team-provided NASA Level 2 data products to the entire SOWG. PDLs will be responsible for delivering Level 1 data products back to the ODS via MIPL's File Exchange Interface (FEI) within 30 minutes after the delivery to the ODS through the DSN. RDR data products are created on the ODS and then deposited into FEI for electronic distribution to remote users via a secure subscription protocol.

After a data validation period, the CheMin RDR data products are collected with other science data and delivered to the Planetary Data System for archiving. [see reference 2 in section 1.3].

The sizes of CheMin RDR data files vary. "Quick-look" CheMin RDRs will be generated 30 minutes after the necessary EDRs have been released by MIPL. CheMin data will be reprocessed only if packets in the original downlink are not received. Partial files are created with missing data filled with zeroes. CheMin RDRs will be reprocessed after all data is retransmitted and received and the original version will be overwritten and placed into FEI for distribution.

2.4.4 Labeling and Identification

There is a file naming scheme adapted for the MSL image and non-image data products. The scheme applies to the EDR and several RDR data products. The file naming scheme adheres to the Level II 36.3 filename convention to be compliant with PDS standards. The file naming scheme also contains a minimal level of meta-data that retains uniqueness and is searchable.

All MSL EDR or RDR data products can be uniquely identified by incorporating into the product filename the Rover Mission identifier, the Instrument identifier, the Starting Spacecraft Clock count (SCLK) of the analysis event, the data Product Type, the Sol, the Site location, the rover Position within the site, the Sequence number, the venue, the Version and a Product Type extension. For non-camera data, including all CheMin data products, fields such as "camera eye" do not apply and are not used. The metadata fields have been selected based on MER and Phoenix mission lessons learned.

Each CheMin RDR has a detached PDS label associated with the CheMin data file. The file naming scheme for the CheMin EDR and RDR data products is formed by:

<instr><config><sclk><prod><sol><site><drive><seq><venue/who><ver>.<ext>

where,

instr = (2 alpha character) Instrument ID, denoting the source MSL science or engineering instrument that acquired the data.
Valid values for Instrument ID's are:

Valid values for:
"**CM**" - CheMin

Config/ spec = (2 alphanumeric) Instrument Configuration, an operational attribute of the Instrument that assists in characterizing the data.
Valid values for CheMin:

Instrument	Configuration	
	Values	Description
CheMin	"A_" "B_" "__"	RCE String Identification (identification of which Rover Compute Element, A or B, is used)

sclk = (9 alphanumeric) Spacecraft Clock Start Count, in units of seconds.

Which specific SCLK is used depends on the instrument but is generally expected to be the time the data was acquired. For CheMin, it is Data Product SCLK which matches the DVT (Data Validity Time) used for operational data management.

The valid values, in their progression, are as follows (non-Hex):

Range 000000000 thru 999999999 - "**000000000**", "**000000001**", **…** "**999999999**"
Range 1000000000 thru 1099999999 - "**A00000000**", "**A00000001**", **…** "**A99999999**"
Range 1100000000 thru 1199999999 - "**B00000000**", **"B00000001**", **…** **"B99999999**"

•
•
•

Range 3500000000 thru 3599999999 - "**Z00000000**", "**Z00000001**", **…** "**Z99999999**"

prod = (3 char) Product Type identifier.

This field has the following rule-of-thumb:

Beginning "**E**" - Type of **E**DR, which is the first order product with no processing applied,

Beginning "**D,R,** or **M**" - Type of RDR, which includes all products with processing applied,

The 2nd two characters identify the type of CheMin EDR or RDR.

Valid values for Product identifiers are listed below for EDRs:

EDR Product Type Description	Value
CCD Frame	"**ECC**"
Diffraction Single	"**ED1**"
Diffraction Split	"**EDS**"
Diffraction All	"**EDA**"
Energy All	"**EEA**"
Energy Single	"**EE1**"
Energy Split	"**EES**"
Film	"**EFM**"
Housekeeping N	**"EHK"**
Transmit Raw	"**ETR**"

Valid values for Product identifiers are listed below for RDRs:

RDR Product Type Description (see Section 2.4.2)	Value
Diffraction Single, windowed K-alpha or K-beta	"**RD1**"
Diffraction Split, windowed K-alpha or K-beta	"**RDS**"
Diffraction All, windowed K-alpha or K-beta	"**RDA**"
Diffraction All, Raw	"**RTR**"
Diffraction, Film Mode	"**RDF**"
Energy All	"**REA**"
Energy Single	"**RE1**"
Energy Split	"**RES**"
Level 2 product:	
Mineral Identification and Abundance	"**MIN**"

sol = (4 numeric) Sol for the SCLK value of the EDR.

site = (3 alphanumeric) Site location count, from the RMC.

This field has the following rules-of-thumb:

a) Site - If value is any 3 alphanumeric characters, or 3 underscores (denoting value is out-of-range), then content represents Site index extracted from RMC.

The valid Site values, in their progression, are as follows (non-Hex):

Range 000 thru 999 - "**000**", "**001**", **…** "**999**"
Range 1000 thru 1099 - "**A00**", "**A01**", **…** "**A99**"
Range 1100 thru 1199 - "**B00**", **"B01"**, **…** "**B99**"

•
•
•

Range 3500 thru 3599 - "**Z00**", "**Z01**", **…** "**Z99**"

drive = (4 alphanumeric) Drive (position-within-Site) location count, from the RMC.

This field has the following rules-of-thumb:

a) Drive - If value is any 4 alphanumeric characters, or 4 underscores (denoting value is out-of-range), then content represents Drive index extracted from RMC.

The valid Drive values, in their progression, are as follows (non-Hex):

Range 0000 thru 9999 - "**0000**", "**0001**", **…** "**9999**"
Range 10000 thru 10999 - "**A000**", "**A001**", **…** "**A999**"
Range 11000 thru 11999 - "**B000**", **"B001"**, **… "B999"**

•
•
•

Range 35000 thru 35999 - "**Z000**", "**Z001**", **…** "**Z999**"
Range 36000 thru 36099 - "**AA00**", "**AA01**", **…** "**AA99**"
Range 36100 thru 36199 - "**AB00**", "**AB01**", **…** "**AB99**"

•
•
•

Range 38500 thru 38599 - "**AZ00**", "**AZ01**", **…** "**AZ99**"
Range 38600 thru 38699 - "**BA00**", "**BA01**", **…** "**BA99**"

Range 38700 thru 38799 - "**BB00**", "**BB01**", **…** "**BB99**"
•
•
•
Range 41100 thru 41199 - "**BZ00**", "**BZ01**", **…** "**BZ99**"
Range 41200 thru 41299 - "**CA00**", "**CA01**", **…** "**CA99**"
•
•
•
Range 65400 thru 65499 - "**LI00**", "**LI01**", **…** "**LI99**"
Range 65500 thru 65535 - "**LJ00**", "**LJ01**", **…** "**LJ35**"

seq = (7 alphanumeric) Sequence identifier. Composed of a 2-char subfield (CH for CheMin) and a 5-digit numeric subfield representing the 6-bit "Category" and 14-bit numeric components of the commanded Sequence ID, respectively.

venue / who = (1 alpha character) Venue and Product Producer ID shared in the same field.

Venue denotes Flight Model versus Engineering Model in data acquisition. Product Producer ID identifies the institution that generated the product.

This field has the following rules-of-thumb:

a) Venue - A value in the range "**A** - **P**" indicates Flight Model rover. A value in the range "**Q** - **Z**" indicates Engineering (testbed) rover. The range "N - O" is not used.

b) Producer - If value is "**P**" (for Flight) or "**Y**" (for Engineering), the provider of the product is the Principal Investigator. Except for MIPL as the provider ("**M**" for Flight or "**Z**" for Engineering), the remaining characters are assigned to Co-investigator providers at the discretion of the P.I. and will be identified in due time. Within the instrument of the P.I., characters are unique. Across instruments, characters are reusable.

See the following list of valid values:

Venue		***by Producer***
Flight Model	*Eng. Model*	
*"**M**"*	*"**Z**"*	***M**IPL (OPGS at JPL)*
*"**P**"*	*"**Y**"*	***P**rincipal Investigator of Instrument …* *Instrument — Principal Investigator* *SAM — GSFC (Goddard, MD)* *REMS — Ministry of Education & Science (Spain)* *DAN — Federal Space Agency (Russia)* *RAD — SwRI (Boulder, CO)* *CheMin — Ames Research Center (Mountain View, CA)* *APXS — (Canada)* *SA/SPaH — JPL*
*"**A**" - "**L**"*	*"**Q**" - "**X**"*	*Co-Investigators (to be identified by P.I. per instrument)*

ver = *(1 alphanumeric) Version identifier.*

The valid values, in their progression, are as follows (non-Hex):

*Range 1 thru 10 - "**1**", "**2**", … "**9**", "**0**"*
*Range 11 thru 36 - "**A**", "**B**", … "**Z**"*

Range 37 and higher - "_" (underscore)

The Version number increments by one whenever an otherwise-identical filename would be produced. Note that not every version need exist, e.g. versions 1, 2 and 4 may exist but not 3. In general, the highest-numbered Version represents the "best" version of that product.

NOTE: To be clear, this field increments independently of all fields, including the Special Processing field.

ext = *(2 to 3 alpha characters) Product type extension.*

Valid values for nominal operations non-camera data products:

*"**IMG**" - Non-imaging instrument data Image.*
*"**LBL**" - Detached label in PDS format*
*"**TAB**" - fixed-width table data*
***"CSV"** comma-separated-value table data*

Example #1: CMA_013760215D1A00010930008CH01066M1.CSV

where,

instr	=	"**CM**"	=	CheMin
config	=	"**A**"	=	From RCE A string
spec	=	"_"	=	Always underscore
sclk	=	"**013760215**"	=	Spacecraft Clock Start Count of 13760215 secs
prod	=	"**D1A**"	=	Diffraction single, K-alpha
sol	=	"0001"	=	Sol 1
site	=	"**093**"	=	Site 93
drive	=	"**0008**"	=	Drive (Position-within-Site) 8
seqid	=	"**CH01066**"	=	CheMin sequence 01066
venue / who	=	"**M**"	=	Flight Model data / produced by MIPL (at JPL)
ver	=	"**1**"	=	Version 1
ext	=	"**CSV**"	=	CSV table data product with PDS label

2.5 Standards Used in Generating Data Products

2.5.1 PDS Standards

The CheMin RDRs comply with Planetary Data System standards for file formats and labels, as specified in the PDS Standards Reference [reference 7 in section 1.3] and the Planetary Science Data Dictionary Document [reference 9 in section 1.3].

2.5.2 Time Standards

The following time standards and conventions are used throughout this document, as well as the MSL project for planning activities and identification of events.

Time Format	*Definition*
SCET	Spacecraft event time. This is the time when an event occurred on-board the spacecraft, in Coordinated Universal Time (UTC). It is usually derived from SCLK.
SCLK	Spacecraft Clock. This is an on-board 64-bit counter, in units of nanoseconds and increments once every 100 milliseconds. Time zero corresponds to midnight on 1-Jan-1980.
ERT	Earth Received Time. This is the time when the first bit of the packet containing the current data was received at the Deep Space Network (DSN) station. Recorded in UTC format.
Local Mean Solar Time	Local Mean Solar Time (LMST). This is the local solar time defined by the local solar days (sols) from the landing date using a 24 "hour" clock within the current local solar day (HR:MN:SC). Since the Mars day is 24h 37m 22s long, each unit of LMST is slightly longer than the corresponding Earth unit. LMST is computed using positions of the Sun and the landing site from SPICE kernels. If a landing date is unknown to the program (e.g. for calibration data acquired on Earth) then no sol number will be provided on output LMST examples: SOL 12 12:00:01 SOL 132 01:22:32.498 SOL 29
RCT	Record Creation Time. This is the time when the first telemetry packet, containing a give data, set was created on the ground. Recorded in UTC format.
Local True Solar Time	This is related to LMST, which is also known as the mean solar time. It is the time of day based on the position of the Sun, rather than the measure of time based on midnight to midnight "day". LTST is used in all MIPL/OPGS generated products.
SOL	Solar Day Number, also known as PLANET DAY NUMBER in PDS label. This is the number of complete solar days on Mars since landing. The landing day therefore is SOL zero.

2.5.3 Coordinate Systems

The coordinate systems defined for MSL surface operations are listed in Table 5 and illustrated in Figure 6 below. Refer to the Pointing, Positioning, Phasing and Coordinate Systems document [reference 4 in section 1.3] for more details. The Rover coordinate system is defined relative to the Rover XYZ coordinate system (ROVER_FRAME). In terms of CheMin sample locations, each sample selected for drill or scoop collection will be tied to one or more images and other data (e.g., APXS) linked to a specific named target. The target name will be carried forward in the metadata within the CheMin sample label. Further discussion of coordinate systems can be found in Section 2.4.3 of the CheMin EDR SIS.

Table 4: Coordinate Frames Used for MSL Surface Operations

Frame Name (Label Keyword Value)	SHORT NAME (SAPP FDD)	REFERENCE FRAME (USED TO DEFINE)	Coordinate Frame	
			Origin	Orientation
ROVER_NAV_FRAME	RNAV	Enclosing SITE_FRAME	Attached to rover	Aligned with rover
ROVER_MECH_FRAME	RMECH	Enclosing SITE_FRAME	Attached to rover	Aligned with rover
LOCAL_LEVEL_FRAME	LL	Enclosing SITE_FRAME	Attached to rover (coincident with Rover Nav Frame)	North/East/Nadir
SITE_FRAME	SITE(n)	Previous SITE_FRAME	Attached to surface	North/East/Nadir
RSM_HEAD_FRAME	RSM_HEAD	ROVER_NAV_FRAME	Attached to mast head	Aligned with pointing of mast head. This corresponds to RSM_HEAD in the Frame Manager
Arm Frames: ARM_TURRET_FRAME ARM_DRILL_FRAME ARM_DRT_FRAME ARM_MAHLI_FRAME ARM_APXS_FRAME ARM_PORTION_FRAME ARM_SCOOP_TIP_FRAME ARM_SCOOP_TCP_FRAME	Arm Frames: TURRET DRILL DRT MAHLI APXS PORTION SCOOP_TIP SCOOP_TCP	ROVER_NAV_FRAME	Attached to the tool; see PPPCS for the specific tool frame.	Aligned with tool in some way; see PPPCS [Ref 1] for the specific tool Frame.

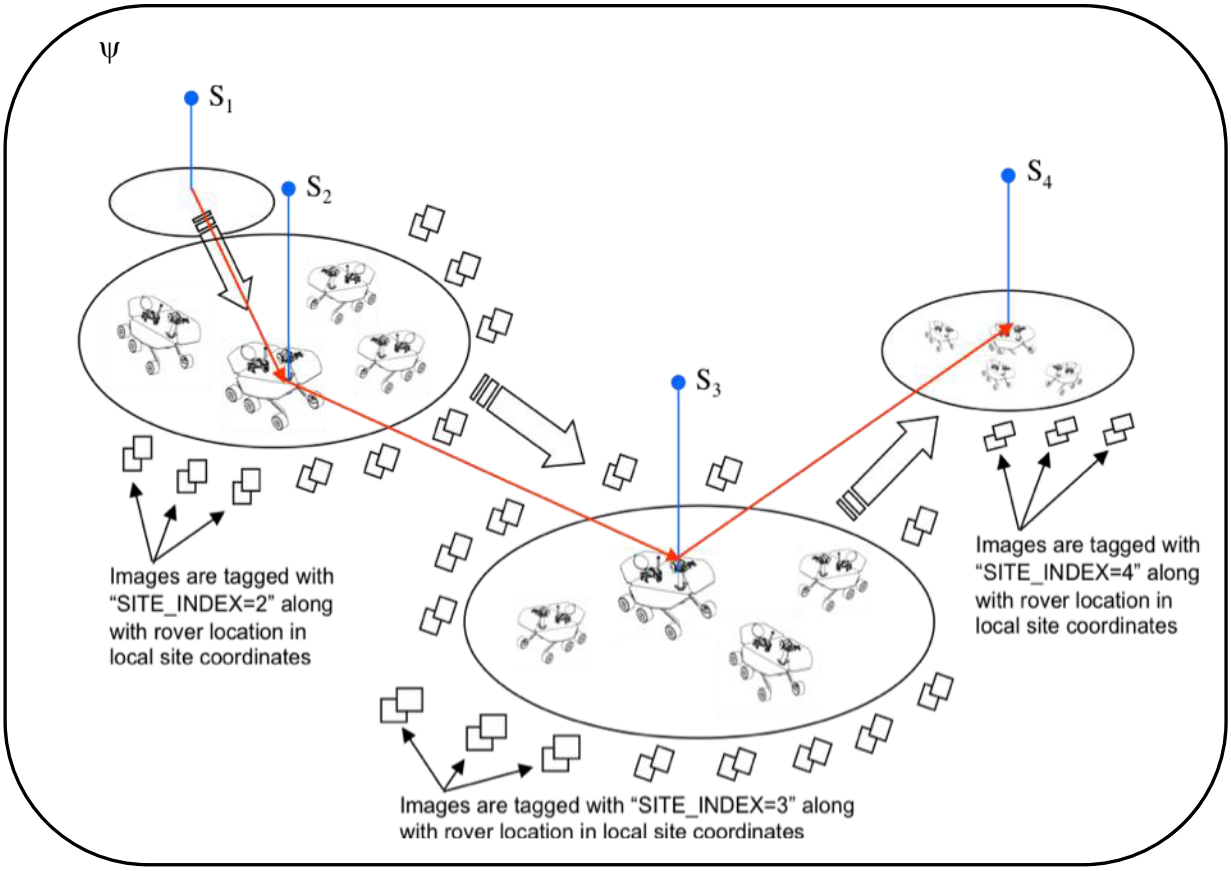

Figure 6: S, S_R, and Image Tagging System

2.5.4 Data Storage Conventions

The CheMin RDR data files contain ASCII text data only. The detached PDS labels for CheMin RDRs are also stored as ASCII text.

2.6 Data Validation

Validation of the MSL RDRs will not be automated and will be done manually.

Manual validation of the data will be performed throughout the life of the mission before each RDR is submitted to PDS, with follow-on validation of subsets of the data as the mission progresses and the MSL science team learns from and revisits accumulated data. These products will be viewed by a human. Initial validation of NASA Level 1B RDRs (all of which are 1D histograms) will include inspection of peak positions versus two-theta (XRD) or versus energy (Energy Dispersive Histogram (EDH)s) to insure that conversion from EDR products was done correctly. If necessary, the NASA Level 0 EDR products will be reprocessed to produce acceptable RDR products. Follow-on validation of Level 1B RDRs will be undertaken if accumulated experience in processing CheMin data leads to improved methods.

Initial validation of NASA Level 2 RDRs will be somewhat different, for these tables of mineral abundances are not 1D histograms but rather ASCII listings of mineral identifications with abundances and estimated errors in weight percent. Here the initial validation will consist of achieving the best possible fit of minerals to a Level 1B 1D diffraction pattern. If more than one type of Level 1B XRD product is obtained, the patterns may be each processed to Level 2 and compared to select a list of mineral abundances that is in the CheMin team's judgment most accurate. Follow-on validation of Level 2 RDRs may be common and will depend on several possible developments. Many such developments may be anticipated but three are listed here. First, it is likely that as the mission progresses the increasing familiarity with site mineralogy may prove that minerals first recognized by few or small diffraction peaks may be either confirmed or made suspect; in the latter case, earlier identifications may need to be revised. Second, there are many methods of background selection, pattern processing, and pattern fitting; the methods chosen for initial processing may be reevaluated as the mission progresses. Third, integrated results from the full instrument suite on *Curiosity* may indicate that certain minerals may be present but not initially uniquely identified in a particular sample; based on such information a revised XRD analysis may be produced based on fewer or smaller diffraction peaks than would be possible without such supporting data.

The design of MSL CheMin RDR products as specified in this document and in example products will be submitted to the PDS for peer review before operations begin, according to PDS policy for mission-generated data sets. When actual products are delivered to PDS during mission operations they will be checked to ensure that they conform to the peer-reviewed design.

3. DETAILED DATA PRODUCT SPECIFICATIONS

3.1 Data Product Structure and Organization

The structure of a CheMin RDR consists of a detached ASCII PDS label and an ASCII data file as shown in Figure 7.

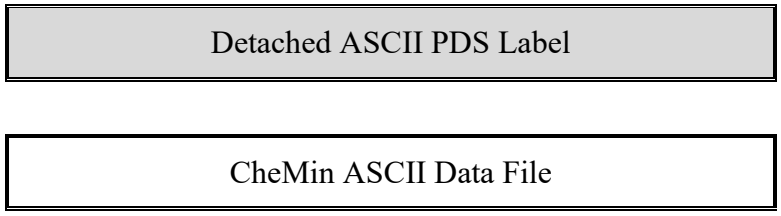

Figure 7: The CheMin RDR consists of two files.

3.2 Data Format Descriptions

CheMin RDR data files are derived from CheMin EDR files that consist of CheMin binary data with a detached ASCII PDS label. See the CheMin EDR SIS, Applicable Document 5, for descriptions of the EDR data files. All CheMin RDR files are ASCII text files formatted as

comma-separated-value (CSV) tables that are easily opened in many spreadsheet applications and in a text editor. Each CheMin RDR data file is accompanied by a detached PDS label in a file with the same name but the extension .LBL. The label describes the content and format of the data file. A label may include a pointer to another file, e.g. ^STRUCTURE = "CHEMIN_XRD.FMT". This file is called a format file, and includes specific information about the columns or fields in a data table.

Examples of CheMin labels, format files, and data files are given in Appendices A through D.

All of the CheMin RDRs listed in section 2.4.4, except tables of mineral identity and abundance (data product 'MIN'), are processed in histogram format, either as 1D diffraction patterns produced by image processing of 2D diffraction EDRs or as 1D energy histograms processed from EDR energy products by conversion of DN to energy.

3.3 Label and Header Descriptions

3.3.1 PDS Label

CheMin RDR data products have detached PDS labels stored as ASCII. A PDS label is object-oriented and describes the objects in the data file. The PDS label contains keywords for product identification and for data object definitions. The label also contains descriptive information needed to interpret or process the data objects in the file.

PDS labels are written in Object Description Language (ODL) [7]. PDS label statements have the form of "keyword = value". Each label statement is terminated with a carriage return character (ASCII 13) and a line feed character (ASCII 10) sequence to allow the label to be read by many operating systems. Pointer statements with the following format are used to indicate the location of data objects in the file:

^object = location

where the carat character (^, also called a pointer) is followed by the name of the specific data object. The location is the starting record number for the data object within the file.

A pointer may also be used with the STRUCTURE keyword to indicate another file containing a piece of the label, e.g.:

^STRUCTURE = "CHEMIN_XRD.FMT"

This example indicates that the contents of the file CHEMIN_XRD.FMT should be read as if they were inserted in the label at that point. This construct is often used with format files (extension .FMT) that contain the part of the label that describes the columns in a table or fields in a CSV file. As this description is the same for every product, it may be convenient to store it once in a separate file rather than repeat it in every label.

Each PDS keyword defined for the CheMin label will always be included in the PDS label. If a keyword does not have a value, a value of N/A will be given as the keyword value.

3.3.2 PDS Data Objects

A CheMin RDR consists of data objects that are described in the PDS label as spreadsheets. A PDS spreadsheet object is the same as a table except that the values in each row are separated by a delimiter character such as a comma, rather than being in a fixed column position in the row. The first record in the RDR data file contains the column names.

4. APPLICABLE SOFTWARE

4.1 Utility and Processing Programs

The X-ray fluorescence EDH data from CheMin are provided in PDS for analysis by comparison with standard tables of XRF energy for identification of fluorescence energies associated with specific elements (e.g., standard XRF energy tables available at lbl.gov/xray/). Further processing of EDH data to assess quantitative chemical composition is currently not within the scope of CheMin analysis. However, quantification of EDH results is a developmental topic and appropriate processing tools will be addressed in this RDR SIS if they become available.

Several standard programs for handling CheMin X-ray diffraction data are listed below. The examples provided are listed in an approximate sequence of increasing complexity and user knowledge, although ease of use is often subjective and depends on the user's background and training. Other programs for handling standard 1D X-ray diffraction data are available and use of such programs depends only on user preference.

XPOWDER. XPowder is a commercial program that supports phase identification using Powder Diffraction File (PDF) standard patterns. XPowder can also be used with data in Crystallographic Information File (CIF) format available in the American Mineralogist Crystal Structure Database (rruff.geo.arizona.edu/AMS/amcsd.php). Quantification of multi-phase samples may be performed using least-squares pattern-fitting and normalized reference intensity ratio (RIR) methods.

JADE. JADE is a commercial program that supports the comprehensive analysis of X-ray diffraction patterns, including phase identification, peak profile fitting, indexing, unit cell refinement, Rietveld analysis, etc. JADE has been developed by Materials Data Incorporated (MDI) and requires a commercial license.

TOPAS. TOPAS is a commercially available program (from Bruker Instruments) that uses 1D XRD data to conduct qualitative and quantitative mineralogical analysis. TOPAS requires a commercial license. TOPAS integrates many available profile fitting techniques as well as related applications such as single line fitting, whole powder pattern decomposition, ab-initio structure determination, and quantitative Rietveld analysis.

FULLPAT. FULLPAT is a quantitative X-ray diffraction methodology developed at Los Alamos National Laboratory (LANL) that merges the advantages of existing full-pattern fitting methods with the traditional reference intensity ratio (RIR) method. FULLPAT uses complete 1D diffraction patterns generated by CheMin, including the background, in conjunction with standard patterns measured on an equivalent instrument on Earth (the CheMin demonstration model, or DM). FULLPAT uses as standard-reference patterns either instrument-specific measured patterns (preferred) or calculated patterns. FULLPAT has been coded into Microsoft EXCEL using standard spreadsheet functions. The program is available through freeware and has already been tested by the CheMin team. Effective use of FULLPAT with CheMin data will require measurement of a set of standard patterns in the laboratory on the CheMin DM. Calculated patterns can also be used if they include simulation of the CheMin profile shapes. Practical use of FULLPAT and free access to FULLPAT for CheMin data will be cited in publications by the CheMin team as DM results become available.

GSAS. GSAS (General Structure Analysis System) is a comprehensive package for the refinement of structural models using both X-ray and neutron diffraction data. The GSAS

package can use 2D/1D powder diffraction data to perform Rietveld analysis, including analysis of peak profiles, refinement of unit-cell parameters, quantitative multicomponent analysis, and, in some cases, full crystal structure refinement. The package is in the public domain and requires no development. Similar analyses will be done with the program TOPAS (see above), which can include the ability to define fundamental instrument parameters that control measured profiles.

5. BACKGROUND TO APPENDICES

Appendices A through D include examples of representative diffraction and EDH RDRs, including label, format, and data table components of the RDRs. These examples include two forms of X-ray diffraction RDRs (Appendix A windowed for Co Kα and Appendix B in film mode), an EDH RDR in Appendix C, and an example of a level-2 mineral identification and abundance table in Appendix D. Although there are basically three different types of CheMin RDRs (diffraction, energy, and mineralogy), two forms of diffraction are included to clarify differences between energy-filtered (windowed to select for Co Kα or Kβ) and unfiltered "film mode" diffraction data. Figure 8 contains diffraction patterns in these two different modes for the mixture of 88% beryl and 12% quartz. This figure shows that the data product selected for Kα has a lower total background (by exclusion of all fluorescence) and an improved (greater) peak/background ratio. Moreover, the film-mode example has both a full Kβ pattern as well as a full Kα pattern (only the 100 beryl Kα and Kβ peaks are labeled, to avoid cluttering the figure). Although standard laboratory practice in X-ray diffraction is to filter out Kβ, handling a pattern with both Kα and Kβ patterns is tractable (though not preferred). Also, note that even with energy filtering for Kα a small amount of Kβ diffraction is still detectable in strong lattice scattering such as the beryl 100 diffraction. In these examples the data are normalized to 10,000 counts for the Kα line of the beryl 100 peak; notable is the higher relative intensity of the less intense peaks of the Kα diffraction pattern. This is because the Kα mode rejects more photon hits in the most intense diffraction, where photon flux is highest and there is more chance of photon events piling up in a pixel (multiple hits are rejected by the energy filter). These differences in relative intensities between the Kα and film mode should be taken into consideration in XRD analysis by full-pattern fitting.

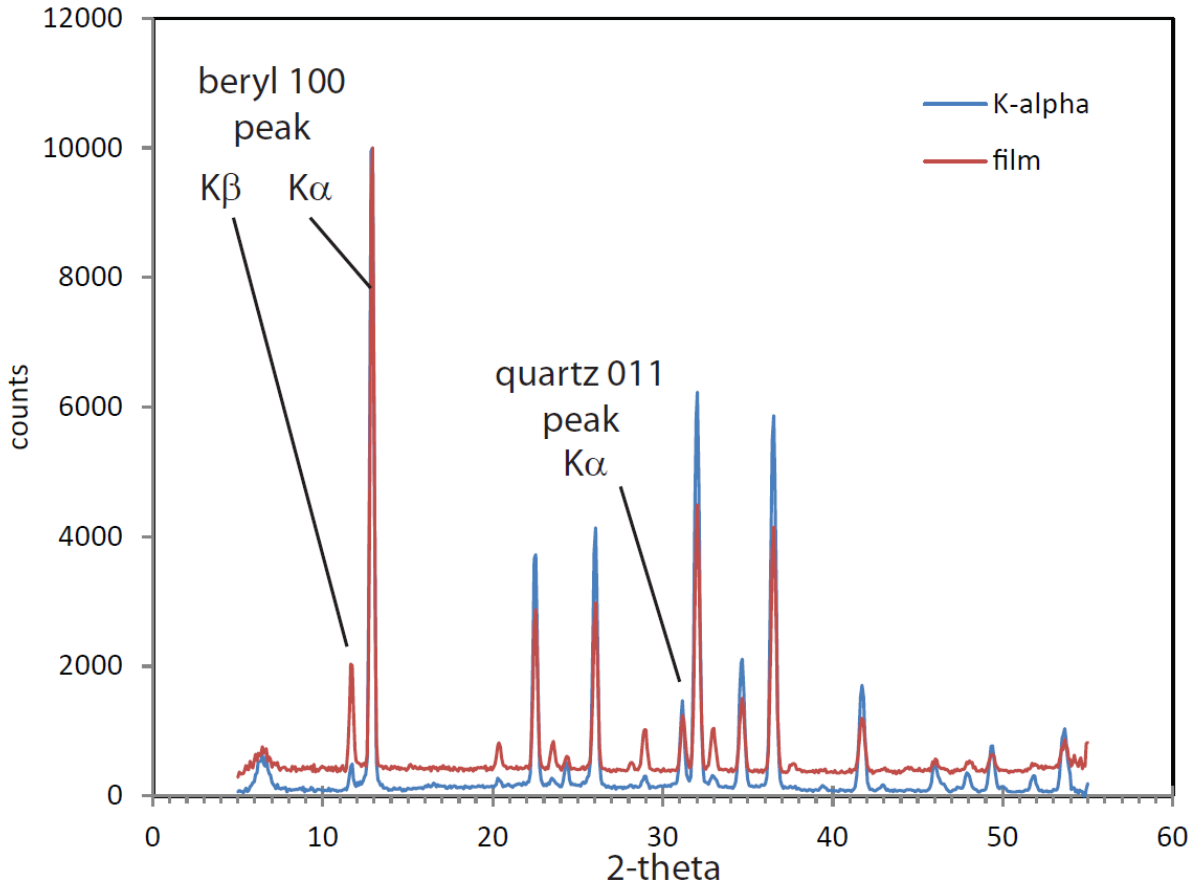

Figure 8: Comparative plots of windowed Kα diffraction (product type RD1, Appendix A3) and film-mode diffraction (product type RDF, Appendix B3). The two patterns represent the same 230 frames from an analysis of the 88:12 beryl:quartz standard on the CheMin FM (data collected in thermovac testing, Mars-pressure Ar atmosphere, 0 °C RAMP, -45 °C CCD). The beryl 100 Kα peak is normalized to 10,000 counts in both samples, in order to compare relative heights of peaks at higher 2-theta.

In practice, energy filtering for Kα will be the primary diffraction product from a CheMin analysis. Film mode, however, may be valuable in conditions where the CCD cannot be cooled to a low enough temperature to discriminate Kα energy, or in terminal stages of the mission where neutron damage to the CCD may have compromised performance.

Appendix A1: Example CheMin Product Type RD1 XRD RDR Label (Example using 88:12 Beryl:Quartz mixture, Single Pixel Co Kα)

```
PDS_VERSION_ID                = PDS3
RECORD_TYPE                   = STREAM
RECORD_BYTES                  = 12288
FILE_RECORDS                  = 981
^HEADER                      =
("CMA_987654321RD100090090009XXXXYYYYYP1.CSV",1)
^SPREADSHEET                 =
("CMA_987654321RD100090090009XXXXYYYYYP1.CSV",2)
DATA_SET_ID                   = "MSL-M-CHEMIN-4-RDR-V1.0"
PRODUCT_ID                    = "CMA_987564321RD100090090009XXXXYYYYYP1"
SOURCE_PRODUCT_ID             = "CMA_987654321ED100090090009XXXXYYYYYP1"
PRODUCT_TYPE                  = "CHEMIN_D1A"
INSTRUMENT_HOST_ID            = "MSL"
INSTRUMENT_HOST_NAME          = "MARS SCIENCE LABORATORY"
INSTRUMENT_ID                 = "CHEMIN"
TARGET_NAME                   = "MARS"
MSL:CALIBRATION_STANDARD_NAME = "BERYL_QUARTZ_88_12_STANDARD"
MISSION_PHASE_NAME            = "PRIMARY SURFACE MISSION"
PRODUCT_CREATION_TIME         = 2011-11-07T00:00:00
START_TIME                    = 2011-11-06T00:00:00
STOP_TIME                     = 2011-11-06T23:59:59
SPACECRAFT_CLOCK_START_COUNT  = "987654320"
SPACECRAFT_CLOCK_STOP_COUNT   = "987654321"

OBJECT                        = HEADER
 BYTES                        = 19
 HEADER_TYPE                  = TEXT
 DESCRIPTION                  = "This header record contains column
headings
 for the following table."
END_OBJECT                    = HEADER

OBJECT                        = SPREADSHEET
 ROWS                         = 981
 ROW_BYTES                    = 12288
 FIELDS                       = 2
 FIELD_DELIMITER              = "COMMA"
 ^STRUCTURE                   = "CHEMIN_XRD.FMT"
 DESCRIPTION                  = "This table contains single-pixel K-alpha
 diffraction data for the beryl:quartz 88:12 standard in CheMin cell
number 6b. The table represents a major frame consisting of 230 individual
CCD frames. CCD temperature was ~-45 degrees centigrade. Column 1 of the
table lists 2-theta from 3.00 to 51.95 degrees cobalt K-alpha, in
increments of 0.05 degrees (980 entries). Column 2 lists the intensity of
the diffraction for each 2-theta value in column 1."
END_OBJECT                    = SPREADSHEET
END
```

Appendix A2: Example CheMin Product Type RD1 XRD RDR Format File (Example using 8:12 Beryl:Quartz mixture, Single Pixel Co Kα)

```
OBJECT       = FIELD
 NAME        = "2-THETA"
 DATA_TYPE   = ASCII_REAL
 UNIT        = "DEGREES"
 BYTES       = 6
 FORMAT      = "F6.2"
END_OBJECT   = FIELD

OBJECT       = FIELD
 NAME        = "INTENSITY"
 DATA_TYPE   = ASCII_REAL
 UNIT        = "COUNTS"
 BYTES       = 7
 FORMAT      = "F7.0"
END_OBJECT   = FIELD
```

Appendix A3: Example CheMin Product Type RD1 XRD RDR Table (Example using 88:12 Beryl:Quartz mixture, Single Pixel Co Kα)

```
2-THETA,INTENSITY
3.00,57
3.05,83
3.10,71
3.15,65
3.20,50
3.25,79
3.30,69
3.35,67
3.40,82
3.45,70
...[lines omitted]...
51.55,58
51.60,88
51.65,35
51.70,87
51.75,71
51.80,65
51.85,90
51.90,88
51.95,101
```

Appendix B1: Example CheMin Product Type RDF XRD RDR Label (Example using 88:12 Beryl:Quartz mixture, from film data)

```
PDS_VERSION_ID               = PDS3
RECORD_TYPE                  = STREAM
RECORD_BYTES                 = 12288
FILE_RECORDS                 = 981
^HEADER                     =
("CMA_987654321RDF00090090009XXXXYYYYYP1.CSV",1)
^SPREADSHEET                =
("CMA_987654321RDF00090090009XXXXYYYYYP1.CSV",2)
DATA_SET_ID                  = "MSL-M-CHEMIN-4-RDR-V1.0"
PRODUCT_ID                   = "CMA_987564321RDF00090090009XXXXYYYYYP1"
SOURCE_PRODUCT_ID            = "CMA_987654321EFM00090090009XXXXYYYYYP1"
PRODUCT_TYPE                 = "CHEMIN_DAR"
INSTRUMENT_HOST_ID           = "MSL"
INSTRUMENT_HOST_NAME         = "MARS SCIENCE LABORATORY"
INSTRUMENT_ID                = "CHEMIN"
TARGET_NAME                  = "MARS"
MSL:CALIBRATION_STANDARD_NAME = "BERYL_QUARTZ_88_12_STANDARD"
MISSION_PHASE_NAME           = "PRIMARY SURFACE MISSION"
PRODUCT_CREATION_TIME        = 2011-11-07T00:00:00
START_TIME                   = 2011-11-06T00:00:00
STOP_TIME                    = 2011-11-06T23:59:59
SPACECRAFT_CLOCK_START_COUNT  = "987654320"
SPACECRAFT_CLOCK_STOP_COUNT   = "987654321"

OBJECT                       = HEADER
 BYTES                       = 19
 HEADER_TYPE                 = TEXT
 DESCRIPTION                 = "This header record contains column
headings
 for the following table."
END_OBJECT                   = HEADER

OBJECT                       = SPREADSHEET
 ROWS                        = 981
 ROW_BYTES                   = 12288
 FIELDS                      = 2
 FIELD_DELIMITER             = "COMMA"
 ^STRUCTURE                  = "CHEMIN_XRD.FMT"
 DESCRIPTION                 = "This table contains film diffraction data
for the beryl:quartz 88:12 standard in CheMin cell number 6b. The table
represents a major frame consisting of 229 individual CCD frames. CCD
temperature was ~-45 degrees centigrade. Column 1 of the table lists 2-
theta from 3.00 to 51.95 degrees cobalt K-alpha, in increments of 0.05
degrees (980   entries). Column 2 lists the intensity of the diffraction
for each 2-theta value in column 1."
END_OBJECT                   = SPREADSHEET
END
```

Appendix B2: Example CheMin XRD Product Type RDF RDR Format File (Example using 8:12 Beryl:Quartz mixture, from film data)

```
OBJECT       = FIELD
 NAME        = "2-THETA"
 DATA_TYPE   = ASCII_REAL
 UNIT        = "DEGREES"
 BYTES       = 6
 FORMAT      = "F6.2"
END_OBJECT   = FIELD

OBJECT       = FIELD
 NAME        = "INTENSITY"
 DATA_TYPE   = ASCII_REAL
 UNIT        = "COUNTS"
 BYTES       = 7
 FORMAT      = "F7.0"
END_OBJECT   = FIELD
```

Appendix B3: Example CheMin Product Type RDF XRD RDR Table ((Example using 88:12 Beryl:Quartz mixture, from film data)

```
2-THETA,INTENSITY
3.00,287
3.05,370
3.10,331
3.15,347
3.20,339
3.25,396
3.30,467
3.35,408
3.40,401
3.45,535
3.50,456
3.55,430
3.60,550
3.65,639
...[lines omitted]...
51.20,728
51.25,577
51.30,558
51.35,454
51.40,508
51.45,628
51.50,527
51.55,500
51.60,453
51.65,552
51.70,588
51.75,414
51.80,489
51.85,459
51.90,809
51.95,826
```

Appendix C1: Example CheMin Product Type RE1 EDH Histogram RDR Label (Example using Amphibole Single Pixel Data)

```
PDS_VERSION_ID                = PDS3
RECORD_TYPE                   = STREAM
RECORD_BYTES                  = 255
FILE_RECORDS                  = 4096
^HEADER                      =
("CMA_987564321RE100090090009XXXXYYYYYP1.CSV",1)
^SPREADSHEET                  =
("CMA_987564321RE100090090009XXXXYYYYYP1.CSV",2)
DATA_SET_ID                   = "MSL-M-CHEMIN-4-RDR-V1.0"
PRODUCT_ID                    = "CMA_987564321RE100090090009XXXXYYYYYP1"
SOURCE_PRODUCT_ID             = "CMA_987564321EE100090090009XXXXYYYYYP1"
PRODUCT_TYPE                  = "CHEMIN_RE1"
INSTRUMENT_HOST_ID            = "MSL"
INSTRUMENT_HOST_NAME          = "MARS SCIENCE LABORATORY"
INSTRUMENT_ID                 = "CHEMIN"
TARGET_NAME                   = "MARS"
MSL:CALIBRATION_STANDARD_NAME = "AMPHIBOLE_STANDARD"
MISSION_PHASE_NAME            = "PRIMARY SURFACE MISSION"
PRODUCT_CREATION_TIME         = 2011-11-07T00:00:00
START_TIME                    = 2011-11-06T00:00:00
STOP_TIME                     = 2011-11-06T23:59:59
SPACECRAFT_CLOCK_START_COUNT  = "987654320"
SPACECRAFT_CLOCK_STOP_COUNT   = "987654321"

OBJECT                        = HEADER
 BYTES                        = 18
 HEADER_TYPE                  = TEXT
 DESCRIPTION                  = "This header record contains column
headings
   for the following table."
END_OBJECT                    = HEADER

OBJECT                        = SPREADSHEET
 ROWS                         = 4095
 ROW_BYTES                    = 255
 FIELDS                       = 2
 FIELD_DELIMITER              = "COMMA"
 ^STRUCTURE                   = "CHEMIN_EDH.FMT"
 DESCRIPTION                  = "This table contains single-pixel
fluorescence energy data for the amphibole standard in CheMin cell
number 14a. The table represents total fluorescence intensities
from a major frame consisting of 740 10-second exposures.
CCD temperature was ~-40 degrees centigrade. Column 1 of the table
lists energy in keV derived from CCD DN values that result in an
energy range from 0.00735 to 30.10235 keV, in increments of 7.351
eV (4095 entries). A conversion factor of 7.351 eV per DN was used,
based on the Co Ka peak position. EDH intensities have been set to
zero for all energies of 0.35020 and lower, to suppress CCD noise at
energies well below the range of CCD measurement. Column 2 lists the
intensity of EDH counts for each energy value in column 1."
END_OBJECT                    = SPREADSHEET
```

```
END
```

Appendix C2: Example CheMin Product Type RE1 EDH Histogram RDR Format File (Example using Amphibole Single Pixel Data)

```
OBJECT      = FIELD
 NAME       = "ENERGY"
 DATA_TYPE  = ASCII_REAL
 UNIT       = "KEV"
 BYTES      = 8
 FORMAT     = "F8.5"
END_OBJECT  = FIELD

OBJECT      = FIELD
 NAME       = "INTENSITY"
 DATA_TYPE  = ASCII_REAL
 UNIT       = "COUNT"
 BYTES      = 8
 FORMAT     = "F8.6"
END_OBJECT  = FIELD
```

Appendix C3: Example CheMin Product Type RE1 EDH Histogram RDR Table (Example using Amphibole Single Pixel Data)

```
ENERGY,INTENSITY
0.00735,0
0.01470,0
0.02205,0
0.02940,0
0.03676,0
0.04411,0
0.05146,0
0.05881,0
0.06616,0
0.07351,0
0.08086,0
0.08821,0
0.09556,0
0.10291,0
0.11027,0
...[lines omitted]...
29.98473,0
29.99208,0
29.99943,0
30.00678,0
30.01413,0
30.02148,0
30.02884,0
30.03619,0
30.04354,0
30.05089,0
30.05824,0
30.06559,0
30.07294,0
30.08029,0
30.08764,0
30.09499,0
30.10235,0
```

Appendix D1: Example CheMin Mineral Identification and Abundance RDR Label, Product Type MIN, using FMFT Sample Harder

```
PDS_VERSION_ID                = PDS3
RECORD_TYPE                   = STREAM
RECORD_BYTES                  = 255
FILE_RECORDS                  = 6
^HEADER                      =
("CMA_987654321MIN00090090009XXXXYYYYYP1.CSV",1)
^TABLE                       =
("CMA_987654321MIN00090090009XXXXYYYYYP1.CSV",2)
DATA_SET_ID                   = "MSL-M-CHEMIN-5-RDR-V1.0"
PRODUCT_ID                    = "CMA_987654321MIN00090090009XXXXYYYYYP1"
MSL:SAMPLE_NAME               = "HARDER"
PRODUCT_TYPE                  = "CHEMIN_MIN"
INSTRUMENT_HOST_ID            = "MSL"
INSTRUMENT_HOST_NAME          = "MARS SCIENCE LABORATORY"
INSTRUMENT_ID                 = "CHEMIN"
TARGET_NAME                   = "MARS"
MISSION_PHASE_NAME            = "PRIMARY SURFACE MISSION"
PRODUCT_CREATION_TIME         = 2011-11-07T00:00:00
START_TIME                    = 2011-11-06T00:00:00
STOP_TIME                     = 2011-11-06T23:59:59
SPACECRAFT_CLOCK_START_COUNT  = "987654320"
SPACECRAFT_CLOCK_STOP_COUNT   = "987654321"

OBJECT                        = HEADER
 BYTES                        = 23
 HEADER_TYPE                  = TEXT
 DESCRIPTION                  = "This header record contains column
headings
   for the following table."
END_OBJECT                    = HEADER

OBJECT                        = SPREADSHEET
 ROWS                         = 5
 ROW_BYTES                    = 255
 FIELDS                       = 3
 FIELD_DELIMITER              = "COMMA"
 ^STRUCTURE                   = "CHEMIN_MIN.FMT"
 DESCRIPTION                  = "
   This table contains mineral identifications and abundances for
   the Fast Motion Field Test (FMFT) sample named “Harder”. The analysis
   uses single-pixel XRD data in a major frame consisting of ten
   minor frames, each minor frame consisting of 50 15-second exposures.
   The 2D CCD pattern was processed into a 1D pattern of intensity
   versus 2-theta using the program FILMSCAN. Mineral identifications
   and abundances were determined using single-pixel K-alpha diffraction
   data for Harder. Mineral identifications and mineral abundance
   determinations were obtained using the program TOPAS. Abundances
   are based on full-pattern fitting for all phases except smectite;
   smectite abundance is estimated based on comparative measure of 001
   diffraction. Column 1 of the table lists mineral identifications.
   Column 2 lists the mineral abundances in weight percent. Column 3
   lists estimated errors in mineral abundance determinations.
```

```
"
END_OBJECT                      = SPREADSHEET
END
```

Appendix D2: Example CheMin Mineral Identification and Abundance Product Type MIN RDR Format File using FMFT Sample Harder

```
OBJECT         = FIELD
 NAME          = "MINERAL"
 DATA_TYPE     = CHARACTER
 UNIT          = "TEXT"
 BYTES         = 16
 FORMAT        = "A16"
END_OBJECT     = FIELD

OBJECT         = FIELD
 NAME          = "PERCENT"
 DATA_TYPE     = ASCII_REAL
 UNIT          = "WEIGHT_PERCENT"
 BYTES         = 9
 FORMAT        = "F9.2"
END_OBJECT     = FIELD

OBJECT         = FIELD
 NAME          = "ERROR"
 DATA_TYPE     = ASCII_REAL
 UNIT          = "ESTIMATED_ERROR"
 BYTES         = 7
 FORMAT        = "F7.2"
END_OBJECT     = FIELD
```

Appendix D3: Example CheMin Mineral Identification and Abundance Product Type MIN RDR Table using FMFT Sample Harder

```
MINERAL,PERCENT,ERROR
QUARTZ,40.00,0.81
SMECTITE,15.00,5.00
KAOLINITE,42.00,0.81
PYRITE,0.25,0.23
ANATASE,1.80,0.34
```