

# James Webb Space Telescope Project

## Science Requirements Document

**December 3, 2004**

**JWST GSFC CMO**

March 11, 2005

**RELEASED**



**Goddard Space Flight Center  
Greenbelt, Maryland**

**National Aeronautics and  
Space Administration**

CHECK WITH JWST DATABASE AT:  
<https://ngst1.hst.nasa.gov/>  
TO VERIFY THAT THIS IS THE CORRECT VERSION PRIOR TO USE.

## CM FOREWORD

This document is a James Webb Space Telescope (JWST) Project Configuration Management (CM)-controlled document. Changes to this document require prior approval of the JWST Project Manager. Proposed changes shall be submitted to the JWST CM Office (CMO), along with supportive material justifying the proposed change. Changes to this document will be made by complete revision.

Questions or comments concerning this document should be addressed to:

JWST Configuration Manager  
JWST Configuration Management Office  
Mail Stop 443  
Goddard Space Flight Center  
Greenbelt, Maryland 20771

CHECK WITH JWST DATABASE AT:

<https://ngst1.hst.nasa.gov/>

TO VERIFY THAT THIS IS THE CORRECT VERSION PRIOR TO USE.

**Signature Page**

***Prepared by:***

Original signed by \_\_\_\_\_ 3/8/2005  
Dr. John Mather Date  
JWST Senior Project Scientist  
NASA/GSFC 685

***Approved by:***

Original signed by \_\_\_\_\_ 3/8/2005  
Phil Sabelhaus Date  
JWST Project Manager  
NASA/GSFC 443

***Concurred by:***

Original signed by \_\_\_\_\_ 3/11/2005  
Eric Smith Date  
JWST Program Scientist Universe Division,  
Science Mission Directorate  
NASA Headquarters

**JAMES WEBB SPACE TELESCOPE PROJECT**

**DOCUMENT CHANGE RECORD**

Sheet: 1 of 1

REV LEVEL	DESCRIPTION OF CHANGE	APPROVED BY	DATE APPROVED
Initial	Release per JWST-CCR-000209	J. Decker	2/18/2005

CHECK WITH JWST DATABASE AT:

<https://ngst1.hst.nasa.gov/>

TO VERIFY THAT THIS IS THE CORRECT VERSION PRIOR TO USE.

**List of TBDs/TBRs**

<b>Item No.</b>	<b>Location</b>	<b>Summary</b>	<b>Ind./Org.</b>	<b>Due Date</b>
1	Section 8.18	SR-31, When requested the Observatory shall track targets which exhibit any angular velocity in the range of 30 mas/s over a total motion 30 arcsec with respect to the guide star ( <b>TBR</b> ).	P. Sabelhaus/GSFC	5/31/2005
2	Appendix B	SR-22, Add to ISIM ( <b>TBD</b> )	Pam Sullivan/GSFC	5/9/2005
3	Appendix B	SR-40, Add to ISIM ( <b>TBD</b> )	Pam Sullivan/GSFC	5/9/2005
4	Appendix B	SR-23, Add to ISIM ( <b>TBD</b> )	Pam Sullivan/GSFC	5/9/2005

CHECK WITH JWST DATABASE AT:

<https://ngst1.hst.nasa.gov/>

TO VERIFY THAT THIS IS THE CORRECT VERSION PRIOR TO USE.

## TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
<b>1.0 PURPOSE AND SCOPE.....</b>	<b>1-1</b>
<b>2.0 SCIENTIFIC INTRODUCTION .....</b>	<b>2-1</b>
2.1 The End of the Dark Ages: First Light and Reionization .....	2-1
2.2 The Assembly of Galaxies .....	2-2
2.3 The Birth of Stars and Protoplanetary Systems .....	2-3
2.4 Planetary Systems and the Origins of Life .....	2-3
<b>3.0 THE END OF THE DARK AGES: FIRST LIGHT AND REIONIZATION.....</b>	<b>3-1</b>
3.1 Impact on Science .....	3-1
3.1.1 General Understanding .....	3-3
3.1.2 Major Open Questions .....	3-4
3.1.3 Specific Objectives .....	3-5
3.2 Previous Investigations .....	3-6
3.2.1 Observational Foundation .....	3-6
3.2.2 Theoretical Predictions .....	3-9
3.3 Measurements Required from JWST .....	3-11
3.3.1 Detection of First Luminous Objects at Very High Redshifts .....	3-11
3.3.2 Establishing when Reionization Occurred and Identifying the Reionizers.....	3-13
3.4 Summary .....	3-16
<b>4.0 THE ASSEMBLY OF GALAXIES.....</b>	<b>4-1</b>
4.1 Impact on Science .....	4-1
4.2 Previous Investigations .....	4-2
4.3 Measurements required from JWST .....	4-4
4.3.1 Where were Stars in the Hubble Sequence Galaxies Formed, When did Luminous Quiescent Galaxies Appear, and How does this Process Depend on the Environment? .....	4-4
4.3.2 Where and When are the Heavy Elements Produced and to What Extent do Galaxies Exchange Material with the Intergalactic Medium? .....	4-6
4.3.3 When and How are the Global Scaling Relations for Galaxies Established?.....	4-7
4.3.4 Do Luminous Galaxies Form through the Hierarchical Assembly of Dark Matter Halos?.....	4-8
4.3.5 What are the Redshifts and Power Sources of the High Redshift Ultra Luminous Infrared Galaxies?.....	4-9
4.3.6 What is the Relation Between the Evolution of Galaxies and the Growth and Development of Black Holes in Their Nuclei? .....	4-10
4.4 Summary .....	4-11
<b>5.0 THE BIRTH OF STARS AND PROTOPLANETARY SYSTEMS .....</b>	<b>5-1</b>
5.1 How do Protostellar Clouds Collapse? .....	5-2
5.1.1 Scientific Rationale.....	5-2
5.1.2 Key JWST Observations.....	5-5

5.2	What is The Early Evolution of Protostars? .....	5-6
5.2.1	Scientific Rationale.....	5-6
5.2.2	Key JWST Observations.....	5-8
5.3	How do massive stars form and affect their Environment? .....	5-8
5.3.1	Scientific Rationale.....	5-8
5.3.2	Key JWST Observations.....	5-10
5.4	What is the Initial Mass Function at sub-stellar Masses? .....	5-12
5.4.1	Scientific Rationale.....	5-12
5.4.2	Key JWST Observations.....	5-14
5.5	How do Protoplanetary Systems form? .....	5-16
5.5.1	Scientific Rationale.....	5-16
5.5.2	Key JWST Observations.....	5-18
5.6	What are the life cycles of gas and dust? .....	5-18
5.6.1	Scientific Rationale.....	5-18
5.6.2	Key JWST Observations.....	5-20
5.7	Summary .....	5-22
<b>6.0</b>	<b>PLANETARY SYSTEMS AND THE ORIGINS OF LIFE.....</b>	<b>6-1</b>
6.1	Origins of Planetary Systems.....	6-1
6.1.1	Questions.....	6-1
6.1.2	Observations .....	6-4
6.1.3	Observatory Capabilities.....	6-11
6.2	Early Evolution of Planetary Systems .....	6-12
6.2.1	Questions.....	6-12
6.2.2	Observations .....	6-13
6.2.3	Observatory Capabilities.....	6-15
6.3	Transition to Life .....	6-16
6.3.1	Questions.....	6-16
6.3.2	Observations .....	6-17
6.4	Summary .....	6-18
<b>7.0</b>	<b>REQUIREMENTS RATIONALE .....</b>	<b>7-1</b>
7.1	Wavelength range .....	7-1
7.2	Imaging and Coronagraphy.....	7-1
7.3	Spectroscopy .....	7-2
7.4	Near-Infrared Camera .....	7-2
7.5	Near-Infrared Spectrograph .....	7-2
7.6	Mid-Infrared Instrument .....	7-2
7.7	Tunable Filters .....	7-3
7.8	Minimum Collecting Area .....	7-3
7.9	Image Quality.....	7-3
7.9.1	Strehl Ratio .....	7-3
7.9.2	Encircled Energy.....	7-3
7.9.3	Point-Spread Function Symmetry.....	7-4
7.9.4	Point-Spread Function Stability.....	7-4
7.10	Sensitivity .....	7-4

7.11	Stray light.....	7-5
	7.11.1 Diffuse Stray Light .....	7-5
	7.11.2 Local Stray Light and Dynamic Range.....	7-5
7.12	Long Exposures .....	7-6
7.13	Photometric Accuracy and Calibration.....	7-6
7.14	Field of View .....	7-7
7.15	Angular Resolution.....	7-7
7.16	Pointing Accuracy.....	7-7
7.17	Field of Regard .....	7-8
7.18	Moving Object Tracking.....	7-8
7.19	Mission Lifetime.....	7-9
7.20	Observing Efficiency .....	7-9
7.21	Data Rate.....	7-9
7.22	Targets of Opportunity.....	7-9
7.23	Public Access.....	7-10
<b>8.0</b>	<b>SCIENCE REQUIREMENTS.....</b>	<b>8-1</b>
8.1	Wavelength Range.....	8-2
8.2	Imaging and Coronagraphy.....	8-2
8.3	Spectroscopy.....	8-2
8.4	Near-Infrared Camera .....	8-3
8.5	Near-Infrared Spectrograph .....	8-3
8.6	Mid-Infrared Instrument .....	8-3
8.7	Medium and Narrow-band Imaging.....	8-3
8.8	Primary Mirror Area .....	8-3
8.9	Image Quality.....	8-4
	8.9.1 Strehl Ratio .....	8-4
	8.9.2 Encircled Energy.....	8-4
	8.9.3 Point Spread Function Symmetry .....	8-4
	8.9.4 Point Spread Function Stability .....	8-4
8.10	Sensitivity .....	8-4
8.11	Stray Light .....	8-5
	8.11.1 Diffuse Stray Light .....	8-5
	8.11.2 Local Stray Light .....	8-5
8.12	Long Exposures .....	8-5
8.13	Photometric Accuracy and Calibration.....	8-6
8.14	Field of View .....	8-6
8.15	Angular Resolution.....	8-6
8.16	Pointing Accuracy.....	8-6
8.17	Field of Regard .....	8-6
8.18	Moving Object Tracking.....	8-7
8.19	Mission Lifetime.....	8-7
8.20	Observing Efficiency .....	8-7
8.21	Data Rate.....	8-7
8.22	Targets of Opportunity.....	8-7

8.23 Public Access ..... 8-8

**APPENDIX A: ABBREVIATIONS AND ACRONYMS ..... A-1**

**APPENDIX B: SCIENCE TO MISSION REQUIREMENT TRACEABILITY  
SUMMARY .....B-1**

**FIGURES**

<b><u>Figure</u></b>	<b><u>Page</u></b>
Figure 3-1. Schematic view of cosmic history.....	3-2
Figure 3-2. Cumulative galaxy counts for $z \sim 6$ .....	3-4
Figure 3-3. The Hubble Ultra Deep Field.....	3-8
Figure 3-4. SEDs of Type II Supernovae at Maximum Light .....	3-13
Figure 4-1. Analysis of HST Images .....	4-5
Figure 4-2. Spectrum of a galaxy at $z \sim 0.5$ .....	4-7
Figure 4-3. ISO Circinus Spectrum .....	4-10
Figure 5-1. How a Single, Isolated, Low-Mass Star and its Planetary System are Formed .....	5-2
Figure 5-2. The Low-Mass Dark Cloud Barnard 68.....	5-4
Figure 5-3. Azimuthally Averaged Radial Dust Column Density Profile for B68.....	5-5
Figure 5-4. The Spectral Energy Distribution of the Prototypical Class 0 Protostar.....	5-7
Figure 5-5. The M 16 Elephant Trunks .....	5-10
Figure 5-6. Color-Magnitude Diagram for the Orion Trapezium Cluster .....	5-13
Figure 5-7. Young Circumstellar Disks in Orion.....	5-17
Figure 5-8. Mid-Infrared Spectra of Young Stars and Circumstellar Disks .....	5-19
Figure 6-1. Comets and Circumstellar Disks.....	6-3
Figure 6-2. Detectability of Extra-Solar Giant Planets and Brown Dwarfs .....	6-5
Figure 6-3. Spectra of the Closest T-dwarfs .....	6-6
Figure 6-4. Detectability of a Jovian-sized Exoplanet.....	6-7
Figure 6-5. Guide to Spectral Features in Extrasolar Giant Planets .....	6-8
Figure 6-6. Seeing the Effect of Planets on Dust Disks.....	6-9
Figure 6-7. Models and Observations of a Dusty Disk.....	6-11
Figure 6-8. Simulated KBO Spectrum.....	6-14
Figure 6-9. Sensitivity of MIRI for KBOs.....	6-15

**TABLES**

<b><u>Table</u></b>	<b><u>Page</u></b>
Table 3-1. JWST Measurements for the End of the Dark Ages Theme .....	3-16
Table 3-2. Required Capabilities for the End of the Dark Ages Theme.....	3-17
Table 4-1. JWST Measurements for the Assembly of Galaxies Theme.....	4-12
Table 4-2. Required Capabilities for the Assembly of Galaxies Theme .....	4-13

Table 5-1. Predicted Fluxes for Sub-Stellar Objects ..... 5-15

Table 5-2. JWST Measurements for the Birth of Stars Theme..... 5-22

Table 5-3. Required Capabilities for the Birth of Stars Theme ..... 5-23

Table 6-1. Model Fluxes of Bound Giant Planets..... 6-4

Table 6-2. JWST Measurements for the Planetary Systems Theme..... 6-18

Table 6-3. Required Capabilities for the Planetary Systems Theme ..... 6-19

Table 8-1. Required Sensitivity Values ..... 8-5

Table 8-2. Required Calibration Accuracies..... 8-6

## 1.0 **PURPOSE AND SCOPE**

The purpose of this document is to define the scientific objectives and requirements of the James Webb Space Telescope (JWST) Project. JWST will be a large, cold, infrared-optimized space telescope designed to enable fundamental breakthroughs in our understanding of the formation and evolution of galaxies, stars, and planetary systems.

This document provides detail for the Level 1 requirements which form part of the JWST Program Plan controlled at the National Aeronautics and Space Administration (NASA) Headquarters. It was written by the JWST Science Working Group under the direction of the JWST Senior Project Scientist, and is controlled by the JWST Project. It provides the rationale for the requirements contained in the Mission Requirements Document (MRD). If there is a conflict between this document and the MRD, the MRD requirement prevails.

This document takes precedence over the instrument Science Rationale and Analysis documents (SRAs) which are controlled by the instrument science teams. The Near-Infrared Camera (NIRCam) SRA is controlled by the NIRCam Investigation Definition Team. The Mid-Infrared Instrument (MIRI) SRA is controlled by the MIRI Science Team. The Fine Guidance Sensor-Tunable Filter (FGS-TF) SRA is controlled by the Canadian Space Agency (CSA). The Near-Infrared Spectrograph (NIRSpec) SRA is controlled by the European Space Agency (ESA).

We base these requirements on current theoretical understanding, interpretation of Hubble Space Telescope (HST) observations, and results from studies using ground-based and other space-based facilities. Many classes of targets for JWST have never been observed before, so the scientific requirements are necessarily derived from theoretical predictions and extrapolations from known objects. In these cases we outline the basis of the predictions and derive the needed observatory capabilities to verify them.

At the time of writing, the Spitzer Space Telescope has recently been launched, and promises great advances in all of the scientific areas outlined in this document. We expect that results from Spitzer, additional results from HST, and other advances in theory and observation will further refine the observational plans for JWST. The science requirements in this document define the required measurement capabilities of the telescope, but they do not require that particular observations be made. A mission which provides these capabilities will support a wide variety of astrophysical investigations. JWST is a facility-class mission, and most of the observing time will be allocated to investigators from the international astronomical community through competitively selected proposals.

This document supersedes previous editions developed by the Ad-Hoc Science Working Group (ASWG) and the Interim Science Working Group (ISWG). In particular, it supersedes the Design Reference Mission (DRM) document, spreadsheets, and computer programs, which were developed by the ASWG and ISWG as examples of observing programs. JWST is consistent with the scientific program for a “Large Infrared-Optimized Space Telescope”, described in the report “HST and Beyond,” (Dressler 1996), and the program for the “Next Generation Space Telescope”, which was given top priority by the National Academy of Sciences survey “Astronomy and Astrophysics in the New Millennium” (McKee & Taylor, 2001).

## **2.0 SCIENTIFIC INTRODUCTION**

The scientific objectives of the JWST fall into four themes:

1. The End of the Dark Ages: First Light and Reionization
2. The Assembly of Galaxies
3. The Birth of Stars and Protoplanetary Systems
4. Planetary Systems and the Origins of Life

These themes require observations specified in the JWST Program Plan Mission Success Criteria (Level 1 Baseline Science Requirements), quoted here:

“4.1.1.1 Measure the space density of galaxies to a 2  $\mu\text{m}$  flux density limit of  $1.0 \times 10^{-34} \text{ W m}^{-2} \text{ Hz}^{-1}$  via imagery within the 0.6 to 27  $\mu\text{m}$  spectral band to enable the determination of how this density varies as a function of their age and evolutionary state.

4.1.1.2 Measure the spectra of at least 2500 galaxies with spectral resolutions of approximately 100 (over 0.6 to 5  $\mu\text{m}$ ) and 1000 (over 1 to 5  $\mu\text{m}$ ) and to a 2  $\mu\text{m}$  emission line flux limit of  $5.2 \times 10^{-22} \text{ W m}^{-2}$  to enable determination of their redshift, metallicity, star formation rate, and ionization state of the intergalactic medium.

4.1.1.3 Measure the physical and chemical properties of young stellar objects, circumstellar debris disks, extra-solar giant planets, and Solar System objects via spectroscopy, and imagery within the 0.6 to 27  $\mu\text{m}$  spectral band to enable determination of how planetary systems form and evolve.”

### **2.1 THE END OF THE DARK AGES: FIRST LIGHT AND REIONIZATION**

Theory and observation have given us a simple picture of the early Universe. The Big Bang produced (in decreasing order of present mass-energy density): dark energy (the cosmic acceleration force), dark matter, hydrogen, helium, cosmic microwave and neutrino background radiation, and trace quantities of lithium, beryllium, and boron. As the Universe expanded and cooled, hydrogen molecules formed, enabling the formation of the first individual stars. The Universe has expanded by a factor of about 20 since that time, the mean density was about 8000 times greater than it is now, and the age was about 180 million years. According to theory, these first stars were 30 to 300 times as massive as the Sun and millions of times as bright, burning for only a few million years before meeting a violent end. Each one produced either a core-collapse supernova (type II) or a black hole. The supernovae enriched the surrounding gas with the chemical elements produced in their interiors, and future generations of stars contained these heavier elements. The black holes started to swallow gas and other stars to become mini-quasars, growing and merging to become the huge black holes now found at the centers of nearly all massive galaxies. The supernovae and the mini-quasars should be observable by the JWST. Both might also be sources of gamma ray bursts and gravity wave bursts that could be discovered by other observatories and then observed by JWST.

Some time after the appearance of the first sources of light, hydrogen in the intergalactic medium was reionized. Results from the Wilkinson Microwave Anisotropy Probe (WMAP) combined with data from the Sloan Digital Sky Survey show that this reionization occurred at two or more epochs, perhaps with a complex history. Although there are indications that galaxies produced the majority of the ultraviolet radiation which caused the reionization, the contribution of quasars could be significant.

JWST must address several key questions:

- What are the first galaxies?
- When and how did reionization occur?
- What sources caused reionization?

To find the first galaxies, JWST must make ultra-deep near-infrared surveys of the Universe, and follow-up with low-resolution spectroscopy and mid-infrared photometry. To study reionization, high resolution near-infrared spectroscopy is required.

## **2.2 THE ASSEMBLY OF GALAXIES**

Galaxies are the basic building blocks of the Universe. Theory and observation show that galaxies are assembled through a process of the hierarchical merging of dark matter concentrations. Small objects formed first, and were drawn together to form larger ones. This dynamical build-up of massive systems is accompanied by chemical evolution, as the gas and dust within the galaxies are processed through successive generations of stars. The interaction of these luminous components with the invisible dark matter produces the beautiful appearance and diverse properties of present-day galaxies. This process is still occurring today, as the Magellanic Clouds fall into the Milky Way, and as the Andromeda Nebula heads toward the Milky Way for a future collision. Galaxies have been observed back to times about one billion years after the Big Bang. While most of these early galaxies are smaller and more irregular than present-day galaxies, some early galaxies are very similar to those seen nearby today. This is a surprise.

Despite all the work done to date, many questions are still open. We do not really know how galaxies are formed, what controls their shapes, what makes them form stars, how the chemical elements are generated and redistributed through the galaxies, whether the central black holes exert great influence over the galaxies, or what are the global effects of violent events as small and large parts join together in collisions.

JWST must address several key questions:

- When and how did the Hubble Sequence form?
- How did the heavy elements form?
- How were the global scaling relations established?
- What is the role of ULIRGs and AGN in galaxy evolution?

To answer these questions, JWST must observe galaxies back to their earliest precursors so that we can understand their growth and evolution. JWST must provide imaging and spectroscopy over the 0.6 to 27  $\mu\text{m}$  band to meet this objective.

### **2.3 THE BIRTH OF STARS AND PROTOPLANETARY SYSTEMS**

While stars have been the main topic of astronomy for thousands of years, only in recent times have we begun to understand them with detailed observations and computer simulations. A hundred years ago we did not know that they are powered by nuclear fusion, and 50 years ago we did not know that stars are continually being formed. We still do not know the details of how they are formed from clouds of gas and dust, or why most stars form in groups, or how planets form with them. The stars within a star-forming region interact with each other both chemically and dynamically in complex ways. The details of how they evolve and liberate the “metals” back into space for recycling into new generations of stars and planets remains to be determined through a combination of observation and theory.

Observation shows that most stars are formed in multiple star systems and that many stars have planets. However, there is little agreement about how this occurs, and the discovery of large numbers of massive planets in very close orbits around their stars was very surprising. We now know that planets are common around late-type (cooler and less massive than the Sun) stars, and that debris disks might reveal the presence of planets.

JWST must address several key questions:

- How do protostellar clouds collapse?
- How does environment affect star formation and vice versa?
- What is the low-mass initial mass function of stars?
- How do gas and dust coalesce to form planetary systems?

To unravel the birth and early evolution of stars, from infall onto dust-enshrouded protostars, through the genesis of planetary systems, JWST must provide near- and mid-IR imaging and spectroscopy to observe these objects.

### **2.4 PLANETARY SYSTEMS AND THE ORIGINS OF LIFE**

Understanding the origin of the Earth and its ability to support life is a key objective for astronomy and is central to the JWST science program. Key parts of the story include understanding the formation of small objects and how they combine to form large ones, learning how they reach their present orbits, learning how the large planets affect the others in systems like ours, and learning about the chemical and physical history of the small and large objects that formed the Earth and delivered the necessary chemical precursors for life. The cool objects and dust in the outer Solar System are evidence of conditions in the early Solar System, and are directly comparable to cool objects and dust observed around other stars.

JWST must address several key questions:

- How do planets form?
- How are circumstellar disks like our Solar System?
- How are habitable zones established?

JWST must determine the physical and chemical properties of planetary systems including our own, and investigate the potential for the origins of life in those systems. JWST must provide near- and mid-IR imaging and spectroscopy to observe these objects.

### **3.0 THE END OF THE DARK AGES: FIRST LIGHT AND REIONIZATION**

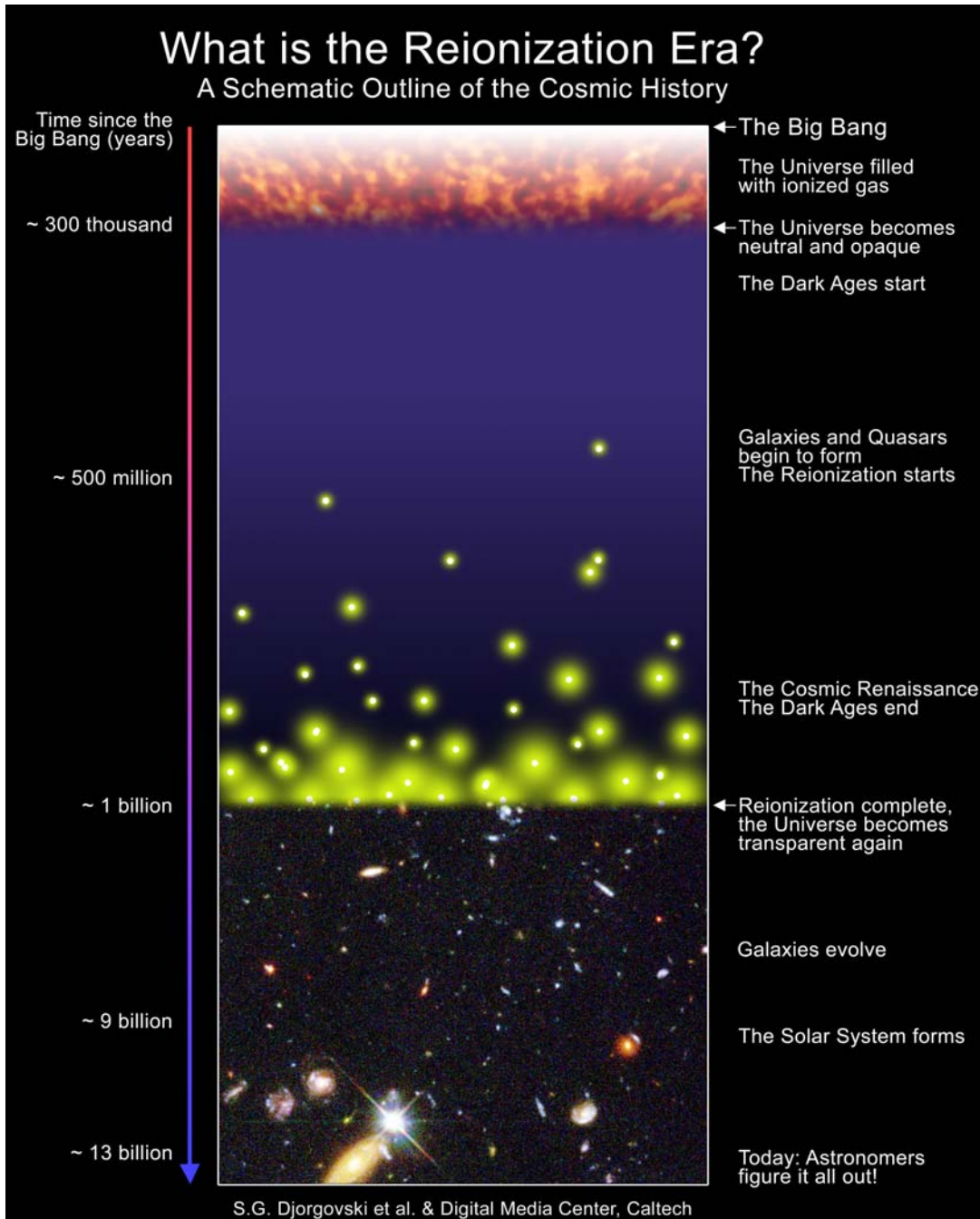
The key objective of The End of the Dark Ages: First Light and Reionization theme is to identify the first luminous sources to form and to determine the ionization history of the early Universe.

#### **3.1 IMPACT ON SCIENCE**

The emergence of the first sources of light in the Universe marks the end of the "Dark Ages" in cosmic history (Figure 3-1), a period characterized by the absence of discrete sources of light. Understanding these first sources is critical since they greatly influence subsequent structures. The current leading models for structure formation predict a hierarchical assembly of galaxies and clusters. The first sources of light act as seeds for the successive formation of larger objects and from their study we will learn the processes that formed the nuclei of present day giant galaxies.

This epoch is currently under intense theoretical investigation. The formation of structure in the Dark Ages is easier to study theoretically than similar processes occurring at other epochs because: i) the formation of the first structures is directly linked to the growth of linear perturbations, and ii) these objects have known metal abundances set by the end-product of the primordial nucleosynthesis. By studying this epoch it is possible to probe the power spectrum of density fluctuations emerging from recombination at scales smaller than those accessible by current cosmic microwave background experiments.

Some time after the appearance of the first sources of light, hydrogen in the Universe is reionized. We do not know the time lag between these two events nor whether reionization is brought about by the first light sources themselves or by subsequent generations of objects. Reionization is by itself a period in cosmic history as interesting as the emergence of first light. The epoch of reionization is the most recent and perhaps the most accessible of the global phase transitions undergone by the Universe after the Big Bang.



**Figure 3-1. Schematic view of cosmic history**

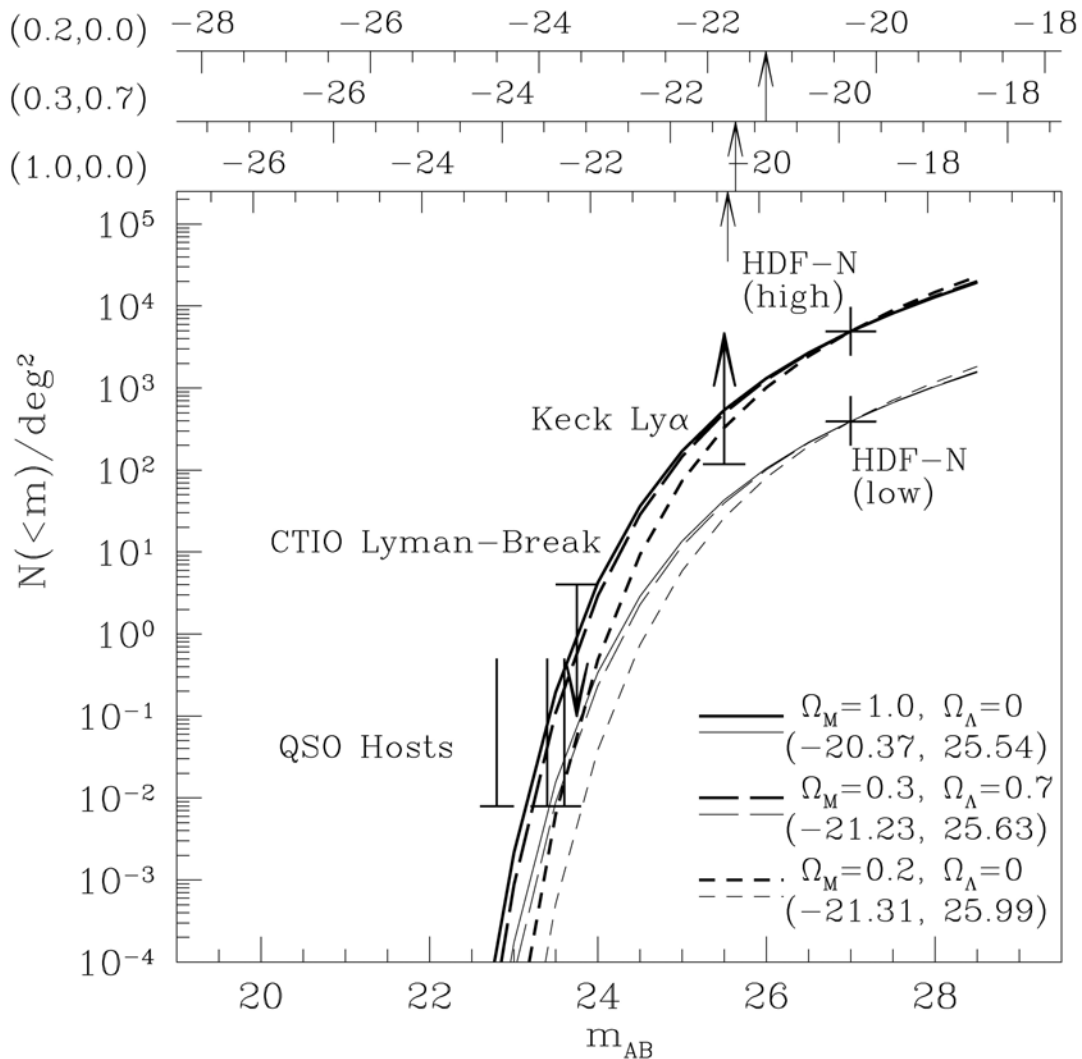
NOTE: Cosmic history is shown from recombination 300 kyr after the Big Bang ( $z \sim 1089$ ), through the first light sources era ( $< 500$  Myr after the Big Bang, or  $z > 9$ ), through the epoch of reionization (1 Gyr after the Big Bang, or  $z \sim 6$ ), to the present day (13.7 Gyr after the Big Bang, or  $z = 0$ ).

### 3.1.1 General Understanding

Before the first light, ionizing photons and metals are essentially absent. Thus, hydrogen molecules can form and survive to become the primary cooling agent of the first perturbation to collapse. We expect stars formed from this process to be very massive and very hot. Historically this primordial stellar population has been given the name of Population III. The most overdense peaks in the perturbations emerging from recombination will collapse first. Hierarchical-clustering models predict that these overdense regions sit in larger mass, lower contrast but still overdense regions. We may expect the first stars to be markers of, and possibly reside in, stellar clusters or even small galaxies.

At the end of their short lives, the very massive first stars will leave black holes as remnants. These black holes will begin accreting gas and form mini-AGN. In more exotic models, black holes can form directly from the collapse of perturbations rather than as stellar remnants. Thus, in these models the first sources of light would be powered by gravitational accretion rather than nuclear fusion.

Soon after the first light sources appear, both stars and accretion onto black holes become viable sources of ionizing radiation. We do not know which are primarily responsible for reionizing hydrogen in the surrounding intergalactic medium (IGM). AGN produce an energetic spectrum, and would reionize helium as well as hydrogen. Because observational evidence reveals that helium is reionized at a much later time, hydrogen was reionized by stars. However, it is possible that helium recombines after being reionized for the first time together with hydrogen. A second reionization of helium would occur during the epoch when quasar activity peaks. More recently, a combination of observations by the WMAP of the cosmic microwave background polarization (Kogut et al 2003) with spectra of  $z=6$  quasars found by the Sloan Digital Sky Survey (SDSS) (Fan et al 2001; 2002) has revealed the possibility that there were two reionization epochs for hydrogen (Cen 2003a; 2003b). In these models, the completion of the reionization epoch that is seen at  $z = 6$  would be that of the second reionization, with the first reionization taking place during or after the peak of the first light epoch at higher redshift.



**Figure 3-2. Cumulative galaxy counts for  $z \sim 6$**

NOTE: Galaxy counts for  $z \sim 6$  are predicted on the basis of lower-redshift measurements. In this figure, the AB magnitudes refer to a band from 9100 to 9800 Å. Different curves refer to different cosmologies and different normalizations of the luminosity function. At AB=28, one expects a few  $z \geq 6$  galaxies per arcmin square but this number is uncertain by about an order of magnitude. JWST will go fainter by three magnitudes and reach completely uncharted territory (Yan et al. 2002).

**3.1.2 Major Open Questions**

- When did the first luminous sources arise and what was their nature? What were their clustering properties?
- When did reionization occur? What is the reionization history of the Universe prior to the final reionization?

- What were the sources responsible for reionization? Were they powered by nuclear fusion or gravitational accretion?

### **3.1.3 Specific Objectives**

#### **3.1.3.1 Detection of First Luminous Objects at Very High Redshifts**

To directly detect the first luminous objects and to identify the redshift when they appear we will study the evolution of the number density of objects  $N(z)$ , and the evolution of the star formation rate  $SFR(z)$  as a function of redshift. A complementary method is to study the evolution of the mean metallicity of galaxies  $\langle Z \rangle(z)$ . Once candidates first light objects are identified, they will be studied in detail to confirm their nature.

##### **3.1.3.1.1 Evolution of $N(z)$**

There will be no objects more distant than the first objects that formed, and so  $N(z)$  will reach zero beyond the redshift of formation. A strong upper limit on the number density of objects at redshifts greater than that of the most distant object observed is an indication that we have detected the first light objects.

##### **3.1.3.1.2 Evolution of $SFR(z)$**

Instead of detecting galaxies one can aim at determining the star formation rate as a function of redshift by measuring the SN rate, and if GRB's are sufficiently numerous, the GRB rate.

##### **3.1.3.1.3 Evolution of $\langle Z \rangle(z)$**

The metallicity of first light objects should be zero, while non-zero metallicity indicates that the object formed from gas that has already been enriched. For the brightest objects one is able to obtain spectra. At low metallicity ratios of oxygen lines to Balmer lines, as, e.g.,  $[OIII]/H\beta$  are a linear measure of metallicity.

##### **3.1.3.1.4 Confirm the Nature of First Light Objects**

A small sample of candidate first light objects will be studied in detail in order to place strong upper limits on their metal content and to prove the absence of an older stellar population by measuring their optical rest frame spectral energy distribution. Alternatively, identifying the age of an older stellar population sets a lower limit to the redshift of the first star formation. This limit becomes easier to measure at high redshift, since stellar ages of tens to hundreds of million years are an increasingly larger fraction of the cosmic time.

#### **3.1.3.2 Establishing When Reionization Occurred and Identifying the Reionizers**

The classical way of identifying reionization is through the detection of signatures in the Lyman  $\alpha$  forest: a black Lyman  $\alpha$  Gunn-Peterson trough, islands in the Lyman  $\alpha$  forest, and appearance of a Lyman  $\alpha$  damping wing. In addition to these techniques, the epoch of reionization can be identified as the redshift at which there is fast evolution of the properties of Lyman  $\alpha$  emitters. However, a sharp transition in the Lyman  $\alpha$  luminosity function can be suppressed if, for

instance, a relatively long reionization onset is coupled to a smooth increase in metal content. Sources at higher redshifts will have increasingly more absorbed Lyman  $\alpha$  but also increasingly stronger intrinsic equivalent widths because of the lower metallicity. It is easy to build models where the two effects cancel out. Alternative methods, not sensitive to this limitation, are the study of the evolution of the ratio between Lyman  $\alpha$  and Balmer lines. Once the reionization epoch is identified, one needs to find a population of objects that have sufficient ionizing continuum to ionize all of the hydrogen. Once these sources are identified, one can derive their properties and determine their nature and energy source.

### **3.1.3.2.1 Lyman $\alpha$ Forest Diagnostics**

Deep observations of QSOs or bright galaxies would detect black Lyman  $\alpha$  Gunn-Peterson troughs, islands in the Lyman  $\alpha$  forest, and appearance of Lyman  $\alpha$  damping wings.

### **3.1.3.2.2 Transition in the Properties of Lyman $\alpha$ Sources**

When the Universe is still neutral, Lyman  $\alpha$  is efficiently scattered over a large volume or absorbed by dust. The faintest Lyman  $\alpha$  sources and those with narrow Lyman  $\alpha$  emission will not be visible before reionization. Thus, at reionization, one expects a fast evolution of the faint end of the Lyman  $\alpha$  luminosity function of star-forming objects.

### **3.1.3.2.3 Transition in the Ratio between Lyman $\alpha$ and Balmer Lines**

If neither the metallicity nor the dust content of the Universe change abruptly at reionization, then detection of a rapid change in the Lyman  $\alpha$  to H $\alpha$  (or H $\beta$ ) ratio can be used to identify the reionization epoch.

### **3.1.3.2.4 Measuring the Ionizing Continuum**

The ionizing continuum can be derived indirectly by estimating its slope and intensity. The slope can be derived from the ratio between hydrogen and helium lines. The hydrogen Balmer lines can provide the intensity.

### **3.1.3.2.5 Determine the Source Nature**

A combination of spectroscopic diagnostics (line shapes, line widths, and line intensity ratios) can be used to distinguish between stellar and non-stellar photo-ionization.

## **3.2 PREVIOUS INVESTIGATIONS**

### **3.2.1 Observational Foundation**

The most direct observational evidence of re-ionization is the detection of a Gunn-Peterson trough (Gunn & Peterson 1965) in the spectrum of high redshift quasars. Neutral hydrogen clouds along the line of sight (the Lyman  $\alpha$  forest) produce increasing absorption as the redshift increases. Still, even at  $z \approx 5$  some signal is detected below Lyman  $\alpha$ , suggesting that re-ionization occurs at higher redshifts.

Recently, Fan et al (2001) detected high redshift quasars using the Sloan Digital Sky Survey, including a few around  $z \cong 6$ . QSO SDSSp J103027.10 0552455.0 at  $z = 6.28$  shows a drop in continuum flux below Lyman  $\alpha$  by a factor 150. Other QSOs at slightly lower redshift show a much smaller continuum drop. This is evidence that a Gunn-Peterson trough has been detected in this object (Becker et al. 2001, Fan et al. 2002).

We cannot conclude that the reionization epoch has been determined on the basis of a single object, particularly since even a very modest neutral hydrogen column density could produce the observed trough. However, this detection opens up the possibility that re-ionization occurs at a relatively low redshift. There are few constraints on the number density of galaxies at redshift greater than 6. By extrapolating the luminosity function of Lyman break galaxies at lower redshift (Steidel et al. 1999) one can obtain predictions for the number of galaxies at  $z \sim 6$  (Figure 3-2; Yan et al. 2002), which are at the level of a few  $AB = 28$  galaxies per arcmin square.



**Figure 3-3. The Hubble Ultra Deep Field**

NOTE: The Hubble ultra-deep field was imaged with the Advanced Camera for Surveys onboard HST. The UDF will probably be the deepest survey before JWST.

In the years before the launch of JWST, progress with HST and large ground based telescopes will allow us to study the bright end of the luminosity function of galaxies at  $z > 6$ . However, these facilities will be unable to push beyond  $z = 8$ , measure the internal properties of these objects, or characterize the population of galaxies.

Cosmic microwave background fluctuations as measured by WMAP support a much earlier reionization of hydrogen (at  $z \cong 20 \pm 10$ , Kogut et al. 2003). This may be an indication that hydrogen recombined, perhaps partially, to be reionized at a lower redshift.

In contrast to the reionization of hydrogen reionization, the epoch of helium reionization has been firmly identified at  $z \sim 3$  through the detection of a Gunn-Peterson trough in quasar spectra (Jakobsen et al. 1994, Davidsen et al. 1996, Heap et al. 2000).

### 3.2.2 Theoretical Predictions

In the standard Cold Dark Matter (CDM) cosmology galaxies are assembled through hierarchical merging of building blocks with smaller mass. The first such building blocks, with  $M \geq 10^4 M_{\odot}$  form in these models at  $z \geq 15$  (Couchman & Rees 1986, Haiman & Loeb 1997, Ostriker & Gnedin 1996, Haiman, Thoul & Loeb 1996, Abel et al. 1998, Abel et al. 2000).

While we do not know whether the first sources of light are powered by nuclear energy from fusion reactions or by gravitational accretion (Haiman & Loeb 1999), it is possible that population III stars are responsible for the reionization of hydrogen at  $z \cong 6-20$  (Madau & Shull 1996, see also Haiman & Loeb 1999, Gnedin & Ostriker 1997, Chiu & Ostriker 2000). Efficient cooling by  $H_2$  molecules and an early, vigorous formation of massive objects could result in reionization at redshifts as early as  $z \cong 20$  (Cen 2003b, Haiman & Holder 2003).

It is generally believed that the harder UV spectrum of AGNs is responsible for the reionization of helium (Jakobsen et al. 1994) at lower redshifts.

Nuclear processing of only a  $2 \times 10^{-6}$  fraction of baryons would be sufficient to reionize the Universe (Loeb and Barkana 2001), leading to a minimum average metallicity of the Universe at reionization of  $\sim 10^{-3} Z_{\text{sun}}$ . It is not clear what the mean metallicity of objects observed at these redshifts would be. Indeed, the metallicity of the first objects and that of the Inter-Galactic Medium (IGM) could be very different. If population III stars are formed in halos of sufficiently low mass they can enrich the IGM by SN-driven winds (Madau et al. 2001, Mori et al. 2002). When a halo undergoes a SN-driven outflow, the ejection of metals can be very effective. However, it is not clear how effective this process is when averaged over all halos. It is possible that the most massive halos retain most of their metals and have much higher metallicities at the epoch of reionization.

It is possible to compute the minimum surface brightness required to reionize the Universe, under the assumptions that the Universe is reionized by hot population III stars, and all UV photons can escape the system. This minimum surface brightness of ionizing sources at  $z > 6$  is  $AB_{1400} \cong 29 \text{ mag arcmin}^{-2}$  in a redshifted  $\lambda = 1400 \text{ \AA}$  band (Stiavelli et al. 2004), when counted as the typical ionizing flux seen per unit area. For a luminosity function similar in shape to that of  $z = 3$  Lyman break galaxies and with  $M_{\star}$  not fainter than  $-15 \text{ mag}$ , this implies a few sources per square arcmin with  $AB=31$  or brighter.

If the power source for reionization is not nuclear fusion but rather gravitational accretion onto massive black holes, the higher efficiency of gravitational accretion requires a smaller fraction of material to be processed. This scenario does not place any constraint on the metallicity of the

Universe at reionization. Even if reionization is caused by stellar UV radiation, it is natural to expect that some fraction of these stars will leave black holes as remnants (Heger & Woosley 2002, Madau & Rees 2001). Both scenarios lead to the presence of seed black holes at the end of re-ionization, with implications for the formation of AGN and galaxies (Silk & Rees 1998, Stiavelli 1998).

Even though one often refers to reionization as if it were a sudden transition, the time elapsed between the epochs when 10% and 90% of hydrogen was reionized can last a significant fraction of the age of the Universe at reionization. If the preliminary measurement from WMAP of a significant Compton opacity is confirmed, this would be evidence of either a very extended reionization process or of two distinct reionization epochs (Cen 2003a, 2003b; Haiman & Holder 2003; Holder et al. 2003, Stiavelli et al. 2004). Regardless of the specifics of the reionization process, inhomogeneities along the line of sight may create dispersion in optical depth shortwards of Lyman  $\alpha$ . Moreover, only a very low residual fraction of neutral hydrogen is needed to produce a Gunn-Peterson trough in the spectra of high redshift quasars. In addition, the opacity near Lyman  $\alpha$  would be modified in the neighborhood of ionizing sources (Miralda-Escudé & Rees 1994), in analogy to the proximity effect in QSOs (Moller & Kjaergaard 1992).

The very first stars (population III) have zero metallicity. In the absence of any metals, cooling is dominated by the less effective  $H_2$  cooling, which leads to the formation of very massive objects, with masses exceeding  $100 M_{\odot}$  (Bromm et al. 1999, Bromm et al. 2002) and possibly going up to  $1000 M_{\odot}$ . The spectral energy distribution (SED) of these massive stars resembles a black body with an effective temperature around  $10^5$  K (Bromm et al. 2001). Due to their high temperatures, these stars are very effective at ionizing both hydrogen and helium. It should be noted that, even at lower mass, zero-metallicity stars are expected to be much hotter than their solar metallicity analogues (Tumlinson & Shull 2000). A single  $1000 M_{\odot}$  star at  $z=30$  has a magnitude of  $AB \approx 36$  in its non-ionizing continuum.

Two direct consequences of the high effective temperature of zero-metallicity stars are their effectiveness in ionizing hydrogen (and helium) and their low optical-to-UV fluxes. Both tend to make the direct detection of the stellar continuum much harder than the detection of the associated HII region. In the surrounding HII region, electron temperatures exceed 20,000 K and 45% of the total luminosity is emitted in the Lyman  $\alpha$  line, resulting in a Lyman  $\alpha$  equivalent width (EW) of 3000 Å (Bromm et al. 2001, Panagia et al. 2003). The helium lines are also rather strong, with the intensity of HeII  $\lambda 1640$  comparable to that of H $\beta$  (Panagia et al. 2003, Tumlinson et al. 2001).

An interesting feature of these models is that the HII region emission longward of Lyman  $\alpha$  is dominated by a strong two-photon nebular continuum. The H $\alpha$ /H $\beta$  ratio for these models is 3.2. Both the red continuum and the high H $\alpha$ /H $\beta$  ratio could be naively (and incorrectly) interpreted as a consequence of dust extinction, even though no dust is present in these systems.

By estimating the brightness of the sources that enriched the IGM to  $10^{-2} Z_{\text{sun}}$  (Miralda-Escudé and Rees 1998), one finds a combined surface brightness of about  $AB = 32 \text{ mag arcsec}^{-2}$ . This value of surface brightness is about 2 orders of magnitude higher than the values estimated above. For any reasonable luminosity function these sources would be either easily detected

directly, or by exploiting amplification by gravitational lensing from an intervening cluster of galaxies. Their large number offers the promising prospect of identifying first light by observing a decrease in the number of sources identified for increasing redshifts.

The number of supernovae expected before reionization also depends crucially on the assumptions made about the nature of the ionizing sources. Based on relatively normal stellar populations and a metallicity of  $5 \times 10^{-3} Z_{\text{sun}}$  at the end of reionization, one arrives at an estimate of about one SN arcmin<sup>-2</sup> at any give time (Miralda-Escude and Rees 1997). If the ionizers are very massive population III stars and the metallicity at the end of reionization is lower, then one could have a SN rate up to one hundred times smaller. However, the SNe with a very massive population III progenitor could be much brighter than regular type II SNe.

While models differ significantly in the details of how the reionization was started by these various possible First Light populations at  $15 < z < 25$ , they all converge to produce roughly the same cosmic star-formation history (SFH) of population II stars in the mini halos of dwarf galaxies at  $6 < z < 10$ . The latter is simply the consequence of the need to fit the nearly complete Gunn-Peterson troughs now seen in the spectra of at least four SDSS quasars in the range  $6.05 < z < 6.43$  (Fan et al. 2003). While these indicate non-zero flux shortward of  $0.8 \mu\text{m}$ , there is essentially zero flux longwards of  $0.810 \mu\text{m}$ . Hence, there was significant HI in front of these quasars at  $z > 5.7$ , although the HI-fraction at  $z = 6$  was still very small (of order  $10^{-4}$  to  $10^{-5}$ ).

In most models, the conclusion of this reionization epoch is modeled by dwarf galaxies producing an increasing number of Population II stars at  $6 < z < 11$ . Most models are rather similar in their predictions of the cosmic SFH at  $6 < z < 8$ , in order to match the SDSS Gunn-Peterson troughs seen at  $z = 6$ . For example, the Cen (2003) models predict an increase in the cosmic SFH of a full factor of 10 over  $16 < z < 11$  and another factor of 10 increase over  $11 > z > 6$ . In other words, most of the Pop II stars that we see today were born in dwarf galaxies, but most were not born until about  $z = 8$  (consistent with the best oldest ages of Pop II measured today of 12.8 Gyr), and it was precisely the high-mass end of those Pop II stars that completed the epoch of reionization by  $z = 6$ . In WMAP cosmology, there was only 300 Myr between at  $6 < z < 8$  and another 170 Myr at  $8 < z < 10$ , so the stellar population that was formed in those galaxies, and whose O, B and A stars helped complete the reionization of the Universe by  $z = 6$ , is still visible as the low-mass Pop II stars seen today.

### 3.3 MEASUREMENTS REQUIRED FROM JWST

#### 3.3.1 Detection of First Luminous Objects at Very High Redshifts

Three basic observing programs can be used to identify the first light objects and derive their properties. The detection of first light from the direct detection of a population III star is not feasible even for JWST. However, JWST can detect super star clusters or dwarf galaxies made of population III stars as well as supernovae with population III progenitors.

##### 3.3.1.1 Identify a Sample of High-Z Galaxies

To identify a sample of high redshift galaxies, we must make an ultra-deep imaging survey using several broad-band filters (Figure 3-3). We will use the Lyman break technique to identify

objects at increasing redshifts up to  $z = 30$  or higher. For dwarf galaxies with  $10^6 M_{\odot}$  of zero-metallicity massive stars at  $15 < z < 30$ , the expected  $AB_{1400}$  magnitude is  $\sim 31$ . A similar value ( $AB = 31$  mag) is obtained by redshifting the brightest local super star clusters to  $z = 20$ . To enable this survey, JWST must have the sensitivity to reach  $AB = 31$  mag in a feasible (although long) exposure time, defined as  $< 5 \times 10^5$  s to reach  $SNR=10$  per filter. The expected number densities are of about 1 object  $\text{arcmin}^{-2}$ , thus a significant sample requires deep observations over several  $10 \text{ arcmin}^2$ . Another driver for a large area is the fact that in order to average over cosmic structures it would be desirable to cover a volume of at least 50 Mpc on the side. Such a volume is obtained at  $z = 15$  over a  $\Delta z = 3$  for an area of  $35 \text{ arcmin}^2$ . By focusing on volumes with  $\Delta z/z = 0.2$ , one obtains roughly the same comoving volume per unit redshift at all redshifts  $z > 5$ .

The sample will allow derivation of  $N(z)$  (see 3.1.3.1.1). The expected density of first light sources is much lower than the density of sources needed to enrich the metals in the IGM so that one should be able to see a drop in number counts. The intensity of the non-ionizing continuum can be calibrated to SFR to yield  $SFR(z)$  (see 3.1.3.1.2). The required observations are deep broad-band imaging in the near infrared (for  $z < 30$ ) and mid infrared (especially for  $z > 20$ ). One practical way to carry out this survey is to reach the depth of  $AB = 31$  mag for one field and integrate only to the depth of  $AB = 30$  mag for an additional three fields. For this program it will be essential to be able to integrate on the same target for at least  $10^6$  seconds.

### 3.3.1.2 In-Depth Study of First Light Sources

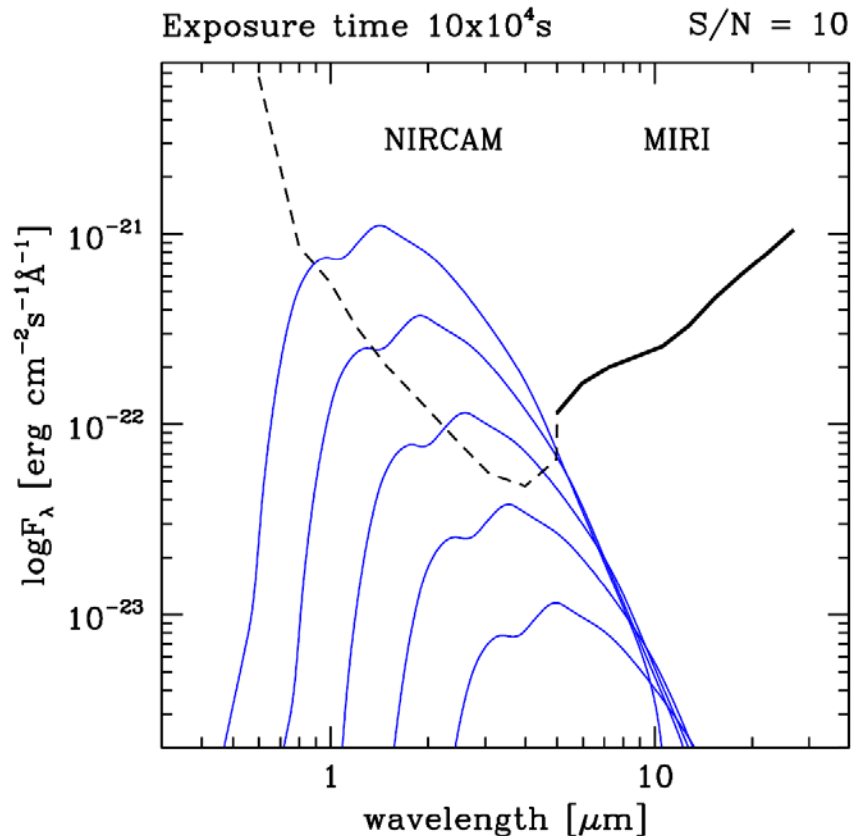
The brightest first light source candidates (or those amplified by intervening gravitational lenses) will be suitable for more detailed follow-ups. Near-infrared spectroscopy at  $R = 100$  will be needed to verify the photometric redshifts. This will only be possible at a limit much brighter than that of the deep imaging. Observations at rest-frame wavelengths longer than  $0.4 \mu\text{m}$  (i.e. at observer's wavelengths up to  $12.4 \mu\text{m}$ ) will establish the absence of an older generation of stars, confirming the nature of the sources as first generation objects (3.1.3.1.4). Spectroscopic follow up at  $R = 1000$  aimed at measuring the Balmer line intensities will provide star formation rates and estimates of the dust content. Metal lines can be used to derive metallicities and the mean metallicity as a function of redshift (see 3.1.3.1.3).

This program combined deep near-infrared spectroscopy and deep mid-infrared imaging, using integrations of  $10^6$  seconds. It is possible that in order to achieve the required S/N it will be necessary to exploit the gravitational lensing effect of a cluster of galaxies. Therefore, low scattered light and spectral purity are a requirement for the near-infrared spectrograph.

### 3.3.1.3 High-Z Supernova Searches

Individual Population III stars are too faint to be detected, but supernovae can be identified up to very high redshift, since they are expected to peak brighter than  $AB = 31$  mag at  $z = 14$  for the median of a population similar to local type II SNe (Figure 3-4). The brightest known SN would be visible up to  $z = 30$ , and it is possible that population III stars will produce bright type II supernovae. Detection of a number of SNe at high- $z$  will require multiple visits and will yield a SN-based star formation rate (see 3.1.3.1.2). The redshift of each supernova will in general be determined photometrically. The expected number of SNe is very uncertain and could be as low

as one per year per 100 arcmin<sup>2</sup>. Large areas need to be surveyed in order to obtain a significant sample. This program requires broad-band near-infrared imaging. Because of time dilation, the time between visits of the search field will need to be up to 6 months. While the first two visits produce only one search epoch, each successive visit produces another search epoch. Thus it is convenient to search the same field for an extended period of time. This could be accomplished by combining the SN search program with the ultra-deep observation. The existence of a continuous viewing zone including areas with low cirrus and high galactic latitude is a requirement.



**Figure 3-4. SEDs of Type II Supernovae at Maximum Light**

Note: From top to bottom the curves show SEDs of Type II supernovae with absolute rest magnitude  $M_B = -17.5$  at  $z = 5, 7, 10, 14$  and  $20$ . The JWST sensitivity requirements, shown in dashed black line in the near-infrared, and a solid black line in the mid-infrared, assumes a  $10^5$  second exposure, a signal to noise of 10 and a resolution of  $R \sim 4$ . At  $1 \mu\text{m}$ ,  $10^{22} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ \AA}^{-1} = 0.33 \text{ nJy}$ , or  $AB = 32.6 \text{ mag}$ . At  $10 \mu\text{m}$ , the same level is  $33 \text{ nJy}$ ,  $AB = 27.6 \text{ mag}$ .

### **3.3.2 Establishing when Reionization Occurred and Identifying the Reionizers**

Five programs are needed to firmly establish the epoch of reionization, to probe the possibility that a first reionization took place at very high redshifts, and to determine the properties of the ionizing sources. A starting sample of Lyman-break selected galaxies will be obtained from the

ultra-deep observations required to identify the first light sources. In order to extend this program to low redshift it is necessary to have sensitivity at wavelengths  $< 1 \mu\text{m}$ .

### 3.3.2.1 Lyman $\alpha$ Forest Diagnostics

We must obtain high signal-to-noise,  $R = 1000$ , near-infrared spectra of QSOs or bright galaxies identified in other surveys in order to determine the presence of a Gunn-Peterson trough or of a Lyman  $\alpha$  damping wing. High S/N is needed to discriminate between optical depths  $\tau$  of a few and  $\tau \gg 10$ . A damping wing should be present for a few million years, before the ionizing radiation is sufficient to create a large Strömngren sphere around each ionizing source. Failure to detect a damping wing does not necessarily imply that the Universe is ionized.  $R = 100$  spectra will be able to determine the presence of a Lyman  $\beta$  island (see 3.1.3.2.1). This is relevant if reionization occurs relatively abruptly. In this case, objects at redshifts between the redshift of reionization,  $z_{\text{reion}}$ , and  $z = \lambda_{\alpha}/\lambda_{\beta} * (1+z_{\text{reion}}) - 1$ , will show an island of normal, finite, forest absorption between the Lyman  $\alpha$  and the Lyman  $\beta$  forests. Here,  $\lambda_{\alpha}$  and  $\lambda_{\beta}$  are the rest frame wavelengths of Lyman  $\alpha$  and Lyman  $\beta$ , respectively.

If there are indeed two distinct reionization epochs, the (possibly partial) recombination following the first reionization may be detectable from the absorption signature left in the continuum of higher redshift objects.

Detection at  $z = 30$  of an object with the same Lyman  $\alpha$  intensity as the Hu et al. (2002) object at  $z = 6.56$  requires a sensitivity of  $2 \times 10^{-19} \text{ erg cm}^{-2} \text{ s}^{-1}$  at  $\lambda = 3.5 \mu\text{m}$ .

### 3.3.2.2 Transition in the Properties of Lyman $\alpha$ Sources

To detect a transition in the properties of Lyman  $\alpha$  sources at the epoch of reionization, we must select Lyman  $\alpha$  emitters at a variety of increasing redshifts by using the narrow band excess technique in near-infrared images. Given the high probability of interlopers, the sources would need to be confirmed either by detecting a second emission line with images at another wavelength, spectroscopically, or by using the Lyman-break technique. The aim is to detect rapid evolution of the Lyman  $\alpha$  luminosity function at one or two specific redshifts. Such evolution would be indicative of reionization. Line intensities will be fainter than  $6 \times 10^{-18} \text{ erg cm}^{-2} \text{ s}^{-1}$ . The need to verify the identification of a line as Lyman  $\alpha$  forces one to attempt the detection of a second line, e.g., H $\beta$ . This will in general be 30 times fainter than Lyman  $\alpha$ . An alternative method for finding Lyman  $\alpha$  emitters would be to search in the spectral domain with spectroscopy of blank areas.

By following the properties of Lyman  $\alpha$  emitters to the highest redshifts, we will be able to identify a period of partial recombination that would appear as a statistical brightening following by dimming of Lyman  $\alpha$  sources in the intermediate non fully-ionized period. This might be a more sensitive test than the equivalent one based on the absorption of the ionizing continuum photons, since for the latter a very small neutral fraction is already sufficient to produce very high opacity.

### 3.3.2.3 Transition in the Ratio between Lyman $\alpha$ and Balmer Lines

By measuring the hydrogen Balmer lines in addition to Lyman  $\alpha$ , it is possible to determine the amount that Lyman  $\alpha$  is suppressed due to either scattering or absorption. Any rapid evolution in this ratio as a function of redshift might indicate a change in the mean ionization state of the Universe (see 3.3.2.3). This measurement requires R = 1000 spectroscopy of Lyman  $\alpha$ , and H $\alpha$  or H $\beta$  and requires to measure line intensities down to  $2 \times 10^{-19}$  erg cm $^{-2}$  s $^{-1}$  at  $\lambda > 3$   $\mu$ m. This flux limit corresponds to a Lyman  $\alpha$  intensity  $6 \times 10^{-18}$  erg cm $^{-2}$  s $^{-1}$  (detected at 1  $\mu$ m) and a flux ratio of 30 between Lyman  $\alpha$  and H $\beta$ . This program requires spectroscopy in both the near and mid infrared.

### 3.3.2.4 Measuring the Ionizing Continuum

In order to measure the ionizing continuum of a class of sources, we need to measure their hydrogen and helium Balmer lines. Comparison between these lines provides an estimate for the steepness, or hardness of the ionizing continuum. The hydrogen Balmer line intensity provides the normalization. Taken together, the normalization and slope provide a measurement of the rate of production of ionizing photons for any given class of sources under the assumption that the escape fraction is known. The escape fraction can be measured from deep imaging observations or estimated from the line equivalent widths. This program requires near-infrared spectroscopy of very faint objects. The expected observed surface brightness of the sources responsible for reionization ranges between AB=27 and 29 mag arcmin $^{-2}$ , counted as the typical ionizing flux per unit area over which they are detected. The former applies to the case of metal-enriched reionization sources with dust and low escape fraction of ionizing UV, the latter applies to zero-metallicity ionizing sources with 100 per cent escape fraction for a much extended reionization period.

### 3.3.2.5 Determine the Source Nature

Identification of the nature of the ionizing sources requires a combination of diagnostics: line shapes, line widths, line ratios, shape of the continuum. We expect the intrinsic line widths of AGN-powered sources to be broader than those of sources ionized by stellar radiation. The line shapes may also help in distinguishing primordial HII regions from mini-AGNs. This program requires a combination of deep near-infrared R = 1000 spectroscopy and mid-infrared imaging.

### 3.3.2.6 Luminosity Function of Dwarf Galaxies

The Luminosity Function (LF) of dwarf galaxies over the redshift range  $6 < z < 10$  will reveal the completion of reionization and the birth of Pop II stars. High-mass Pop II stars completed the reionization at  $8 > z > 6$ , and low-mass Pop II stars are still visible today.

Dwarf galaxies at  $6 < z < 10$  are best found with the Lyman break or dropout technique. Finding objects in this redshift range requires high sensitivity at wavelengths  $0.8 < \lambda < 1.3$   $\mu$ m. JWST will measure any structural properties of these objects at wavelengths  $> 2.0$   $\mu$ m, where it must be diffraction limited, and we do not require JWST to be diffraction limited at shorter wavelengths. Still, it will be critical for the study of the conclusion of the epoch of reionization at  $6 < z < 10$ ,

that objects can be detected in the  $0.8 < \lambda < 1.3 \mu\text{m}$  range, and that basic properties such as colors and total fluxes can be measured with sufficient S/N.

### 3.4 SUMMARY

Table 3-1 lists the observational programs for the End of the Dark Ages theme. In addition to the high sensitivity required to detect extremely faint sources, there is also a strong requirement on angular resolution, to determine whether the distant objects are points like individual stars and supernovae, or extended collections of stars. We require diffraction-limited imaging resolution at all wavelengths longer than  $2.0 \mu\text{m}$ . Table 3-2 lists the required instrument and observatory capabilities for each observational program.

**Table 3-1. JWST Measurements for the End of the Dark Ages Theme**

Observation	Primary Instrument	Magnitude or flux	Target Density
Ultra-deep survey (3.3.1.1), SN (3.3.1.3)	NIRCam	AB = 31 mag	1 arcmin <sup>-2</sup>
In-depth study (3.3.1.2)	NIRSpec	AB = 28 mag, R=100	1 arcmin <sup>-2</sup>
	MIRI	AB = 28 mag	1 arcmin <sup>-2</sup>
Ly $\alpha$ forest diagnostics (3.3.2.1)	NIRSpec	$2 \times 10^{-19} \text{ erg cm}^{-2} \text{ s}^{-1}$ , R=1000	Individual
Transition in Ly $\alpha$ properties (3.3.2.2)	FGS-TF	$2 \times 10^{-19} \text{ erg cm}^{-2} \text{ s}^{-1}$ , R=100	1 arcmin <sup>-2</sup>
Transition in Ly $\alpha$ /Balmer (3.3.2.3)	NIRSpec	$2 \times 10^{-19} \text{ erg cm}^{-2} \text{ s}^{-1}$ , R=1000	1 arcmin <sup>-2</sup>
Measure ionizing continuum (3.3.2.4)	NIRSpec	$2 \times 10^{-19} \text{ erg cm}^{-2} \text{ s}^{-1}$ , R=1000	1 arcmin <sup>-2</sup>
Ionization source nature (3.3.2.5)	NIRSpec		
	MIRI		
LF of dwarf galaxies (3.3.2.6)	NIRCam		

**Table 3-2. Required Capabilities for the End of the Dark Ages Theme**

	First Light			Reionization					
	UDF	Props	SN	Forest	Ly $\alpha$ LF	Balmer	UV	Source	Dwarf
	3.3.1.1	3.3.1.2	3.3.1.3	3.3.2.1	3.3.2.2	3.3.2.3	3.3.2.4	3.3.2.5	3.3.2.6
<b>NIRCam</b>									
Broad-band	X		X					X	X
Sensitivity	X		X						X
FOV	X		X					X	X
PSF	X		X						X
Dynamic range									
Coronagraphy									
<b>NIRSpec</b>									
Multi-obj FOV	X	X			X	X	X	X	
R = 100	X	X			X				
R = 1000	X	X		X	X	X	X	X	
R = 3000									
Sensitivity		X		X	X	X	X	X	
High Contrast									
<b>MIRI</b>									
Broad-band		X						X	
Sensitivity		X				X	X		
FOV		X						X	
PSF		X						X	
R = 3000 IFU						X	X		
Coronagraphy									
<b>FGS-TF</b>									
Short $\lambda$ TF					X				
Long $\lambda$ TF					X				
Sensitivity					X				
FOV					X				
Coronagraphy									
<b>Observatory</b>									
PSF $\lambda < 1 \mu\text{m}$	X				X		X		X
Stable PSF	X		X						
Stable Image	X		X						
CVZ			X						
Moving Targets									

## **4.0 THE ASSEMBLY OF GALAXIES**

The key objective of The Assembly of Galaxies theme is to determine how galaxies and the dark matter, gas, stars, metals, morphological structures, and active nuclei within them evolved from the epoch of reionization to the present day.

### **4.1 IMPACT ON SCIENCE**

Galaxies are basic building blocks of the Universe. Material within galaxies undergo the vast cycle of stellar birth, life and death that results in the production of the heavy elements, the formation of planets and the emergence of life. Most of the astrophysical complexity of the Universe is manifested within galaxies and the formation of galaxies thus represents a key link between this complexity, and the relative simplicity of the surrounding cosmological Universe. On the one hand, the most basic properties of galaxies reflect the distribution of dark matter in the Universe, which result from very simple quantum processes operating during the earliest moments of the Big Bang. On the other hand, the subsequent complex astrophysical behavior of the baryonic material within these dark matter gravitational potential wells produces the beautiful appearance and diverse properties of present-day galaxies. Therefore, understanding the processes that formed the present-day population of galaxies is central to cosmology, to astrophysics and to our understanding of the emergence of life in the Universe.

$\Lambda$ CDM cosmology provides a conceptual framework for understanding the formation of galaxies through the hierarchical assembly of progressively more massive objects, but many of the most basic questions about this process remain unanswered due to the difficulty of observing small and faint objects at high redshifts. Even the origins of the most fundamental scaling relations for galaxies are not understood and the CDM cosmogony on galactic scales has not yet been tested. On the theoretical side, the superficially successful “semi-analytic” models for galaxy formation and evolution are based on many ad-hoc assumptions, while numerical gravito-hydrodynamic simulations do not yet have the resolution and dynamic range needed to simultaneously cover astrophysical feed-back mechanisms on the scales of individual star-formation events up to the growth of a galaxy in its cosmological environment.

It is clear that the formation and early evolution of galaxies is a complex and multifaceted problem that no single observation or theory will solve. Essential elements of our eventual understanding will almost certainly include the following:

- The fundamental physics of the very early Universe, including the origin of density fluctuations and the nature of the dark matter and dark energy;
- The hierarchical assembly of matter through gravitational instability;
- The formation of stars under a wide range of conditions, including some quite different from those encountered today;
- The origin and growth of black holes at the centers of galaxies;
- The feedback of energy and radiation produced by the first galaxies or pre-galactic objects on the surrounding material, including the re-ionization of the intergalactic medium;

- The exchange of material between galaxies and the surrounding reservoir of baryons. Coupled with these physical processes, a host of observational issues must be understood, including the effects of dust obscuration and the inevitable observational selection effects in wavelength and surface brightness.

Progress requires new observational data, both to characterize the galaxy population at different epochs, and to understand the astrophysics of key processes that are occurring in the Universe at early times. JWST must address the most pressing observational questions.

To gain an understanding of the extremely distant Universe requires a systematic and comprehensive approach. Objects must be detected and identified (i.e. recognized as being at high redshift), and then characterized in terms of their physical properties and of the physical processes occurring in and around them. They must be placed in the context of a global understanding of the other objects and phenomena going on at the same epochs. It is also essential to understand which objects at one epoch evolve into which objects at a subsequent epoch and to understand the relationship at all times between the visible baryonic material and the underlying dark matter.

## 4.2 PREVIOUS INVESTIGATIONS

During the mid-1990s, the simultaneous introduction of efficient multi-object spectrographs on large telescopes, the commissioning of the first 8 to 10m telescopes and the repair of the HST led to a dramatic advance in our direct observational knowledge of the galaxy population at earlier epochs.

At  $z \sim 1$  Universe at optical and near-infrared wavelengths appears roughly similar to that seen today. There is a full range of Hubble types including spirals and ellipticals (e.g. Driver et al. 1995, Schade et al. 1995, Abraham et al. 1996, Brinchmann et al. 1998), a well-developed luminosity function of quiescent red galaxies (Lilly et al. 1995), approximately the same number density of large spiral disks, “normal” Tully-Fisher rotation curves in these disks (Vogt et al. 1997, 1998), and so on. The metallicities of the star-forming gas are close to solar. There are however some clear evolutionary effects in the sense that most of the signatures of vigorous star-formation activity (e.g. blue colors, strong emission lines, irregular morphologies, lower metallicities) are found in more luminous galaxies at  $z \sim 1$  than locally, the so-called "down-sizing" effect (Cowie et al. 1995). In addition the overall luminosity density in the ultraviolet and emission lines is about a factor of five higher at  $z \sim 1$  than it is locally.

Extending beyond  $z \sim 1$ , the known galaxies at  $z \sim 3$  are generally blue with compact or irregular morphologies. Most of these galaxies have been selected in the ultraviolet and it is not yet clear whether there is a real absence of well-developed spiral or quiescent elliptical galaxies at this redshift or when such galaxies first appear (see e.g. Giavalisco et al. 1996, Abraham et al. 1999, Zepf 1997, Dickinson 2000, Frenx et al. 2003). The first deep K-selected samples appear to show large numbers of red galaxies at redshifts approaching  $z \sim 2$  (McCarthy et al. 2003, Abraham et al. 2003), although the stellar masses are sufficiently uncertain that it is not yet clear what fraction of the  $z \sim 1$  population these represent. The handful of ultraviolet-selected galaxies studied in detail at  $z \sim 3$  show evidence for significantly sub-solar metallicities ( $Z \sim 0.3 Z_{\odot}$ ) and

for galactic winds of several hundred  $\text{km s}^{-1}$  indicating substantial ejection of enriched material into the intergalactic medium.

Beyond  $z \sim 3$ , our picture of the galaxy population becomes very fragmentary as we approach the epoch at which reionization appears to have been completed ( $z \sim 7$ ). Small samples of galaxies are known, generally found through their strong Lyman  $\alpha$  emission, (Hu et al. 2002, 2004, Rhoads et al. 2003,) or by extensions of the Lyman break "drop-out" technique to longer wavelengths (Dickinson et al 2004, Bouwens et al. 2003, Yan et al. 2003), but there has been little systematic or detailed study of these exceedingly faint objects.

COBE showed in 1995 that the extragalactic background light has equal energy in the far-infrared as in the optical and near-infrared and that the absorption and re-radiation of light by dust has played a major role in shaping the appearance of the Universe (Puget et al. 1996, Fixsen et al. 1998). Much less is currently known about the sources responsible for the far-IR/sub-mm background than the optical sources described above. At  $850 \mu\text{m}$ , about 50% of the background has been resolved (e.g. Hughes et al. 1998, Barger et al. 1998, Eales et al. 1999, 2000), and these sources are extremely luminous heavily dust-enshrouded galaxies with luminosities greater than several  $10^{12} L_{\odot}$ , comparable or greater to the local ultra-luminous infrared galaxies (ULIRGs) discovered by IRAS. Although little is known reliably about their redshifts, it is clear that these are about 100 times more common at high redshift ( $z > 1$ ) than they are locally. At  $15 \mu\text{m}$ , deep counts are available from ISOCAM surveys and Chary and Elbaz (2001) show that the rapid evolution required to account for these must flatten at  $z \geq 1$ , so as not to overproduce the background seen beyond  $100 \mu\text{m}$  (see also Lagache, Dole & Puget 2002). This is broadly similar to the behavior seen in the ultraviolet, with a possibly steeper rise at low redshifts. It is not yet known definitively whether the energy source in these obscured objects is a massive burst of star-formation or accretion onto a black hole in an active galactic nucleus. It is tempting to associate these objects with major mergers of young galaxies, since the low redshift ULIRGs appear to be triggered by such events.

From the above it is clear that the redshift range  $1 < z < 7$  is the time when the galaxy population acquired most of its present-day characteristics, when a large fraction of the stars we see today were formed, and when a large fraction of the metals were produced. Accordingly, this is the period when the most important astrophysical processes in galaxy formation and evolution occurred.

JWST must observe galaxies at  $1 < z < 7$  to answer these fundamental questions.

- Where were stars in the Hubble Sequence galaxies formed, when did luminous quiescent galaxies appear, and how does this depend on the environment?
- Where and when were the heavy elements produced, and to what extent do galaxies exchange material with the intergalactic medium?
- When and how were the global scaling relations for galaxies established?
- Do luminous galaxies form through the hierarchical assembly of dark matter halos?
- What are the redshifts and power sources of the high redshift Ultra Luminous Infrared Galaxies?

- What is the relation between the evolution of galaxies and the growth and development of their nuclear black holes?

In answering these questions, observations of the interstellar medium within the galaxies by ALMA will be complementary to the contributions of JWST.

### 4.3 MEASUREMENTS REQUIRED FROM JWST

#### 4.3.1 Where were Stars in the Hubble Sequence Galaxies Formed, When did Luminous Quiescent Galaxies Appear, and How does this Process Depend on the Environment?

To answer this question we require observations of the morphologies, stellar populations, and star-formation rates in a very large sample of galaxies observed in deep imaging and spectroscopic surveys. This investigation has substantial overlap with several of the topics below, especially with the chemical enrichment of galaxies (4.3.2), the measurement of masses (4.3.3), and the nature of the highly obscured luminous galaxies (4.3.4).

To solve this question, we must characterize the star-formation rates in individual galaxies, ideally as a function of their mass and environment. We must also determine when the long-lived stars in a typical galaxy were formed, whether in situ or in smaller galaxies that subsequently merged together to form a large galaxy. Direct characterization of the merging rate of galaxies will provide another angle on this question.

The emergence of quiescent red galaxies, which have completed their major episodes of star-formation, at least for the time being, will tell us why star-formation ceases in some galaxies. The importance of chaotic star-formation in star-bursts, as compared with the steady-state star-formation in stable galactic disks, will reveal the modes of star-formation which dominate different phases of galactic evolution and develop the morphological components in the galaxies.

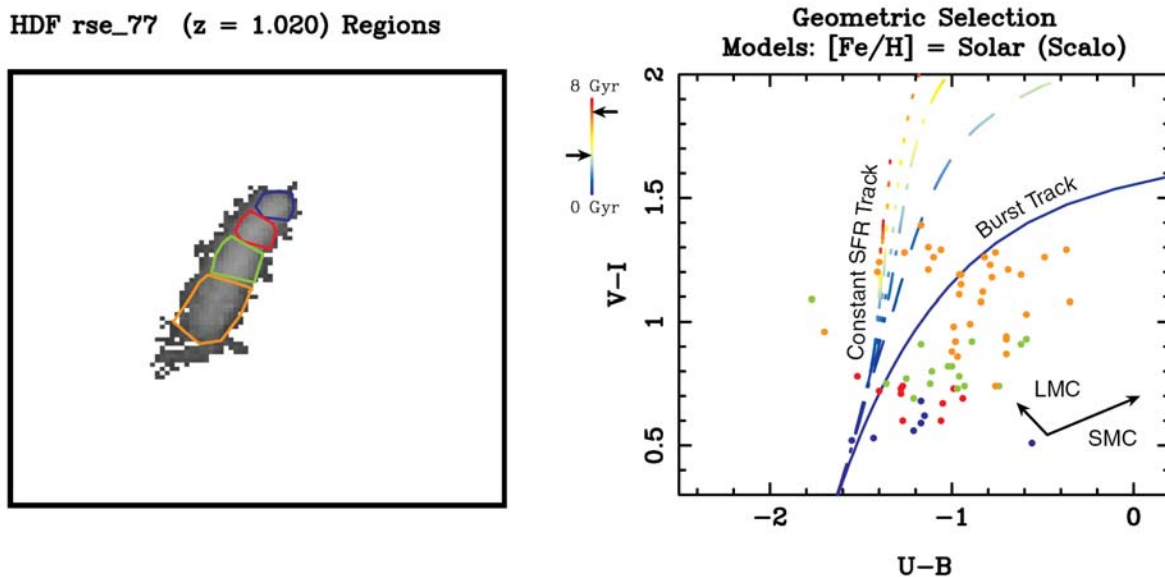
Quantities such as the disk size function, as well as color gradients within galactic disks at different redshifts will show directly how galactic disks grew, while the merger rate of disk galaxies will reveal the rate at which stars, originally formed in disks, are redistributed into the spheroids.

Except in objects with very high levels of dust extinction, the star-formation rate of massive stars in a galaxy can best be estimated from measurements of the H $\alpha$  emission line, complemented by those of other emission lines, the ultraviolet continuum and the bolometric luminosity at longer wavelengths. JWST should have the capability to spectroscopically measure the H $\alpha$  emission ( $5 \times 10^{-22}$  Wm $^{-2}$ ) that would be produced by a star-formation rate of only  $1 M_{\odot}/\text{yr}$  at  $z \sim 5$  (Kennicutt 1999).

The existence of older stellar populations is best revealed by continuum imaging at rest wavelengths  $\lambda > 0.5 \mu\text{m}$  and if possible  $\lambda > 1 \mu\text{m}$ . Based on the Local Group and Milky Way Galaxy, the deepest near-infrared imaging should be able to detect the Small Magellanic Cloud (with  $M_V = -16.2$  mag) if placed at  $z \sim 5$ , where it would be unresolved and have  $L_{AB} \sim 30.3$  mag. JWST should make this measurement at  $5\sigma$  in 30 ksec.

Redshifts for essentially all galaxies on an image above a faint threshold (typically  $S/N \geq 10$ ) may be estimated from multi-band photometric redshifts with a typical accuracy  $\delta_z/(1+z) < 0.1$  and only a few percent falling far from the actual redshift. Confirmation of these will be possible using either  $R \sim 100$  or  $R \sim 1000$  near-infrared spectra, as required.

Broad-band colors on their own can reveal information on the ages and reddening of individual components within a galaxy (Figure 4-1; Abraham et al. 1999), possibly revealing the physical causes of episodes of star-formation, such as sequential star-bursts.



**Figure 4-1. Analysis of HST Images**

NOTE: Galaxies in deep HST images are separated on the basis of color into regions with different star-formation histories (Abraham et al. 1999).

Relative to work on the HDF, GOODS, and the UDF, deep images from the JWST and HST together, will allow a full representation of the spectral energy distribution of each galaxy, and of the distinct morphological components within it, from the Lyman limit at  $0.0912 \mu\text{m}$  out to a solid anchor in any older population in the rest-frame  $0.6 \mu\text{m}$  region, even for galaxies at redshifts as high as  $z \sim 7$ . Only with JWST can the relationship between old and young stellar populations be understood, and only with JWST can a full characterization of the star-formation process at high redshift be made.

The properties of galaxies today depend on their environments, and there is strong observational evidence for a morphology-density relation, showing a clear difference between stellar populations in the field and in rich clusters (e.g. Dressler 1984). It is not understood how these differences came about, and if they were established early in the evolutionary history of galaxies, perhaps in groups prior to the establishment of the full-blown clusters. Carrying out the above studies in a range of environments would show when and why these differences arose.

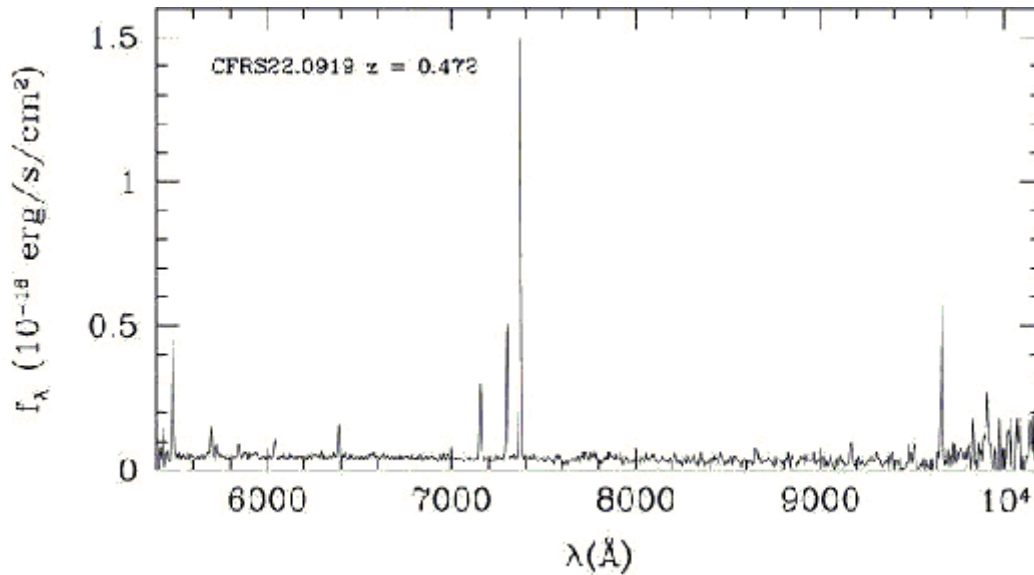
#### **4.3.2 Where and When are the Heavy Elements Produced and to What Extent do Galaxies Exchange Material with the Intergalactic Medium?**

The metallicity of the Universe and of the objects in it provides a fundamental metric reflecting the development of structure and complexity on galactic scales. Metallicity is observable and “long-lived” in the sense that heavy atomic nuclei, once produced, are not readily destroyed. The production of heavy elements is also one of only two cosmologically significant producers of luminosity in the Universe, along with gravitational accretion energy.

For many years, the metallicities of gas at high redshifts have been studied through the analysis of absorption line systems seen in quasar spectra. The lines of sight to quasars probe random regions of the Universe. The study of the metallicities of material in known galaxies at high redshift is at a much earlier stage of development. This is more relevant for models of the chemical evolution of galaxies, and for the use of metallicity estimates to constrain the present-day descendants of high redshift galaxies. The emission line gas in star-forming regions is relevant for the planetary and astro-biological studies, since it is representative of the material out of which the stars and planets are made. The metallicities of star-forming gas, especially of the [O/H] abundance, can be measured using diagnostics such as Pagel's R23 index, which is based on strong emission lines such as [OII]  $\lambda$ 3727, H $\beta$ , [OIII]  $\lambda\lambda$  4959, 5007, H $\alpha$ , [NII]  $\lambda$ 6583, [SII]  $\lambda\lambda$  6717, 6731 (Figure 4-2). Such measurements require  $R \sim 1000$  to separate H $\alpha$  and [NII].

JWST must be able to measure the H $\alpha$ , H $\beta$  and [OII]3727 and [OIII]5007 emission lines from compact low-extinction galaxies at  $z \sim 5$  that are forming stars at the rate of  $3 M_{\odot}/\text{yr}$ . This star-formation rate is comparable to that of the Milky Way today, and requires a line sensitivity of  $5 \times 10^{-22} \text{ Wm}^{-2}$  in the 2-4  $\mu\text{m}$  range at  $10\sigma$  in less than  $10^5$  s. In order to assemble sufficient samples for statistical determinations, JWST must be able to make these measurements with high multiplexing gain.

Extensive studies of the local Universe have recently been extended to  $z \sim 1$  using ground-based facilities. However, beyond  $z \sim 0.5$  these observations become progressively more difficult from the ground as the emission lines are redshifted into the infrared. Only a handful of galaxies have been observed in any fashion at redshift much greater than  $z \sim 1$ . With the JWST, all of these lines will be observable in the near infrared over the redshift range  $1.7 < z < 5.0$  enabling metallicities of individual star-forming galaxies to be measured to a precision of about 0.2 dex.



**Figure 4-2. Spectrum of a galaxy at  $z \sim 0.5$**

NOTE: Galaxy spectrum taken as part of the Canada-France Redshift Survey shows O[II] to H- $\alpha$ , the lines that make up the R23 index (*REFERENCE TBD*).

Metallicities must be determined over this range of redshifts to tie in with the observations of "first light" and the very first enrichment (see Section 3) and to chart the development of metallicity through the epoch when most of the stars and metals were made.

Measurement of the gas metallicity in a very large number of faint galaxies (i.e. the metallicity distribution function), and comparison with the metallicities in neutral absorption line gas will allow the JWST to address the origin of the enriched intergalactic medium, the enrichment histories of different types of galaxies, and the degree to which merging or accretion of galaxies alters the metallicity of growing galaxies.

#### **4.3.3 When and How are the Global Scaling Relations for Galaxies Established?**

Despite the variety of galaxy properties observed today, galaxies obey a number of remarkably tight scaling relations between basic properties of luminosity, size, and kinematics and metal enrichment. These include the Tully-Fisher relation for disk galaxies and the "fundamental plane", and projections thereof, for spheroids. Most recently, a surprising relationship between the mass of the central black hole and the properties of the surrounding spheroid (e.g. the velocity dispersion) has been established. It is not known how or when these were established and whether they represent an asymptotic, late-epoch, state or whether they are obeyed at essentially all epochs (once allowance is made for the simple evolution of the stellar population).

Simulations of galaxy formation have managed to reproduce the slope, but not the normalization of the dynamical relations. The compatibility of scaling relations based on color or metallicity with models in which most stars are formed outside of their eventual parent galaxies is not completely clear. Determination of the nature of these scaling laws at

$1 < z < 7$  and of the scatter about them would likely reveal what physical process was responsible.

Construction of these relations requires deep imaging for structural parameters plus high-resolution spectroscopy for kinematical data. Disk rotation curves at high spatial resolution can be measured in  $H\alpha$  in the observed near-infrared. This is much superior to the lines that are accessible in the observed optical band, such as Lyman  $\alpha$ , which are strongly affected by dust and radiative transfer effects in the interstellar medium, and in outflow regions, making them essentially useless for dynamical studies.

An important stellar kinematical diagnostic will come from the CO band heads at  $2.2 \mu\text{m}$ . These are particularly useful as they appear very quickly in young stellar populations, are largely independent of metallicity, and are little affected by dust extinction. For these reasons they may well be the kinematic diagnostic of choice at high redshift. JWST must have sufficient mid-infrared sensitivity to measure the stellar velocity dispersions using the CO bandheads at  $z \sim 3$  in at least the brighter of the Lyman break galaxies. Scaling from an optical magnitude of  $R_{AB} = 24.5$  mag, and the spectral energy distribution of a present-day Scd galaxy, this requires a continuum sensitivity at  $9 \mu\text{m}$  of  $AB = 21.3$  mag at  $10\sigma$  per resolution element in less than  $10^5$  sec.

#### **4.3.4 Do Luminous Galaxies Form through the Hierarchical Assembly of Dark Matter Halos?**

The fundamental feature of the standard CDM cosmogony is that the mass function of bound structures develops with time, as smaller objects are assembled hierarchically into larger ones, leading to an increase in the characteristic mass  $M^*$  in the Press-Schechter mass function. JWST images and spectra will study the evolution and organization of baryons in galaxies at high redshift, but will not reveal the underlying structures of non-luminous matter which make up the gravitationally bound dark matter halos. It is the development of these halos which is the fundamental test of the CDM theory of galaxy formation.

The dark matter mass function of bound objects at very high redshifts can be uniquely measured in two ways with the JWST. First, the dynamics of groups of galaxies or sub-galactic fragments can be used to determine the typical masses of halos (Zaritsky & White 1994). The Lyman  $\alpha$  line is affected by galactic winds making it essentially useless for determining the kinematics of individual galaxies or groups, and such measurements require observations of emission lines at longer wavelengths, such as [OII] 3727, [OIII] 5007 and  $H\alpha$ . These are very difficult from the ground at all redshifts  $z \gg 1$ .

Second, JWST will measure halo masses through the gravitational bending of light. Using this weak-lensing method, ground-based programs have measured the mass within 200-500 kpc at redshifts of  $z \sim 0.1$  (McKay et al. 2002) and  $z \sim 1$  (Wilson et al. 2001). Using the superior resolution of HST, these measurements are likely to be extended into 30-50 kpc for galaxies at  $z \sim 1$ . While there are some hints of variable halo structures for galaxies of different luminosity and total halo mass, the radial penetration of these surveys and the ability to compare galaxies of

different morphologies are limited by statistics. We can expect that HST will establish the statistical mass functions for spiral and elliptical galaxies at  $z \sim 1$ .

With JWST, we must extend the equivalent measurements of galaxies to  $z \sim 2.5$  and thus determine the development of the dark matter halos during the peak growth of galaxies and star formation. JWST will require near infrared imaging with high spatial resolution and sensitivity achieve this greater depth. Background galaxies with a size comparable to the resolution of JWST must be measured to a  $S/N \sim 20$ .

The same near-infrared sensitivity and resolution will also make JWST superior to those of ground-based facilities and HST for the study of dark matter structures on larger scales, e.g. 1 - 10 arcminutes or 2 - 20 Mpc (co-moving) at  $z \sim 3$ . These volumes measure the clustering of dark matter on cluster or even supercluster scales, and would extend the study of the mass function into the linear regime. The goal of these observations would be to verify the growth of structure between  $z \sim 1000$  (the CMB large scale structure) and  $z \sim 2.5$ , i.e. during the period that dark matter dominated the cosmological expansion of the Universe prior to the emergence of the dark energy.

#### **4.3.5 What are the Redshifts and Power Sources of the High Redshift Ultra Luminous Infrared Galaxies?**

The optical identification of high-redshift ULIRGs, found at sub-mm wavelengths, is well below the limits of ground-based 8-10 m telescopes. At present none of the deepest field samples are securely identified at greater than the 50% level. Intensive efforts with ground-based telescopes will improve this before the JWST launch, but it is almost certain that there will still be many unidentified sub-mm sources when JWST flies. Analogs of known  $z \sim 2$  ULIRGs, if they exist at  $z > 5$ , will have remained unidentified from the ground until JWST flies even though they may well already be present in today's sub-mm samples. Deep imaging with JWST at  $\lambda > 2 \mu\text{m}$  is required to identify these sub-mm sources.

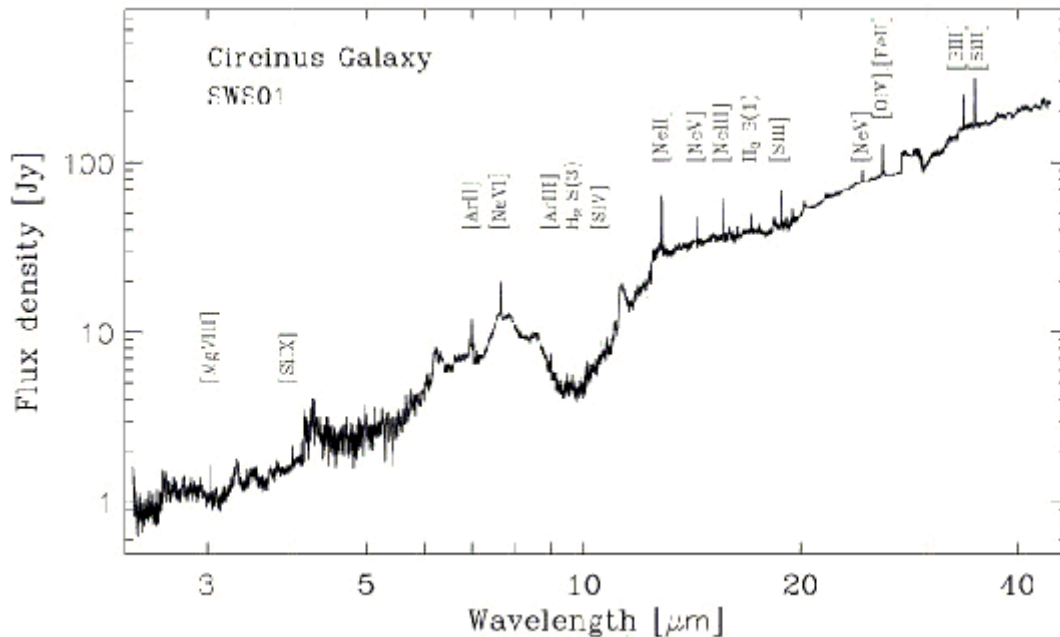
Mid-infrared imaging will test whether mergers are the cause of the energy injection in high redshift ULIRGs. This will penetrate the dust obscuration that is known to be present in these obscured galaxies and will sample the oldest stellar populations in these objects, rather than just knots of recent star-formation. At the median redshifts of  $z \sim 2$  to 3 expected for many of the sub-mm selected ULIRGs, JWST images at  $4 \mu\text{m}$  will sample the stellar populations in these galaxies at wavelengths longwards of  $1 \mu\text{m}$  in the rest-frame, allowing the best possible identification of mergers.

Near-infrared spectroscopy with JWST must have the capability to measure redshifts for identifications that cannot be secured from ground-based spectroscopy. Most ULIRGs at  $z \sim 4$  are too faint to be observed from the ground at  $\lambda < 2 \mu\text{m}$ . The  $H\alpha$  line, which would be expected to be the strongest line in these highly obscured but vigorously star-forming galaxies, leaves the ground-based K-band window at  $z > 2.6$ , but will be readily observable with JWST near-infrared spectroscopy to  $z \sim 6.5$ . Beyond  $z > 2.6$ , it may be possible to observe lines shortward of  $H\alpha$  from the ground (e.g. [OIII] 4939,5007,  $H\beta$  and [OII] 3727), but these will be extremely faint in these highly reddened objects and even these will have left the K-band by  $z \sim 5$ . Use of  $R = 1000$

will yield kinematic information on the merging system, and will separate H $\alpha$  and [NII] allowing some estimate of metallicity to be made.

Finally, at the long wavelength end of the mid-infrared, high resolution spectra will allow detection of narrow emission lines such as [NeVI] 7.66  $\mu\text{m}$  to  $z \sim 2.5$ , while spectra at lower resolution will allow measurement of the equivalent width of the 7.7  $\mu\text{m}$  polycyclic aromatic hydrocarbon (PAH) feature at redshifts as high as  $z \sim 2.5$  and of the 3.3  $\mu\text{m}$  feature to redshifts of  $z \sim 6$ . These emission line and PAH features are good diagnostic of the energy sources in the interior of the system. Star-bursts have strong PAH features, while AGNs have much weaker features because the PAHs are themselves destroyed and the hot dust continuum is stronger. [NeVI] is also much stronger in AGN powered systems. JWST must allow application of the same diagnostic as has proved most useful in the low redshift systems.

JWST must be able to measure the [NeVI] line in an ultra-luminous obscured galaxy with the bolometric luminosity of Arp 220,  $1.3 \times 10^{12} L_{\odot}$ , at  $z \sim 2$ , assuming a line/bolometric luminosity ratio as in the Circinus galaxy, at  $10\sigma$  in less than  $10^5$  s (see Figure 4-3).



**Figure 4-3. ISO Circinus Spectrum**

NOTE: Mid-infrared spectrum of the Circinus galaxy taken with ISO shows an abundance of emission lines useful for diagnosing the energy sources which power ULIRGs.

**4.3.6 What is the Relation Between the Evolution of Galaxies and the Growth and Development of Black Holes in Their Nuclei?**

One of the most surprising discoveries in galactic astrophysics of the last ten years has been that the masses of central black holes and of the bulge stellar population in present-day galaxies are tightly correlated. These estimates have been extended using proxy indicators to redshifts  $z \sim 2$

in QSOs, and indicate that the correlation still holds. Furthermore, the host galaxies of QSOs at redshifts  $z > 2$  appear to be in very high states of star-formation, while the peak in the quasar number density at  $z \sim 2$  suggests that the formation of the central black hole is contemporaneous with the production of the bulk of the stellar population. However, the existence of some bright QSOs at redshifts above 6, with spectra that differ little from those with the lowest redshift, suggests that some massive black holes and their associated stellar populations have formed early in the history of the Universe. The close connection between central black holes and spheroidal populations must be intimately connected with galaxy formation and evolution, and with the events that trigger and fuel AGN over cosmic time.

Black hole masses have been measured by echo or reverberation mapping, maser kinematics, nuclear gas dynamics, nuclear star dynamics, and BLR emission-line widths. They show a good level of agreement, and are probably correct to within a factor of two or three. Many of them will be applicable at high redshifts with the JWST. Bulge stellar populations are characterized by the bulge luminosity profiles, velocity dispersion, and overall flux, with appropriate M/L ratios according to the stellar populations.

There are many questions that still remain about the formation and evolution of super-massive black holes. We don't know if the seed black holes are primordial, if they form through the high end of the population III mass function, or if they form over a wide range of redshifts. We don't know if their evolution traces the hierarchical growth of structure, or through merging within an initial stellar population. We don't know the role of angular momentum, and the role of central engine accretion mechanisms in their growth. Finally, we don't know the redshift dependence of black hole mass growth.

While many of these questions will be addressed using the same type of observations outlined above for non-AGN galaxies, JWST must also observe a range of active galaxy types and luminosities.

#### **4.4 SUMMARY**

Table 4-1 lists the sensitivities needed for JWST observations of the assembly of galaxies. Spatial structure is a fundamental aspect of galaxies, and the observing program depends critically on resolving that structure and classifying galaxies according to shape. As not all the galaxies of interest are at high redshift, the telescope performance at the shortest JWST wavelengths is extremely important. For example, at a redshift of 6, the limit of current observations, the Lyman alpha line at  $0.1216 \mu\text{m}$  is redshifted to  $0.85 \mu\text{m}$ , at the shortest limits of the JWST wavelength range. This theme requires that the image quality be excellent (high Strehl, high encircled energy, very circular), and very stable (to enable long exposures). Table 4-2 summarizes the modes and instrument capabilities used by each program.

**Table 4-1. JWST Measurements for the Assembly of Galaxies Theme**

<b>Topic</b>	<b>Primary Instrument</b>	<b>Key Observation</b>	<b>Magnitude or Flux</b>	<b>Target Density</b>
Faint galaxy identification and morphology	NIRCam	Detect the SMC at $z = 5$ in rest-frame V	$L_{AB} = 30.3$ mag	100 arcmin <sup>-2</sup>
Metallicity determination	NIRSpec	Determine $R_{23}$ from emission line ratios for galaxy with $SFR = 3 M_{\odot}/yr$ at $z = 5$	$5 \times 10^{-19}$ erg s <sup>-1</sup> cm <sup>-2</sup>	100 arcmin <sup>-2</sup>
Scaling relations	MIRI spectroscopy (short wavelength)	Measure stellar velocity dispersion for $R_{AB}=24.5$ Lyman Break galaxy at $z = 3$	AB (9 $\mu$ m) = 21.3 mag	1 arcmin <sup>-2</sup>
Obscured galaxies	MIRI spectroscopy (long wavelength)	Measure [NeVI] in ULIRG with Arp220 $L_{bol}$ assuming Circinus spectrum	line $1.4 \times 10^{-19}$ Wm <sup>-2</sup>	Individual

**Table 4-2. Required Capabilities for the Assembly of Galaxies Theme**

	Hubble Sequence	Heavy Elements	Scaling Relations	Hierarchical Assembly	ULIRGs	AGN
	4.3.1	4.3.2	4.3.3	4.3.4	4.3.5	4.3.6
<b>NIRCam</b>						
Broad-band	X		X	X	X	X
Sensitivity	X		X	X	X	X
FOV	X		X	X	X	X
PSF	X		X	X	X	X
Dynamic range						
Coronagraphy						
<b>NIRSpec</b>						
Multi-obj FOV	X	X	X	X	X	X
R = 100	X			X	X	X
R = 1000	X	X	X	X	X	X
R = 3000						
Sensitivity	X	X	X	X	X	X
High Contrast						
<b>MIRI</b>						
Broad-band	X				X	X
Sensitivity	X	X	X		X	X
FOV	X	X	X		X	X
PSF	X				X	X
R = 3000 IFU	X	X	X		X	X
Coronagraphy						
<b>FGS-TF</b>						
Short $\lambda$ TF	X				X	
Long $\lambda$ TF	X				X	
Sensitivity	X				X	
FOV	X				X	
Coronagraphy						
<b>Observatory</b>						
PSF $\lambda < 1 \mu\text{m}$	X		X	X		X
Stable PSF	X		X	X	X	X
Stable Image						
CVZ						
Moving Targets						

## **5.0 THE BIRTH OF STARS AND PROTOPLANETARY SYSTEMS**

The key objective of The Birth of Stars and Protoplanetary Systems theme is to unravel the birth and early evolution of stars, from infall on to dust-enshrouded protostars, to the genesis of planetary systems.

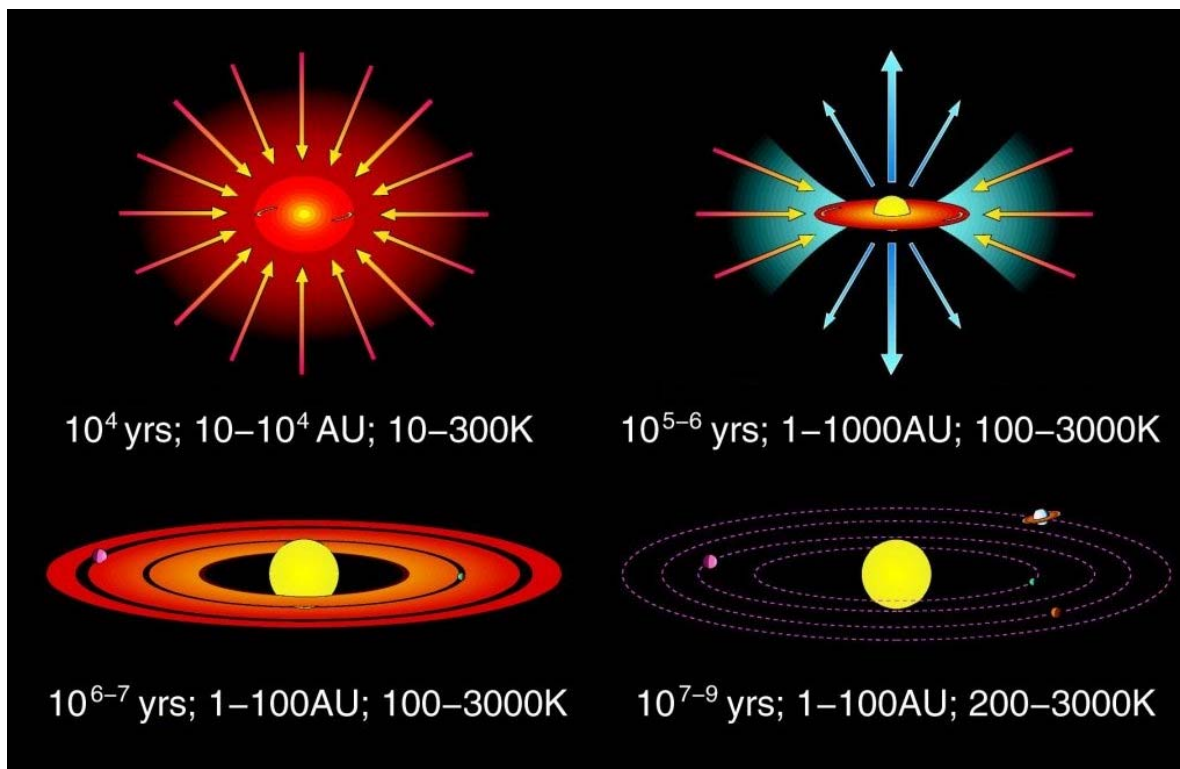
As illustrated in Figure 5-1, the formation of stars and planets is a complex process that starts on large scales, as molecular cloud cores cool and fragment to form highly dynamic clusters of protostars, which span the entire mass spectrum from O and B stars down to planetary-mass brown dwarfs. Within those clusters, individual young sources are generally encircled by disks of warm gas and dust, where material aggregates to form planetary systems. These disks also act as the source of highly-collimated jets and outflows, which remove energy and angular momentum from the infalling material, inject them into the surrounding medium, and clear away the remainder of the birth core. On larger scales, the intense ultraviolet flux and strong winds of the most massive stars can disperse an entire molecular cloud, while simultaneously ionizing and evaporating the circumstellar disks of the surrounding lower-mass stars.

Young stars, brown dwarfs, and circumstellar disks emit the bulk of their radiation at near- and mid-infrared wavelengths, and at the earliest stages, shorter wavelength emission is absorbed by the dust surrounding them. To probe deep into these obscured regions and detect emission from gas and dust at temperatures ranging from 100-10000 K, imaging and spectroscopic observations from 1-27  $\mu\text{m}$  are required. High sensitivity, high spatial resolution, and a large dynamic range are required to study the physical properties, composition, and structure of faint companions, disks, and protoplanets immediately adjacent to their much brighter neighbors. Finally, a large field-of-view is needed to ensure that the diverse range of sources, phenomena, and their interactions within a given star-forming complex can be captured and disentangled.

With these characteristics, JWST must observe star-forming regions to answer these fundamental questions:

- How do protostellar clouds collapse?
- What is the early evolution of protostars?
- How do massive stars form and affect their environment?
- What is the initial mass function at sub-stellar masses?
- How do protoplanetary systems form?
- What are the life cycles of gas and dust?

JWST will be one of the cornerstone facilities available in the decade 2010-2020, which include other space-based missions, and large ground-based optical and near-infrared telescopes and sub-millimeter interferometer arrays. Their combined power should deliver extraordinary new insights into the complex interplay between the various processes that yield stars and their planetary systems.



**Figure 5-1. How a Single, Isolated, Low-Mass Star and its Planetary System are Formed**

Note: A schematic diagram showing the formation of an isolated star (after Shu et al. 1987). Following a deeply embedded protostellar collapse phase (Class 0 YSO; top-left), a circumstellar disk and collimated outflow are established, which renders the central star visible at most orientations (Class I/II; top-right). After accretion from the envelope is terminated, perhaps by environmental influences, planetesimals and protoplanets form in the passive disk via sedimentation and agglomeration (Class III; bottom-left), later leaving a mature planetary system in orbit around the star (bottom-right). The range of temperatures (10 to 3000 K) involved implies that the bulk of the radiation seen in the early phases comes out at near- and mid-infrared wavelengths, as well as the millimeter. Typical size scales (1 to 1000 AU, or 0.002 to 2 arcsec at 500 pc) imply that high spatial resolution is required for such studies.

## 5.1 HOW DO PROTOSTELLAR CLOUDS COLLAPSE?

### 5.1.1 Scientific Rationale

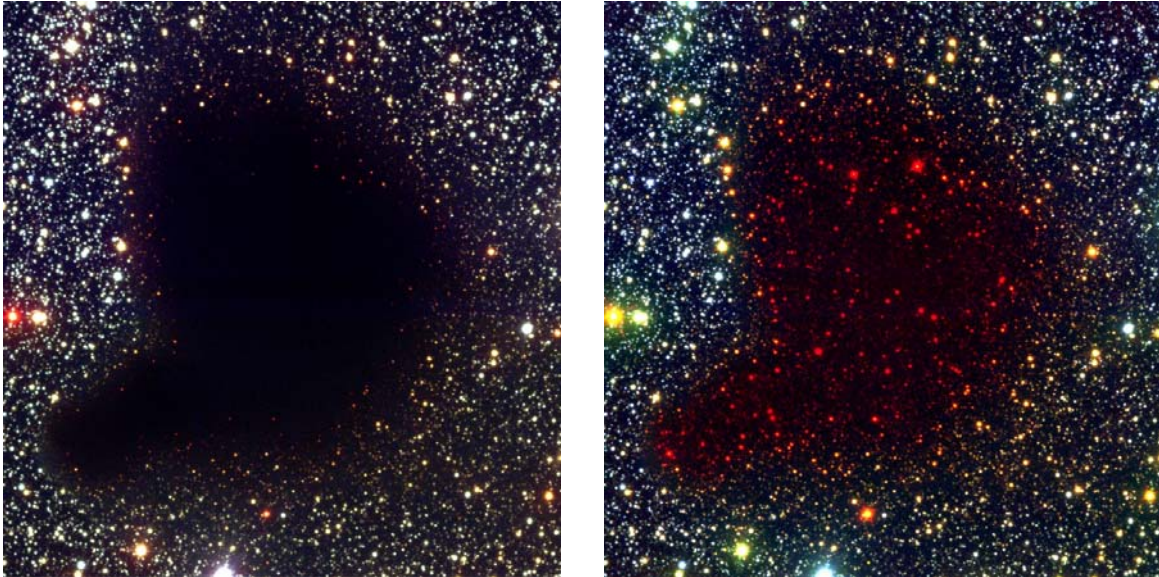
Stars form in small ( $\sim 0.1$  pc) regions undergoing gravitational collapse within larger molecular clouds. These dense cores have densities  $n_{\text{H}_2} > 10^4 \text{ cm}^{-3}$ , roughly a hundred times greater than ambient cloud material. Standard theory predicts that these cores collapse from the inside out (e.g., Shu 1977; Terebey et al. 1984), with the collapse propagating at an effective sound speed of about  $0.3 \text{ km s}^{-1}$ , accounting for gas pressure and support due to magnetic fields and turbulence. The slowly collapsing and slowly rotating core approximates a singular isothermal sphere, breaking down in the center where a protostar and a more rapidly rotating disk are found. However, there are alternatives to the standard picture. Ambipolar diffusion, due to incomplete

coupling of magnetic fields to the gas, can result in rigid, as opposed to differential, rotation of the cloud core (Mouschovias & Paleologou 1981; Crutcher et al. 1994). Furthermore, cores may be externally pressure-confined (Alves et al. 2001), or be altogether more chaotic and dynamic structures formed in the intersections of fractal clouds (Bate, Bonnell, & Bromm 2003). These different models predict quite different density distributions for star-forming cores, and thus by measuring those density distributions for cores in a wide range of environments and evolutionary states, we can hope to understand the relative roles that magnetic fields, turbulence, and rotation play while the clouds collapse to form stars.

Observations of optically thin dust emission at millimeter continuum wavelengths have been used to trace the structure of dense cores, but the inversion of a measured intensity profile into a density profile is difficult, as it relies on an assumed underlying temperature profile and three-dimensional structure. For example, it has not yet been possible to distinguish unambiguously between flattened and peaked central cores (Ward-Thompson et al. 1994, 1999; Evans et al. 2001; Zucconi et al. 2001). The low spatial resolution of current sub-millimeter telescopes (about 10 arcsec) is also a problem that will be partly alleviated by future sub-millimeter interferometers (e.g., SMA, ALMA). An alternative technique involves mapping the extinction seen along various lines-of-sight through a cloud core: the extinction can be directly related to the dust column density, so a 2D projection of the core density profile can be deduced, assuming a fixed gas-to-dust ratio.

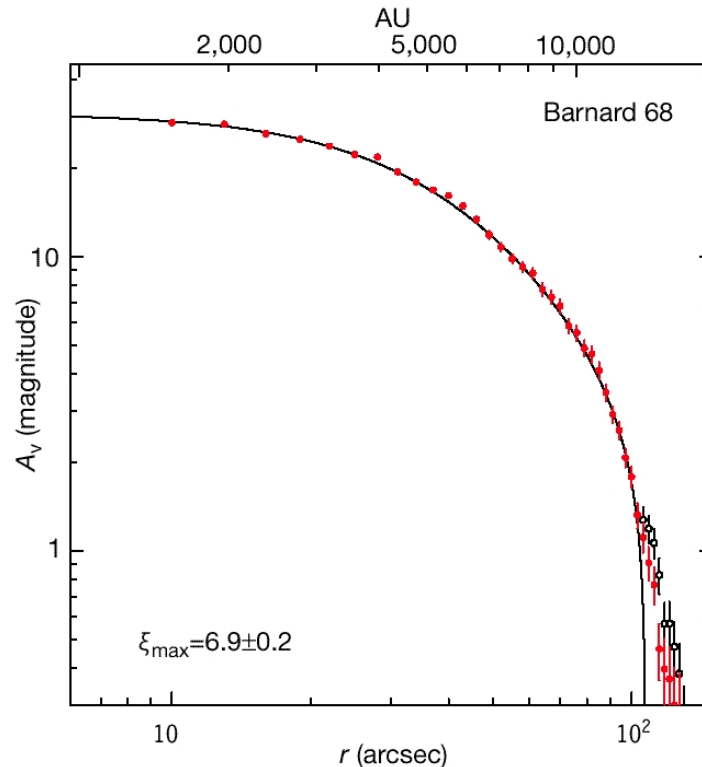
It is also possible to map the extinction profile of a cloud by measuring the colors of discrete background field stars seen through a core at near-infrared wavelengths. Used on large ground-based telescopes and reaching a typical depth of  $K_{AB} = 22$  mag with subarcsecond spatial resolution, this technique provides a resolution of 10-15 arcsec through extinctions of up to  $A_V \sim 60$  mag in dark clouds (e.g., Lada et al. 1994; Alves et al. 1998; Alves et al. 2001; see Figure 5-2)

An alternative approach maps the attenuation of the diffuse mid-infrared background produced by the interstellar radiation field or by hot sources in the same star-forming complex as the core. This background is particularly bright in the 6.2 and 7.7  $\mu\text{m}$  PAH emission features, where dust extinction is also near a minimum. In this manner, Bacmann et al. (2000) used ISOCAM to measure extinction profiles in pre-stellar cores with a spatial resolution of 10 arcsec through extinction of up to  $A_V \sim 50$  mag.



**Figure 5-2. The Low-Mass Dark Cloud Barnard 68**

Note: Barnard 68 is imaged at optical and near-Infrared wavelengths using FORS1 on the ESO VLT and SOFI on the ESO NTT. The left panel shows a color composite of optical B, V, and I images, while the right panel shows a composite of the B and I images with a near-infrared  $K_S$  image. The images cover  $4.9 \times 4.9$  arcmin, or  $0.18 \times 0.18$  pc. At optical wavelengths, the small cloud is completely opaque because of the obscuring effect of dust particles in its interior. Since the light from stars behind the cloud is only visible at the infrared wavelengths, they appear red. Using other infrared images and measuring the extinction on a star-by-star basis, the dust column density profile of B68 could be measured (see Figure 5-3, from Alves et al. 2001).



**Figure 5-3. Azimuthally Averaged Radial Dust Column Density Profile for B68**

Note: The red circles show the averaged profile of a subsample of data that does not include the southeast prominence of the cloud (see Figure 5-2) while the open circles include the prominence. The solid line represents the best fit of a theoretical Bonnor-Ebert sphere to the data. The close match suggests that the internal structure of the cloud is well characterized by a self-gravitating, pressure-confined, isothermal sphere, and the cloud appears to be near hydrostatic equilibrium. (From Alves et al. 2001).

### 5.1.2 Key JWST Observations

Measuring the larger-scale structure in clouds can be done from the ground, but to probe the centers of pre-stellar cores and Class 0 envelopes, substantially higher sensitivity and spatial resolution are required to detect the much fainter and redder stars through the compact core. To detect them, the extinction penetration must be twice that available with ground-based observations, i.e.  $A_V \sim 120$  mag, which requires an additional 7 mag of sensitivity at  $2 \mu\text{m}$ .

To achieve this, JWST must reach a  $10\sigma$  point-source limiting sensitivity of at least  $K_{AB} = 29$  mag, or 9 nJy at  $2 \mu\text{m}$ . Diffraction-limited spatial resolution is required to ensure high-fidelity mapping of the extinction profile with  $\sim 1$ -2 arcsec resolution, i.e., 200-500 AU for clouds at a few hundred parsecs. Finally, in order to map the full extinction profile of a typical 0.1 pc radius cloud core at the same distance, a field-of-view of 2 to 4 arcmin is required.

The centers of cores and Class 0 objects have even more extinction, and thus mid-infrared observations using the extended background emission technique are required. A typical mid-infrared background flux of  $10 \text{ MJy sr}^{-1}$  yields an unattenuated surface brightness of  $240 \text{ } \mu\text{Jy arcsec}^{-2}$ .

In order to observe this surface brightness through  $A_V \sim 300 \text{ mag}$  (or  $A_{7\mu\text{m}} \sim 7 \text{ mag}$ , assuming a standard extinction law), JWST requires a  $10\sigma$  surface-brightness sensitivity of  $1 \text{ } \mu\text{Jy arcsec}^{-2}$  over the  $6.7$  to  $7.7 \text{ } \mu\text{m}$  region where the background is bright and the dust extinction low: binning of the pixels to  $1 \text{ arcsec}$  can be used to increase the surface brightness sensitivity. In order to map the central regions of cores and connect the results with those obtained for the lower-density outskirts in the near-infrared, a mid-infrared field-of-view larger than  $1 \text{ arcmin}$  (i.e.,  $0.15 \text{ pc}$  at  $500 \text{ pc}$ ) is required.

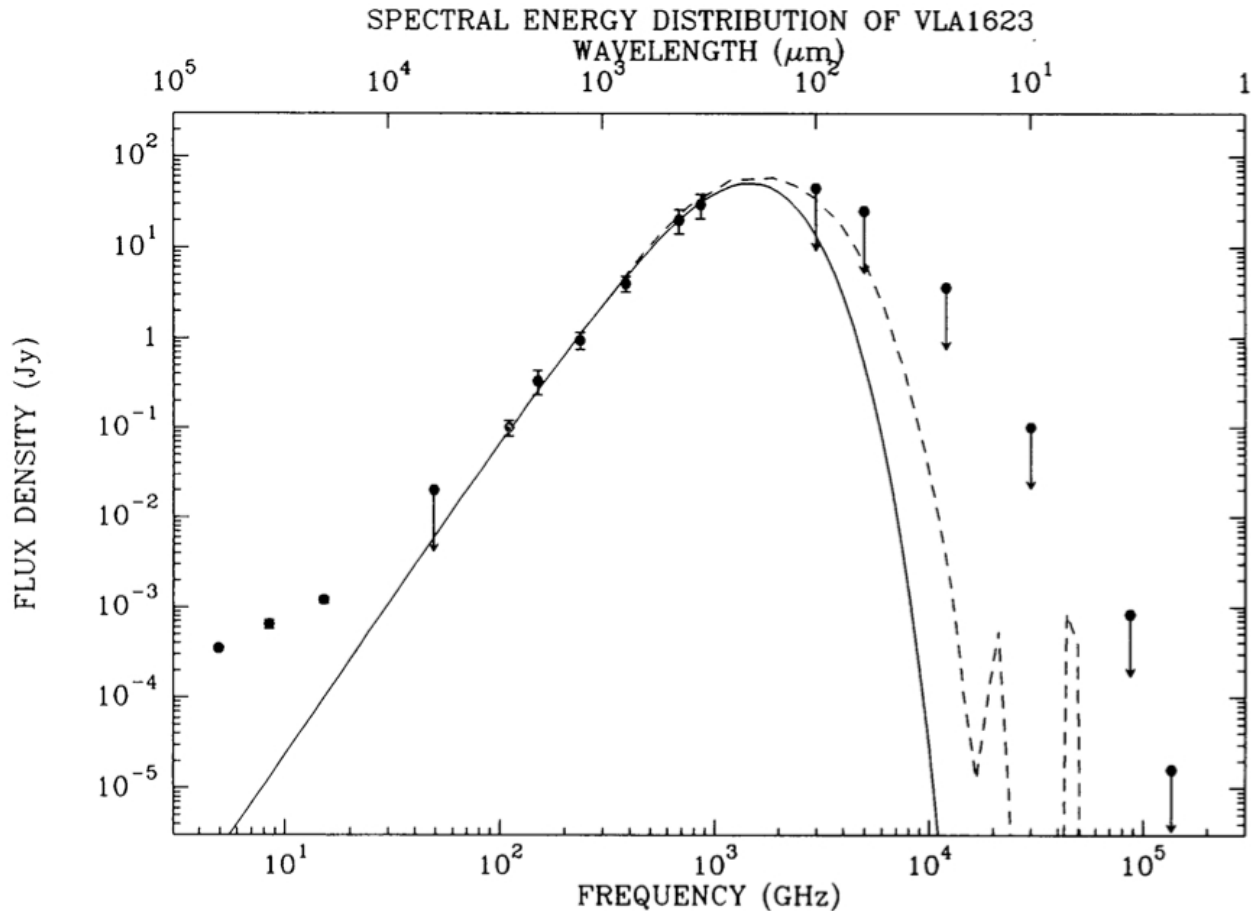
## 5.2 WHAT IS THE EARLY EVOLUTION OF PROTOSTARS?

### 5.2.1 Scientific Rationale

Once self-gravitating molecular cloud cores have formed, they can collapse to form protostellar seeds, which gain material via continuing accretion. The earliest category of protostar, the Class 0 object (André, Ward-Thompson, & Barsony 1993), is deeply embedded in and obscured by the massive envelope from which it is accreting, and its spectral energy distribution is dominated by this cold ( $\sim 20 \text{ K}$ ) material. As a result, these young ( $\sim 10^4 \text{ yrs}$ ) sources emit the bulk of their flux at millimeter and sub-millimeter wavelengths, and are generally undetected at shorter wavelengths to date (Figure 5-4).

Detecting and studying the  $10 \text{ } \mu\text{m}$  to  $20 \text{ } \mu\text{m}$  emission from protostars is important. Radiative transfer models (Wolfire & Cassinelli 1986, 1987; André et al. 1993; Boss & Yorke 1995) predict that there should be a warm ‘shoulder’ in the mid-infrared in the otherwise single  $\sim 20 \text{ K}$  blackbody SED, and that protostars should be roughly 1000 times brighter than the blackbody, as radiation from the warm central source is scattered off dust grains in the inner envelope into the line-of-sight. The degree of scattering is a strong function of the density distribution in the envelope, so the departures from the single blackbody SED at mid-infrared wavelengths would be an important diagnostic of envelope structure, most critically the power law of the density distribution.

Cernicharo et al. (2000) confirmed these predictions with ISOCAM detections of a few luminous Class 0 protostars. Imaging in selected narrow bands ( $5.3$ ,  $6.6$ ,  $7.5 \text{ } \mu\text{m}$ ) between ice and silicate absorption features, warm material ( $\sim 700 \text{ K}$ ) was observed through effective extinctions of  $A_V \sim 80$  to  $100 \text{ mag}$ , and with flux coming from within  $4 \text{ AU}$  of the accreting protostars. More detailed observations of this kind are required for a wider range of protostellar luminosities, in order to constrain the central protostellar parameters in envelope models, so that density distributions can be extracted more accurately from the Class 0 envelope observations.



**Figure 5-4. The Spectral Energy Distribution of the Prototypical Class 0 Protostar**

Note: VLA 1623, in Ophiuchus. Despite a nominal blackbody temperature of only  $\sim 20$  K, radiative transfer models do predict significant mid-infrared flux as emission from the warm core is scattered off dust grains in the inner envelope. Adapted from André et al. 1993.

The dynamics of the protostellar collapse can be diagnosed through imaging spectroscopy of shocks, which form as material accretes onto the inner envelope and disk and the vertical velocity component is dissipated. The models of Yorke & Bodenheimer (1999) predict at least two shock fronts at 500 to 1000 AU from the protostar, with positions changing as a function of evolution in the system, i.e., moving further from the source as the disk grows, but disappearing once accretion terminates.

Finally, the origin of binary stars is not yet understood. There is growing evidence that the majority of all stars form in binary or higher-order multiple systems. While some theoretical predictions of fragmentation models are supported indirectly by statistical studies of evolved binary systems at optical and near-infrared wavelengths, direct observations of the binary formation phase itself became possible only recently with the advent of large, sensitive millimeter interferometers. However, as noted above, millimeter wavelength observations can only probe extended envelopes, not the protostellar cores themselves. Deep high-resolution

imaging at 10  $\mu\text{m}$  is required to observe the central hydrostatic cores in simultaneously turbulent, rotating, fragmenting, and collapsing protostellar clouds. In combination with detailed kinematic data supplied by future millimeter interferometers such as ALMA, such data will provide crucial tests of binary fragmentation models, allowing us to determine true initial binary fractions and separations, and how these properties change as stars evolve through the pre-main-sequence phases.

### **5.2.2 Key JWST Observations**

In order to characterize the density structure in the envelopes and cores of Class 0 sources, broad-band fluxes from 10 to 20  $\mu\text{m}$  and narrow-band imaging in the 5 to 7  $\mu\text{m}$  extinction windows are required. To detect such young protostars and protostellar cores at a distance of  $\sim 150$  pc (the distance of the nearest star-forming regions like Taurus and Chamaeleon), JWST must reach sensitivities of  $1\mu\text{Jy}$  at 6  $\mu\text{m}$  and  $10\mu\text{Jy}$  at 15  $\mu\text{m}$  (Figure 5-4). Going out to a distance of 500 pc, and thus encompassing a much wider selection of star-forming regions, sensitivities of  $0.1\mu\text{Jy}$  and  $1\mu\text{Jy}$  will be required at 6 and 15  $\mu\text{m}$ , respectively.

High spatial resolution ( $<1$  arcsec FWHM) is required at mid-infrared wavelengths, as Class 0 protostars should be slightly extended to dust scattering, with diameters  $\sim 1000$  AU, i.e., 2 to 6 arcsec in star-forming regions at 150 to 500 pc. In addition, the sources need to be resolved out against the larger (2000 to 5000 AU) envelope in which they are embedded (e.g., Boss & Yorke 1995).

Resolving binary protostars at the peak of the separation distribution for pre-main sequence stars ( $\sim 30$  AU) requires an angular resolution less than 0.25 arcsec at 6  $\mu\text{m}$ , where hot dust emission from the inner parts of accretion disks around slightly more evolved (Class I) protostars. The preferred wavelength range to detect embedded protostars directly is  $\sim 20$  to 25  $\mu\text{m}$ ; diffraction-limited imaging ( $\sim 0.8$  arcsec FWHM) would resolve binary protostars at separations that are slightly beyond the peak of the period distribution. ISO and Spitzer do not have sufficient angular resolution to distinguish the protostars from extended emission by externally heated small grains in the envelopes or to resolve typical binary protostellar separations.

For brighter protostars and cores, we require mid-infrared imaging spectroscopy of a range of narrow-band features (extinction windows, shock tracers, and PAHs) and the intervening continuum simultaneously, thus yielding a large range of diagnostics while ensuring accurate spatial registration when calculating temperatures and sizes of the sources. Assuming the necessary spectral resolution of  $R = 2000$  at 15  $\mu\text{m}$ , a continuum sensitivity of  $\sim 7 \times 10^{-22}$   $\text{W m}^{-2}$  per resolution element is required to detect a 10  $\mu\text{Jy}$  source, i.e., a typical Class 0 protostar in Taurus-Auriga.

## **5.3 HOW DO MASSIVE STARS FORM AND AFFECT THEIR ENVIRONMENT?**

### **5.3.1 Scientific Rationale**

The formation of massive stars produces intense winds and ionizing radiation which impacts the surrounding molecular cloud material and the nascent circumstellar disks of adjacent low-mass

stars. And yet we don't know the mechanism by which they form, or whether they are accompanied by circumstellar disks. A simple scaling-up of the standard low-mass star formation paradigm predicts that at some point, the radiation pressure from the growing central source builds up so rapidly that no more material can accrete, thus limiting the mass of the source (Yorke & Sonnhalter 2002).

Various solutions have been proposed to solve this problem. McKee & Tan (2002) have suggested that high accretion rates can overcome the radiation pressure, so a massive star can form 'normally'. Alternatively, Bonnell et al (1998) proposed that massive stars form by the agglomeration of colliding low-mass stars in a very deep potential well. In this model, stellar densities of  $10^7$  to  $10^8$  stars per cubic parsec are required, on the order of 1000 times the density seen in the core of present-day young embedded clusters within their central 0.1 pc. The required high densities would be achieved by shrinking such a cluster by a factor of 10 in linear size, i.e., to a core radius of  $\sim 0.01$  pc, or  $\sim 2$  arcsec at 1 kpc, producing a hypercompact H II region.

Once born, massive stars can be disturbing neighbors. In particular, ionizing photons and strong winds from O and B stars can wreak disproportionate damage on their environment. On small scales, they can destroy disks around young low-mass stars (Johnstone et al. 1998; Bally et al. 2000), while on larger scales, they can simultaneously evaporate and compress surrounding molecular material, either halting or triggering further star formation in surrounding molecular material (Larosa 1983; Bertoldi 1989; Lefloch & Lazareff 1994).

An example of these environmental impacts is seen in M 16, the Eagle Nebula where parsec-scale molecular trunks are evaporated and disrupted by O and B stars of the adjacent NGC 6611 cluster (Hillenbrand et al. 1993). A population of small 1000 AU-scale (0.5 arcsec at M 16) dense globules is seen on the surfaces of the trunks (Hester et al. 1996). These evaporating gaseous globules may contain young stars about to be revealed as the O and B stars evaporate their birth clouds. Near-infrared observations show that about 10 to 20% contained embedded low-mass stellar and brown dwarf candidates (McCaughrean & Andersen 2002; see Figure 5-5), in addition to the tips of the ablating columns also containing higher-mass YSOs (Thompson, Smith, & Hester 2002; Sugitani et al. 2002).

Although columns like these may be associated with recent star formation, it remains unknown whether it is triggered by the passage of an ionization front from the O and B stars, or whether pre-existing YSOs are revealed as the parent cores are eroded away. Do O and B stars have a pro-active role in creating new stars via radiative implosion, or are they simply destructive, exposing stars prematurely, perhaps reducing their final masses and destroying their disks? To what extent is the mass of a star (and by extension, the whole stellar IMF) determined by when and how its envelope of accreting material is stripped away by nearby massive stars rather than processes local to the star itself?



**Figure 5-5. The M 16 Elephant Trunks**

Note: True-Color Near-Infrared (1-2.5  $\mu\text{m}$ ) Image of the M 16 Elephant Trunks made using the ESO VLT. The  $J_s$  data are shown as blue,  $H$  as green, and  $K_s$  as red. The image covers  $158 \times 214$  arcsec ( $1.5 \times 2.0$  pc at 1.9 kpc); north is up, east left. Subimages show more detail, including EGGs from Hester et al. (1996) found to be associated with point sources; E23, an EGG with no near-infrared point source, but thought to contain an embedded protostar driving a collimated jet; YSO1 and YSO2, massive sources in the tips of C1 and C2, respectively; and HH 216, an optically-visible Herbig-Haro object. (From McCaughrean & Andersen 2002).

### 5.3.2 Key JWST Observations

In order to solve the massive star formation paradox, deep, high spatial resolution imaging is required in the mid-infrared to see through the high extinction. These stars are obscured by the considerable density of gas and dust that accompany them: a prototypical hypercompact H II

region might contain a volume density of  $10^7 \text{ cm}^{-3}$  in a region 0.01 pc (2000 AU) in radius, yielding a column density to the center of  $3 \times 10^{23} \text{ cm}^{-2}$  or  $A_V \sim 150 \text{ mag}$ .

On the basis of extinction alone, observations at 8 and 13  $\mu\text{m}$  would be preferred, but some trade-off with spatial resolution must be sought, given the potentially very high density of point sources that would be expected in the agglomeration hypothesis, for example, where as many as 50 to 100 low-mass stars might occupy a region only 4 arcsec across.

The best compromise is found at somewhat shorter wavelengths. At 3.8 and 4.8  $\mu\text{m}$ ,  $A_V \sim 150 \text{ mag}$  would be reduced to 6 and 4 mag, respectively. In order to detect a fiducial  $0.1 M_{\odot}$ , 1 Myr old star at 1 kpc, seen through this dust column, JWST must reach point source limits of  $\sim 2$  and 10  $\mu\text{Jy}$ , respectively (Baraffe et al. 1998). Given the impact of crowding, JWST must combine good sensitivity with large dynamic range and excellent, stable imaging quality. With diffraction-limited imaging at these wavelengths,  $\sim 0.15$  arcsec resolution should yield about 30 fully-sampled resolution elements across 4 arcsec, equivalent to 4000 AU at 1 kpc.

To determine the impact of massive stars on their environments, comprehensive surveys must be made of dark clouds and elephant trunks in regions with recent massive star formation, to reveal populations of young stars and study their properties. By examining the masses and ages of the sources as a function of their distance from the ionization front, it will be possible to determine whether O and B stars simply reveal pre-existing star formation or trigger it directly.

Sensitive mid-infrared observations are required to penetrate the dust in such trunks and reveal pre-main sequence stars and brown dwarfs, as well as very young protostellar sources. Broad-band imaging at 3 to 20  $\mu\text{m}$  is required to provide a census of the photometric properties of deeply embedded and/or very young sources, while follow-up spectroscopy is necessary to compare the mass, ages, and evolutionary status of the sources in a given trunk with standard models. In addition, narrow-band line imaging in a variety of ionized, atomic, and molecular tracers is required to delineate the rate at which photoionization, photoevaporation, and shocks are propagating into the molecular cloud core and to determine the balance between implosion and destruction. All of these observations must be conducted at high spatial resolution to permit unambiguous identification of a given source with a given globule: typical sizes are 200 to 2000 AU or 0.1 to 1 arcsec at 2 kpc.

Young sources embedded in such regions will vary considerably in flux as a function of mass, wavelength, extinction, and circumstellar excess emission. For example, the sensitivity required to detect a 1 Myr old,  $0.02 M_{\odot}$ , low-mass brown dwarf (Baraffe et al. 1998) with typical infrared excess emission due to a circumstellar disk (Kenyon & Hartmann 1995) would be  $\sim 55 \text{ nJy}$ , 2.3  $\mu\text{Jy}$ , and 34  $\mu\text{Jy}$  at 3.5, 4.8, and 20  $\mu\text{m}$ , respectively, assuming a distance of 2 kpc and an extinction of  $A_V = 100 \text{ mag}$ . Under similar conditions, a  $0.075 M_{\odot}$  source at the star to brown dwarf boundary would require sensitivities of 0.65, 24, and 370  $\mu\text{Jy}$ , respectively, at the same wavelengths. Broad-band imaging photometry is required in the near- and mid-infrared down to the  $0.02 M_{\odot}$  limit, with multiobject spectroscopy at  $R = 100$  for the shorter near-infrared wavelengths and  $R=1000$  at the longer. In the mid-infrared, integral field spectroscopy at  $R = 2000$  of selected individual sources down to the star-brown dwarf boundary is required.

## 5.4 WHAT IS THE INITIAL MASS FUNCTION AT SUB-STELLAR MASSES?

### 5.4.1 Scientific Rationale

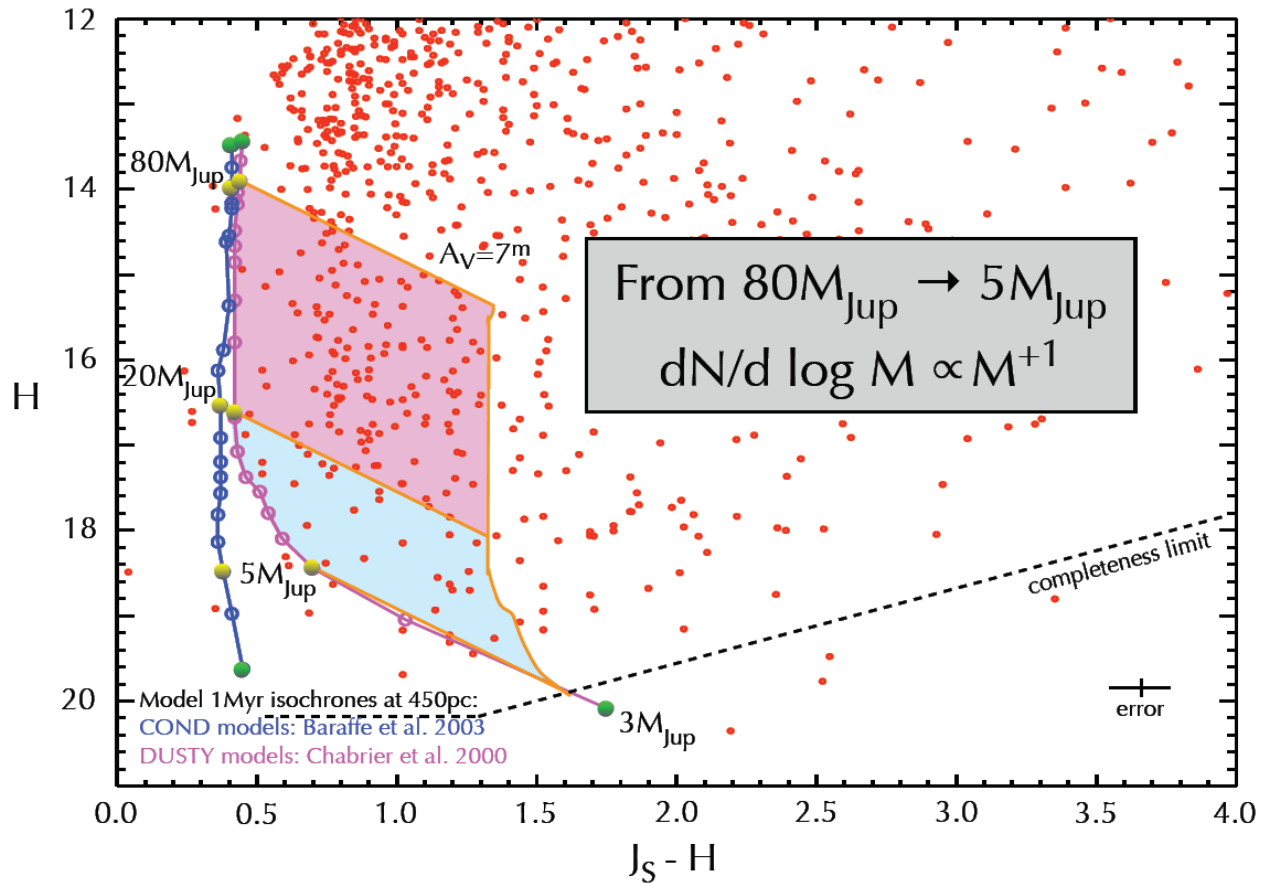
The initial mass function (IMF) is a key product of star formation (Salpeter 1955; Miller & Scalo 1979). Remarkably, the IMF is almost entirely feature-free, all the way from the most massive stars down to  $\sim 0.3 M_{\odot}$ . Below  $0.3 M_{\odot}$ , studies show that the IMF flattens somewhat, but continues to increase. Below  $0.1 M_{\odot}$  ( $100 M_{\text{JUP}}$ ), the mass function starts to decline, and microlensing observations (Alcock et al. 1998) clearly indicate that our Galaxy is not full of sub-Jupiter mass brown dwarfs. Thus, somewhere below  $0.1 M_{\odot}$ , the physics of star formation produces a turnover and decline in the IMF, perhaps with some lower-limit boundary condition.

Theory has long predicted a significant boundary around 3 to  $10 M_{\text{JUP}}$ , below which it is believed that fragmentation and collapse of molecular cloud cores cannot occur (Hoyle 1953; Low & Lynden-Bell 1976; Rees 1976; Silk 1977; Boss 1988). More recent work by Boss (2001) has shown that magnetic fields may lower this limit to about  $1 M_{\text{JUP}}$ . The whole fragmentation scenario at low masses may have to be replaced by a more complex model involving a wide range of physical processes, including supersonic turbulence (Padoan & Nordlund 2002), dynamical interactions between protostars (Bate, Bonnell, & Bromm 2002), and feedback due, for example, to strong bipolar outflows (Adams & Fatuzzo 1996) and ionizing radiation from massive stars (Palla & Stahler 2000).

Thus the form of the substellar IMF can yield important clues in our understanding of the star formation process. Studies of embedded and young open clusters 10 to  $100 M_{\text{JUP}}$  regime yield surprisingly different results. On one hand, a rising substellar IMF has been found in the 100 Myr Pleiades (Bouvier et al. 1999; Zapatero-Osorio et al. 1999) and other young open clusters, while in the 1 Myr old Trapezium Cluster there is a strong turnover of the IMF across the stellar to substellar boundary, with a smaller proportion of brown dwarfs down to 5 to  $10 M_{\text{JUP}}$  (Hillenbrand & Carpenter 2000; Luhman et al. 2000; Muench et al. 2002; Lucas & Roche 2000; McCaughrean et al. 2002).

Searching for and characterizing sources at  $1 M_{\text{JUP}}$  and below in nearby star-forming regions would allow us to constrain the physics of any lower-limit boundary (Figure 5-6). In addition, free-floating objects less massive than  $10 M_{\text{JUP}}$  can serve as important proxies for true planets, yielding crucial insights into their early evolution and complementing JWST studies that will be made of mature giant planets in orbit around nearby stars.

Furthermore, the near-infrared sensitivity and high spatial resolution of the JWST will enable us to search for brown dwarfs in distant clusters with greatly differing metallicity, including globular clusters, clusters in the inner and outer galaxy, and in the Magellanic Clouds. In this way, we will be able to determine whether the substellar IMF is universal, or if it depends on the local Jeans mass.



**Figure 5-6. Color-Magnitude Diagram for the Orion Trapezium Cluster**

Note: Figure shows the H vs.  $(J_s - H)$  color-magnitude diagram for the Orion Trapezium Cluster. An extinction-limited sample of sources in the range  $0.005$  to  $0.08 M_{\odot}$  ( $5$  to  $80 M_{JUP}$ ), assumed to be  $1$  Myr old, shows that the brown dwarf end of the IMF is falling, as roughly characterized by the form  $dN/d \log M \sim M^{+1}$  (McCaughrean et al 2002).

### 5.4.2 Key JWST Observations

At 1 Myr, sources below  $5 M_{\text{JUP}}$  have effective temperatures below 1000 K, implying that the bulk of their flux moves into the mid-infrared, making ground-based observations very difficult. JWST must be sensitive enough at 3 to 10  $\mu\text{m}$ , to survey nearby young clusters to  $1 M_{\text{JUP}}$ .

Broadband imaging observations are required to identify very low luminosity candidates, and to compare their properties with pre-main sequence evolutionary models. Imaging surveys at 2 to 5  $\mu\text{m}$  must be supplemented with 10  $\mu\text{m}$  photometry to characterize circumstellar emission due to disks and thus help eliminate it when calculating the intrinsic source luminosity. Wide-field imaging is required for nearby clusters: a typical cluster is  $\sim 1$  pc in diameter or 3.5 arcmin at 1 kpc. At larger distances, high spatial resolution is required to eliminate crowding and resolve the lowest-mass sources from adjacent brighter, more massive neighbors.

Multiobject spectroscopy at 2 to 5  $\mu\text{m}$  is then required to distinguish between true cluster members and field interlopers. These observations will also break the age-reddening-mass degeneracy found in young clusters, by determining spectral types and thus making it possible to accurately deredden the sources and locate them in theoretical pre-main sequence HR diagrams. Typically, within a given cluster, there will be 100 to 1000 sources with masses in the range 1 to  $100 M_{\text{JUP}}$ .

We must obtain full  $R = 1000$  multiobject spectra across the 2 to 5  $\mu\text{m}$  regime. These data will allow spectral typing with the  $\text{H}_2\text{O}$  and methane absorption bands, atomic Na, K, and Ca lines, and  $\text{H}_2$  collisionally-induced absorption. We will eliminate older, high-gravity, non-cluster members, and assign spectral-type based effective temperatures to the true cluster members.

In extremely crowded regions with bright nebulosity however, contrast and light leakage in the spectrograph becomes an issue, reducing sensitivity to the lowest-mass sources. Narrow-band imaging provides an excellent alternative, using photometry in a set of about 10 select narrow-band features to construct spectral indices with respect to the atmospheric models. These data could be compared to an optimal set of spectral templates across the 2 to 5  $\mu\text{m}$  regime as a function of mass and age (e.g., Burrows et al. 2001; Baraffe et al. 2003).

The mass limit that can reasonably be reached in a given cluster is a function of its age, distance, and the foreground and intracluster reddening, all of which can vary considerably. Table 5-1 shows the sensitivities required to reach mass limits of 1, 10, and  $100 M_{\text{JUP}}$ , assuming a 1 Myr old cluster at 500 pc (e.g., Orion), 5 kpc (inner galaxy), and 50 kpc (Magellanic Clouds), with 10 magnitudes of visual extinction in each case.

**Table 5-1. Predicted Fluxes for Sub-Stellar Objects**

Mass ( $M_{\text{JUP}}$ )	$T_{\text{eff}}$ (K)	Distance (kpc)								
		0.5			5			50		
		2.2 $\mu\text{m}$	3.8 $\mu\text{m}$	4.8 $\mu\text{m}$	2.2 $\mu\text{m}$	3.8 $\mu\text{m}$	4.8 $\mu\text{m}$	2.2 $\mu\text{m}$	3.8 $\mu\text{m}$	4.8 $\mu\text{m}$
1	941	290	270	1300	2.9	2.7	13	0.03	0.03	0.13
10	2251	34000	33000	28000	340	330	280	3.4	3.3	2.8
100	2856	950000	820000	760000	9500	8200	7600	95	82	76

Note: Predicted Fluxes in nJy at near-infrared wavelengths for 1  $M_{\text{JUP}}$ , 10  $M_{\text{JUP}}$ , and 100  $M_{\text{JUP}}$  Sources at 1 Myr, at Distances of 0.5, 5, and 50 Kpc. A typical extinction of 10 magnitudes at V-band is assumed throughout. (Burrows et al. 1997; Burrows et al. 2001; Baraffe et al. 1998; Chabrier et al. 2000; Marley et al. 2002; Baraffe et al. 2003).

For the nearby clusters at  $\sim 500$  pc, JWST must carry out accurate broad-band near- and mid-infrared photometry, as well as classification spectroscopy at 2 to 4  $\mu\text{m}$  down to 1  $M_{\text{JUP}}$ , requiring a limiting continuum sensitivity of 300 nJy. This spectroscopy must be possible using either multiobject  $R = 1000$  spectroscopy or  $R \sim 100$  imaging in (approximately) 10 bands.

Towards the inner galaxy, imaging surveys down to 1  $M_{\text{JUP}}$  are required, with classification spectroscopy down to 10  $M_{\text{JUP}}$ . Finally, at the Magellanic Clouds, imaging down to 10  $M_{\text{JUP}}$  is required, with spectroscopy to 100  $M_{\text{JUP}}$  or 0.1  $M_{\odot}$ , just above the star-brown dwarf boundary. The massive 30 Doradus cluster in the Large Magellanic Cloud is the closest object we have to a starburst template; a detailed study of its low-mass content is crucial.

In all young clusters, 3 to 10  $\mu\text{m}$  imaging is required to measure excess thermal emission indicative of disks around brown dwarfs, and thus assess their mode of formation and whether or not they may build planetary systems (cf. McCaughrean et al. 1996; Muench et al. 2001; Natta & Testi 2001; Liu et al. 2003).

High spatial resolution images of clusters are required to measure the binary frequency function to separations of 30 AU (0.06 arcsec at 500 pc), the peak of the main sequence binary separation distribution. We must determine whether the frequency and properties of field binaries and multiples can be reproduced by mixing the populations of clusters and low-mass star-forming regions in different proportions (Kroupa et al. 1999; Scally et al. 1999).

Similarly, high spatial resolution is required to conduct high accuracy proper motion measurements to analyze the internal dynamical state of a cluster, to determine what fraction of it will remain bound and what effect dynamical mass segregation has had on the measured mass function (Kroupa 1998). Assuming a centroiding accuracy for point sources equal to the Gaussian sigma of the point spread function divided by the signal-to-noise, we require  $S/N = 50$  to yield 0.8 milliarcsec precision at 2  $\mu\text{m}$  or 0.4 AU at 500 pc. Measurements at 3 epochs over 3 years will permit individual proper motions to be measured for sources down to 1  $M_{\text{JUP}}$  to an

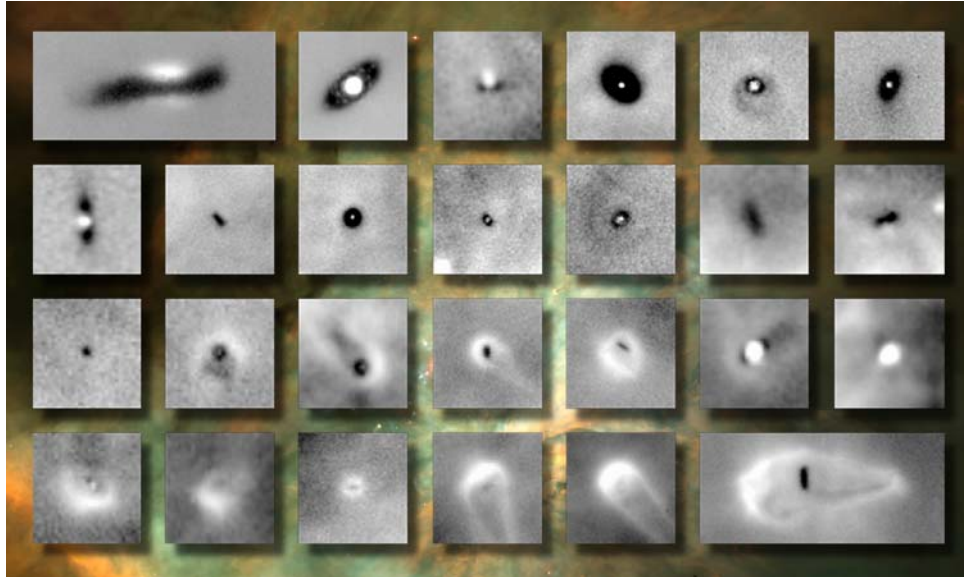
accuracy of about  $1 \text{ km s}^{-1}$  in nearby clusters, resolving the typical cluster velocities of  $3\text{-}4 \text{ km s}^{-1}$  (Jones & Walker 1988).

## 5.5 HOW DO PROTOPLANETARY SYSTEMS FORM?

### 5.5.1 Scientific Rationale

The existence of disks around young low-mass stars was firmly established in the 1990s, initially based on indirect measurements, such as spectral energy distributions, asymmetric wind profiles, polarization mapping, (Beckwith & Sargent 1993; Strom et al. 1993b), and more recently through direct imaging (McCaughrean et al. 2000; Wilner & Lay 2000). Well-resolved direct images of circumstellar disks reveal their internal density and temperature structure, and show how disks are affected by their ambient environment. Since circumstellar disks are both a product and a mediator of the star formation process, as well as the progenitors of planetary systems, it is clear that a fuller understanding of the evolution of circumstellar disks will play a key role in our understanding of these central topics.

Young circumstellar disks have been directly imaged at optical, near-, and mid-infrared wavelengths in a number of nearby star-forming regions, including Taurus-Auriga (Burrows et al. 1996; Koresko 1998; Stapelfeldt et al. 1998; Krist et al. 1999; Padgett et al. 1999), as well as closer to home (Koerner et al. 1998; Jayawardhana et al. 1998; Schneider et al. 1999). The largest sample of young circumstellar disks imaged to date is in the Orion Nebula. HST optical emission-line surveys of the region (O'Dell et al. 1993; O'Dell & Wen 1994, Bally et al. 1995; O'Dell & Wong 1996; Bally et al. 2000) have shown that many of the approximately 2000 young ( $\sim 1 \text{ Myr}$  old) stars of the Trapezium Cluster are either surrounded by compact ionized nebulae, called 'proplyds', which are interpreted as circumstellar disks externally ionized by the central O and B stars, or are surrounded by dark silhouettes, which are disks seen in projection against the bright H II region or within the proplyds (O'Dell & Wen 1994; McCaughrean & O'Dell 1996; Figure 5-7).



**Figure 5-7. Young Circumstellar Disks in Orion**

Note: A collection of young circumstellar disks seen as silhouettes against the bright background emission of the Orion Nebula H II region, imaged using the HST. All of these sources are thought to be associated with young (0.5 to 2 Myr) low-mass ( $0.3$  to  $1.5M_{\odot}$ ) members of the Trapezium Cluster. In several cases, where the disk is oriented close to edge-on and the central star cannot be seen directly, its presence is betrayed by polar reflection nebulae. A number of the disks are also seen to be embedded in ionized tadpole-shaped nebulae (‘proplyds’), which are created as the disk is heated and ablated by the central massive stars in the cluster. The square panels are  $1.67 \times 1.67$  arcsec or  $750 \times 750$  AU in size; the larger panels are proportional. The disks range from 50 to 500 AU in radius. The images were either made through the  $H\alpha$  filter of WFPC2 or the [O III] filter of STIS. Data from McCaughrean & O’Dell (1996), Bally, O’Dell, & McCaughrean (2000).

The approximately 50 silhouette disks that have been observed have diameters ranging from the HST resolution limit of 0.1 arcsec up to 2 arcsec (50 to 1000 AU). The disks are truncated at the outer edge, due either to internal evolution or external effects such as photoionization or star-disk interactions in the dense cluster environment (McCaughrean & O’Dell 1996). The distribution of disk sizes shows that disks inside ionized proplyds tend to be a little smaller than those of the pure (non-ionized) silhouettes (Rodmann 2002), suggesting that they are being rapidly eroded by the O and B stars (Johnstone et al. 1998).

Can these disks form planets before being destroyed? One way of answering this question is to look for evidence of growth in the dust grain population, which would indicate that planetesimal formation is already underway (Beckwith et al. 2000). By comparing the diameter of a silhouette disk at UV, optical, and near-infrared wavelengths, an estimate can be made of the dominant particle size in its outer reaches. HST and ground-based adaptive-optics studies to date have proven ambiguous, with suggestions of ISM-like grains on one hand and much larger  $5 \mu\text{m}$  grains on the other (cf. McCaughrean et al. 1998; Throop et al. 2001; Shuping et al. 2003).

However, due to the relatively limited spatial resolution of the HST at the critical near-infrared wavelengths, it has not been possible to attempt these studies for any but the single largest disk.

### **5.5.2 Key JWST Observations**

JWST must provide high spatial resolution imaging of a substantially larger sample of silhouette disks in selected near-infrared wavelengths where the background H II region is particularly bright. These observations must cover a broad wavelength range to yield maximum leverage with respect to the dust extinction. Most critical are the 1.87  $\mu\text{m}$  Pa $\alpha$  line, inaccessible from the ground, and the 4.05  $\mu\text{m}$  Br $\alpha$  line, where adequate sensitivity is hard to achieve from the ground.

Narrow-band ( $R = 100$ ) diffraction-limited JWST imaging is required to yield angular resolutions of 0.07 and 0.15 arcsec in the Pa $\alpha$  and Br $\alpha$  lines, respectively, corresponding to 30 and 70 AU, respectively, at Orion. These resolutions imply that for the  $>10$  silhouette disks with diameters of 200 AU (0.4 arcsec) or greater, it will be possible to measure the outer disk radial profiles and in combination with similar resolution HST images in the H $\alpha$ , [O III], and [O II] lines, a statistical assessment of the grain sizes in their outer reaches can be made. An important goal is to understand whether grain growth is inhibited or promoted if a disk is embedded in an ionized proplyd.

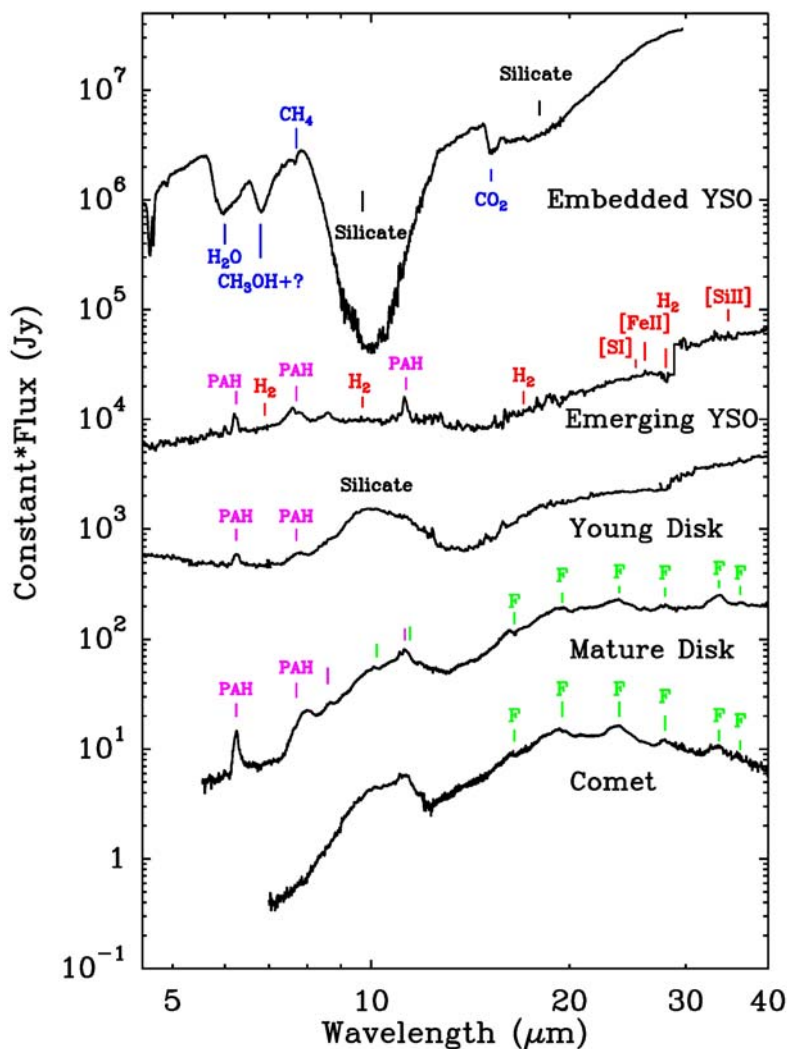
Extrapolating from measured H $\alpha$  fluxes around the Orion silhouette disks and the known foreground dust extinction, predicted Pa $\alpha$  line fluxes in the outer parts of the nebula are  $\sim 10^{-16}$   $\text{Wm}^{-2}$  arcsec $^{-2}$  or  $10^{-19}$   $\text{Wm}^{-2}$  pixel $^{-1}$ , assuming 0.0317 arcsec pixels. In the disk centers, the fluxes will be roughly 10 times fainter, about  $10^{-20}$   $\text{Wm}^{-2}$  pixel $^{-1}$ , and accurate mapping of the disk profiles requires imaging at S/N=100 to this flux level through a 1% narrow-band filter in the near infrared.

A clean, stable, and well-characterized PSF is also required to reduce the ‘PSF-blurring’ that ultimately limits such studies (McCaughrean & O’Dell 1996). In the more general case where the disk is not seen edge-on, the observations must be taken using a coronagraphic spot to occult the central star, which are relatively bright in the near-infrared.

## **5.6 WHAT ARE THE LIFE CYCLES OF GAS AND DUST?**

### **5.6.1 Scientific Rationale**

Generations of both low- and high-mass stars have converted primordial hydrogen and helium into successively heavier elements, including carbon, oxygen, and nitrogen, elements familiar to us as part of the cycle of life. Returned to the interstellar medium via winds and supernovae, they may eventually be incorporated into molecular clouds and later into new stars and their circumstellar disks. A major goal of astrochemistry and astrobiology is to trace the life cycle of gas and dust from pre-stellar cores to planetary systems (van Dishoeck & Blake 1998; Waters 2000; Ehrenfreund & Charnley 2000; see Figure 5-8).



**Figure 5-8. Mid-Infrared Spectra of Young Stars and Circumstellar Disks**

Note: A series of ISO SWS mid-infrared spectra of young stars and circumstellar disks at different stages in their evolution (Gibb et al. 2000; Van Den Ancker et al. 2000; Malfait et al. 1998). From top to bottom, in a rough evolutionary sequence, the spectra change from being dominated by solid state absorption features and shocked gas emission lines, to PAH features and photo-dissociation region lines, to amorphous and crystalline silicates with H I recombination lines. An ISO SWS spectrum of Comet Hale-Bopp is shown for comparison (Crovisier et al. 1997).

In cold quiescent cloud cores, gas-phase molecules stick to dust grains, forming icy mantles on silicate cores. This freeze-out is predicted to be very efficient and it is a paradox that gas-phase molecules are seen in clouds at all. Recent infrared and millimeter observations yield indirect evidence that there is substantial depletion, with up to 90% of the heavy elements frozen out onto grains (Lada et al. 1999; Kramer et al. 1999).

When the heavy elements are frozen out, the chemical composition is initially modified via grain-surface reactions. Then, as a central protostar develops and heats up its envelope, the ices evaporate back to the gas phase and a rich mixture of organic compounds develops. Later, disk accretion events lead to FU Ori type outbursts, producing heating and UV flux. As the envelope material is processed in this way, its thermal and irradiation history is imprinted through irreversible changes in band profiles in spectra, via features including the solid CO<sub>2</sub> bending mode at 15 μm, the OCN<sup>-</sup> band at 4.62 μm, and the unidentified 6.85 μm feature (Ehrenfreund et al. 1998; Gerakines et al. 1999; Schutte & Greenberg 1997; Schutte & Khanna 2003).

Later, gas and dust from the envelope is incorporated into a circumstellar disk surrounding the young star, where it is further processed by UV radiation, X-rays, and thermal processes. In the cold disk midplane, molecules will freeze out again on the dust, while gas-phase molecules such as H<sub>2</sub>, CO, CH<sub>4</sub>, and C<sub>2</sub>H<sub>2</sub> are expected in the warmer regions. Disk spectra should be dominated by features from PAHs, ices, and silicates, both in crystalline and amorphous form; these particles are the building blocks of planets and, as they include water ices and organic materials, potentially of life. By studying the astrochemical evolution in detail, it is possible to trace the formation history of planetesimals and other solid bodies, their composition, their processing, and their possible future evolution.

Gas plays an important role in the formation of giant planets. Warm (~ 100 K) molecular gas located 1 to 50 AU from the central star in nearby debris disks can be traced using the pure rotational lines of H<sub>2</sub>: J = 5-3 S(3) at 9.662 μm, J = 4-2 S(2) at 12.278 μm, and J = 3-1 S(1) at 17.035 μm. Simultaneous measurements of these lines would provide a profile of the temperature and mass of the gas as a function of radius. Dust-to-gas mass ratios can be derived by comparing the H<sub>2</sub> results with dust measurements. These H<sub>2</sub> measurements in different-age disks are sensitive to small masses of molecular gas, and will constrain the age at which gas giants can form. The measurements can also show how gas is cleared from the disks, and on what time scale.

To trace the evolution of these various tracers as they cycle back and forth between the gas and solid phases, we require high dynamic range, medium resolution near- and mid-infrared spectroscopy of individual sources in star-forming regions spanning a range of ages and environments. Spectra from the ISO SWS have shown signs of material evolution in the disks around more massive stars (Malfait et al. 1998; Meeus et al. 2001; Bouwman et al. 2001) and the H<sub>2</sub> S(1) lines at 17 and 28 μm have been tentatively detected by ISO in two nearby debris disks (Thi et al. 2001). JWST will allow this work to be extended to a much more extensive sample, including proto-solar-type stars.

### **5.6.2 Key JWST Observations**

Spectroscopy of individual sources covering 3 to 27 μm (and up to 28.3 μm if possible) is required at spectral resolutions of R = 1000 to 3000, depending on the source and the tracers under study, with smoothing to R ~ 100 (with attendant gain in signal-to-noise) acceptable for detecting very broad features at the faintest possible levels.

Tracers may be seen in absorption against bright continuum sources. For example, in order to map the freeze-out processes in a dark cloud core, solid-state absorption bands are seen in the spectra of background field sources. To detect ices frozen on grains in the midplanes of circumstellar disks, the central young star itself can be used as the continuum source in edge-on geometries. Conversely, the tracers may appear in emission, as in the case of gas-phase molecules, PAHs, ices, and silicates in both amorphous and crystalline forms from warm material in the inner regions of disks.

To study the mineralogy of the dust in nearby debris disks and of their precursors around young stars, a  $5\sigma$  limiting sensitivity of  $10 \mu\text{Jy}$  is required at  $25 \mu\text{m}$  with a smoothed spectral resolution of  $R \sim 100$  and with azimuthal averaging around the disk. This observation will detect a Vega-like system at a distance of  $40 \text{ pc}$ .

To study the evolution of gas in disks, a line flux sensitivity of  $2.6 \times 10^{-20} \text{ Wm}^{-2} \text{ arcsecond}^{-2}$  at  $17 \mu\text{m}$  would enable the detection of about half an Earth mass of  $100 \text{ K}$  molecular gas in a debris disk at  $30 \text{ pc}$ . It is also desirable that the JWST spectroscopy be sensitive beyond  $28.22 \mu\text{m}$  in order to include the lowest  $\text{H}_2$  transition, which will be the brightest line and sensitive to the coldest gas. High spectral resolution is essential at this wavelength to obtain adequate contrast of the emission line relative to the continuum.

Finally, integral field capability in the mid-infrared is required to allow accurate mapping of the emission and absorption components in a given system. A field of at least  $3 \text{ arcsec}$  is required to match the scales in typical nearby debris disks and more massive disks around young stars in nearby star-forming regions. A high-contrast long slit is required for near-infrared spectroscopy, to allow the faint disk material to be well separated from the potentially bright central source and to map out spatial structures in the disks.

## 5.7 SUMMARY

**Table 5-2. JWST Measurements for the Birth of Stars Theme**

<b>Program</b>	<b>Primary Instruments</b>	<b>Magnitude or Flux</b>	<b>Target Density</b>	<b>Field</b>
Cloud Collapse	NIRCam	9nJy at 2 $\mu$ m	Individual	2–4 arcmin
	MIRI (imaging)	1 $\mu$ Jyarcsec <sup>-2</sup> at 7 $\mu$ m	Individual	1–2 arcmin
Evolution of Protostars	MIRI (imaging)	0.1 $\mu$ Jy at 6 $\mu$ m, 1 $\mu$ Jy at 15 $\mu$ m	1 arcmin <sup>-2</sup>	<1 arcmin
	MIRI (spectra)	7x10 <sup>-22</sup> Wm <sup>-2</sup> at 15 $\mu$ m	1 arcmin <sup>-2</sup>	3 arcsec
Massive Stars	NIRCam	2 $\mu$ Jy at 3.8 $\mu$ m, 10 $\mu$ Jy at 4.8 $\mu$ m	1 arcmin <sup>-2</sup>	<1 arcmin
	NIRCam	55nJy at 3.5 $\mu$ m, 2.3 $\mu$ Jy at 4.8 $\mu$ m	10 arcmin <sup>-2</sup>	2–4 arcmin
	NIRSpec		10 arcmin <sup>-2</sup>	2–4 arcmin
	MIRI (imaging)	34 $\mu$ Jy at 20 $\mu$ m	10 arcmin <sup>-2</sup>	1–2 arcmin
IMF	NIRCam	2.9nJy at 2.2 $\mu$ m	100 arcmin <sup>-2</sup>	2–4 arcmin
	NIRSpec	290nJy at 2.2 $\mu$ m	100 arcmin <sup>-2</sup>	2–4 arcmin
	MIRI		10 arcmin <sup>-2</sup>	1–2 arcmin
Protoplanetary Systems	NIRCam	10 <sup>-20</sup> Wm <sup>-2</sup> arcsec <sup>-2</sup> at 1.87 $\mu$ m	10 arcmin <sup>-2</sup>	2–4 arcmin
Astrochemistry	MIRI (spectra)	2.6x10 <sup>-20</sup> Wm <sup>-2</sup> arcsec <sup>-2</sup> at 17 $\mu$ m	Individual	3 arcsec

**Table 5-3. Required Capabilities for the Birth of Stars Theme**

	<b>Cloud Collapse</b>	<b>Protostar Evolution</b>	<b>Massive Stars</b>	<b>IMF</b>	<b>Proto-planetary Systems</b>	<b>Gas and Dust</b>
	5.1	5.2	5.3	5.4	5.5	5.6
<b>NIRCam</b>						
Broad-band	X		X	X		
Sensitivity	X		X	X	X	
FOV	X		X	X	X	
PSF						
Dynamic Range			X	X	X	
Coronagraphy					X	
<b>NIRSpec</b>						
Multi-obj FOV			X	X		
R = 100			X			
R = 1000			X	X	X	
R = 3000						X
Sensitivity			X	X		
High Contrast				X	X	X
<b>MIRI</b>						
Broad-band	X	X	X	X	X	
Sensitivity	X	X	X			
FOV	X		X	X		
PSF						
R = 3000 IFU		X			X	X
Coronagraphy					X	
<b>FGS-TF</b>						
Short $\lambda$ TF			X	X	X	
Long $\lambda$ TF			X	X	X	
Sensitivity						
FOV						
Coronagraphy					X	
<b>Observatory</b>						
PSF $\lambda < 1 \mu\text{m}$				X	X	
Stable PSF					X	
Stable Image				X		
CVZ						
Moving Targets						

## **6.0 PLANETARY SYSTEMS AND THE ORIGINS OF LIFE**

The key objective of the Planetary Systems and the Origins of Life theme is to determine the physical and chemical properties of planetary systems including our own, and investigate the potential for the origins of life in those systems.

To trace the origins of the Earth and life in the Universe, we need to study of planet formation and evolution, including the structure and evolution of circumstellar material. The search for evidence of life in our solar system and beyond is fundamental to the understanding of our place in the cosmos. JWST observations of objects in our own Solar System and planetary systems around other stars will provide data crucial for understanding the origin of planetary systems, the origin of free-floating brown dwarfs and planets, and the potential for stable habitable regions around other stars (Seager & Lunine 2004). Both the near- and mid-infrared capabilities of JWST are required to conduct the investigations described in this theme.

### **6.1 ORIGINS OF PLANETARY SYSTEMS**

#### **6.1.1 Questions**

##### **6.1.1.1 How Do Planets and Brown Dwarfs Form?**

The formation of multiple objects is a common outcome of star formation, including binary or triple star systems, a central star orbited by brown dwarfs and/or planets, or simply a remnant disk of particulates. Brown dwarfs and giant planets might arise from two different formation mechanisms. Brown dwarfs may represent direct collapse of gas, from a molecular cloud clump or from a disk of material, whereas giant planets could be defined as those objects forming from a two-stage process in which growth of a rock-ice core triggers the rapid accretion of gas. In this model, brown dwarf companions are the low-mass tail of binary star formation, while giant planets are the high-mass end of a process that also makes Earths and Neptunes, and the two processes may produce distinct initial mass functions. Even if there is an overlap in the masses generated by the two processes, they may be distinguished by the metallicities of the objects generated relative to the parent star. Giant planet formation by two-stage accretion increases the metallicity of the planets by a factor of several relative to the parent star; direct collapse does not need to do so.

Formation of giant planets is a signpost, detectable with JWST, of a process that may also generate terrestrial planets. In contrast, direct collapse formation of brown dwarfs may signal systems in which terrestrial planet formation is rare or impossible, because of the required disk masses, angular momenta and subsequent disk evolution.

##### **6.1.1.2 How Common are Giant Planets and What is their Distribution of Orbits?**

Planets of Saturnian mass or larger exist in orbits detectable by the current radial velocity surveys around 5% of F, G, and K type stars in the solar neighborhood. Extrapolating to larger semi-major axes, and guided by our own dynamically crowded outer solar system (the region from 5 to 30 AU), one can infer that about 10% of mature F, G, and K type stars should possess at least one giant planet. Theoretical studies of migration of giant planets through interactions

with the gaseous disk, or with massive remnant particulate disks, suggest that as many as 80% of low mass stars generate giant planets during their pre-main sequence phase. Many or most of these giant planets are lost through inward migration and merging with the central star, or by ejection. We don't know how many nearby suns possess giant planets in orbits too large for detection by the radial velocity approach, or with periods too long for planned space-borne astrometric surveys. Giant planet formation may be a process favored in the colder outer regions of protoplanetary disks, where water ice exists. Alternatively, giant planets could be formed over a wide range of semi-major axes, even in the warm inner parts of disks.

### **6.1.1.3 How Do Giant Planets Affect the Formation of Terrestrial Planets?**

Given a particular distribution of orbital semi-major axes of extrasolar giant planets, one can assess the dynamical consequences of these bodies on both the stability of the orbits of putative terrestrial planets in the habitable zones of low mass stars, and on the delivery of water to the habitable zone from colder regions. Jupiter and Saturn, forming within a few million years of the birth of our protoplanetary disk, both ejected remnant planetesimals from their own orbits, and increased the inclinations and eccentricities of planetesimals, so that they reached the region where the Earth and other terrestrial planets formed later (roughly 50 - 100 million years based on radioisotopic dating). If this is a general characteristic of giant planet formation, then in many disks, both processes may lead to delivery of colder, water- and organic-rich planetesimals to the inner planet forming region, and accelerate the growth of terrestrial planets. On the other hand, radial migration of giant planets inward through the zone of formation of the terrestrial planets would sweep material there into the parent star, along with the migrating planet. It is possible that giant planets do both in a given system: an early generation of rapidly forming Jovian mass objects sweeping inward through the disk in the first million years or so, followed by formation of a second generation of giant planets that triggers the formation of terrestrial planets and delivers volatiles.

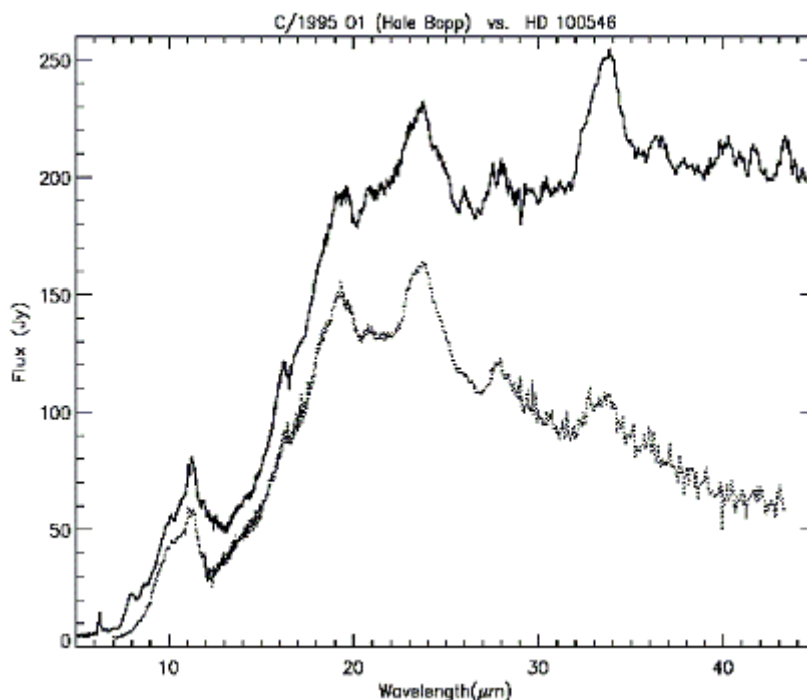
### **6.1.1.4 What Comparisons, Direct or Indirect, can be made between our Solar System and Circumstellar Disks (Forming Solar Systems) and Remnant Disks?**

Detailed observations of disks around other stars, both during and after planet formation and disk clearing, provides a global view of the distribution of major condensables and the effect of planets on the distribution of the dusty, small body material. A detailed discussion of the observations of disks is given in section 5, but disks are both the product of star formation and the initial step in the formation of planets.

The remnant of the circumstellar disk that formed our Solar System is observable today as the smaller planets, moons, asteroids and comets, along with the interplanetary gas. We have samples of this material in meteorites and interplanetary dust particles. Studies of samples provide accurate chemical and (radioactive isotope) age determinations for events, but are difficult to relate to specific locations in our own protoplanetary disk because of dynamical stirring of material. Further, the primitive material in our meteorite collection does not seem to correspond in detail to the abundances in the Earth's interior, suggesting that much of the "primitive inventory" is missing.

Spectroscopy and photometry of small Solar System bodies, particularly comets, at wavelengths and sensitivities unavailable from the ground can identify isotopic ratios and molecular and elemental abundances. These can be compared with remnant and planet-forming disks, providing direct measurements of the smaller components of circumstellar disk formation. For example, IR and radio spectroscopic observations of cometary parent molecules, the species that sublime directly from the nucleus, suggest a strong similarity in chemical composition between cometary nuclei and the icy dust in protostellar environments (Irvine et al. 2000). In both cases, composition is dominated by H<sub>2</sub>O ice, CO and CO<sub>2</sub> are usually the next most abundant. The spectra also show CH<sub>3</sub>OH, H<sub>2</sub>CO, and CH<sub>4</sub>. In addition, C<sub>2</sub>H<sub>6</sub> and C<sub>2</sub>H<sub>2</sub> have also been observed in comets, but they have not yet been detected in icy ISM grains. High concentrations of deuterated species are another indication that comets may retain pristine interstellar matter (Meier and Owen 1999).

Since the mid 1980's apparition of Comet Halley, it has been known that the silicate mineralogy in cometary dust is similar to that in circumstellar disks (Spinrad 1986). This was beautifully illustrated with ISO observations of C/Hale-Bopp, which showed that the mid-infrared reflectivity of the comet's dust is strikingly similar to that in a protoplanetary disk surrounding the young star HD 100546 (Figure 6-1; Malfait et al. 1998). In particular, strong emission features of carbon- and oxygen-rich dust are seen in both spectra, with the most prominent being attributed to crystalline silicates. This observation confirms the key role of comets in understanding the chemical nature of dust in debris disks.



**Figure 6-1. Comets and Circumstellar Disks**

Note: An ISO spectrum of Herbig Ae/Be star HD 100546 (upper curve) shows emission features in the star's circumstellar disk as predicted if the disk is heated by radiation from the central source. The particle composition in the circumstellar disk appears to be remarkably similar to that of comet Hale-

Bopp (lower curve) (Malfait et al. 1998). JWST will only reach the shorter wavelengths depicted here, but the similarities are most prominent at those wavelengths.

## 6.1.2 Observations

### 6.1.2.1 Extra-Solar Giant Planets

JWST must provide spectra of isolated giant planets and brown dwarfs as well as filter photometry of giant planets and brown dwarfs in bound systems. JWST must detect giant planets in large orbits unreachable by other techniques, and provide spectroscopic and photometric information on their thermal properties and constrain mass and age. The statistics of the metallicities of isolated brown dwarfs and giant planets will be compared with those of similar-age low mass stars (F, G, K) to determine whether the isolated population is biased toward higher metallicities, and hence formed in a process wherein large amounts of metals were incorporated during formation.

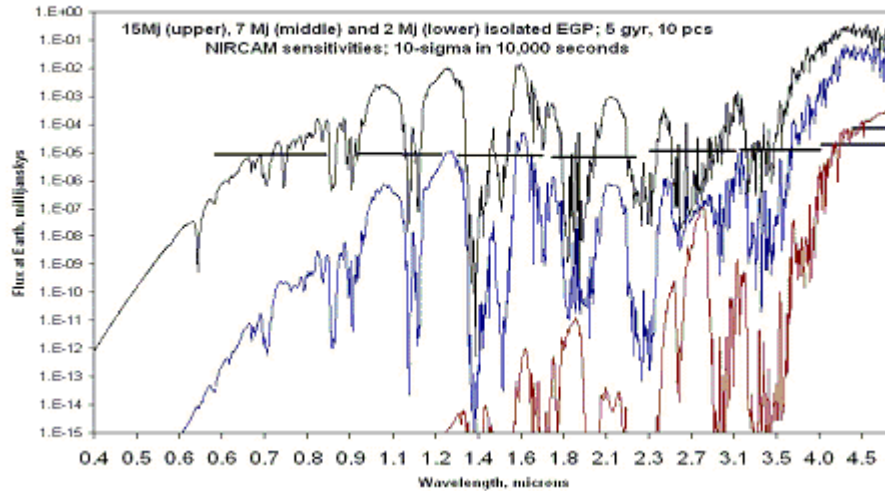
JWST must make coronagraphic images at 2.7  $\mu\text{m}$  and 4.44  $\mu\text{m}$  to find nearby mature Jovian companions and more distant young Jovian companions. Table 6-1 illustrates that the 2.7- $\mu\text{m}$  to 4.44- $\mu\text{m}$  flux ratio is extremely sensitive to the temperature of planetary companions, and provides a rapid diagnosis of an object's mass, along with a rough estimate of the orbital parameters. One modeling uncertainty that could dramatically affect this ratio is the existence of clouds of various condensable species, ranging from silicates to water ice according to the background temperature-pressure profile. Extensive additional modeling work will be required to understand the impact, which in general will be to render the flux ratio less sensitive to temperature.

**Table 6-1. Model Fluxes of Bound Giant Planets**

	<b>2.7 <math>\mu\text{m}</math> flux (millijanskys)</b>	<b>4.4 <math>\mu\text{m}</math> flux (millijanskys)</b>	<b>Ratio of Fluxes (2.7 <math>\mu\text{m}</math>/4.4 <math>\mu\text{m}</math>)</b>
55 Cnc b	0.188	1.18	0.159
55 Cnc d	$1.55 \times 10^{-8}$	$5.38 \times 10^{-3}$	$2.88 \times 10^{-6}$

NOTE: Values show the ratio of 2.7  $\mu\text{m}$  to 4.4  $\mu\text{m}$  fluxes for models of Bound Giant Planets (*REFERENCE TBD*).

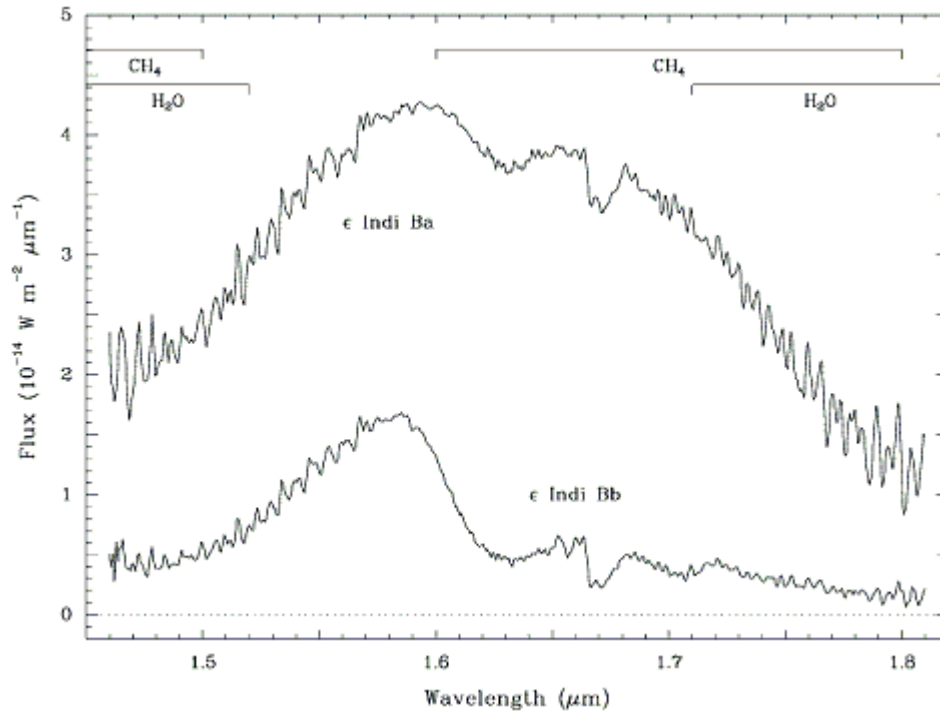
JWST must make coronagraphic images at  $R \sim 100$  spectral resolution to determine effective temperature and radius. The spectrum of an extrasolar planet reveals not only temperature and composition but also to gravity and the presence of clouds as well. The relative differences between absorption features will provide diagnostic information on all of these parameters (Burrows et al. 2003). With the mass obtained from the temperature and age of the companion star, we can then infer the radius. Again, this determination will be affected by the presence of clouds, and the clouds in turn will be revealed by the relative changes in absorption features compared with cloud-free giant planets. JWST coronagraphy will also make it possible to study bound giant planets in Jovian-type orbits around nearby stars as well.



**Figure 6-2. Detectability of Extra-Solar Giant Planets and Brown Dwarfs**

Note: Model spectra (Sudarsky et al., 2003) of a 15, 7 and 2 Jupiter mass extrasolar giant planet, free-floating at 10 parsecs from Earth, age 5 billion years, with JWST broad-band filter sensitivity plotted. The 5- $\mu$ m window makes the smallest mass planet detectable with JWST, while the 15 Jupiter mass object (transitional to the T dwarfs) is bright enough to be studied spectroscopically with narrow-band imaging or spectroscopy.

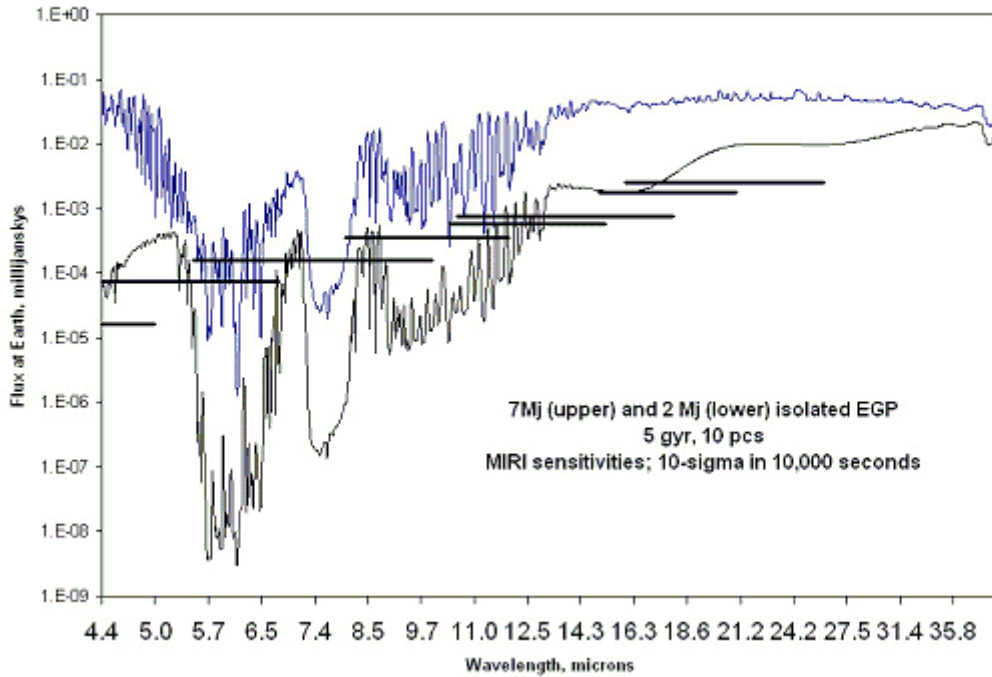
JWST must make  $R \sim 1000$  spectroscopy at 1 to 5  $\mu$ m of isolated or widely separated giant planets or sub-brown dwarfs (Figure 6-2). With this spectral resolution, it is possible to unambiguously determine basic physical parameters of a brown dwarf atmosphere such as gravity, composition, the temperature profile (gradient, inversions) and the effect of clouds. Much of the information that may be contained in key features such as methane absorption feature is poorly understood because few T dwarf spectra are available, and of these telluric absorption renders analysis difficult. JWST must provide much higher signal-to-noise for T-dwarfs down through warm Jupiters with full coverage through the regions obscured by telluric absorption (Figure 6-3).



**Figure 6-3. Spectra of the Closest T-dwarfs**

Note: In these data, obtained by McCaughrean et al. (2004) the closest brown dwarfs to Earth,  $\epsilon$  Indi B a and b (originally thought to be one object), show distinctly different spectra. The differences between the two low-resolution spectra relate primarily to effective temperature (1250 K vs. 850 K) through a range of potential effects including molecular composition, presence of grains, and the effect of physical temperature on the band shapes themselves.

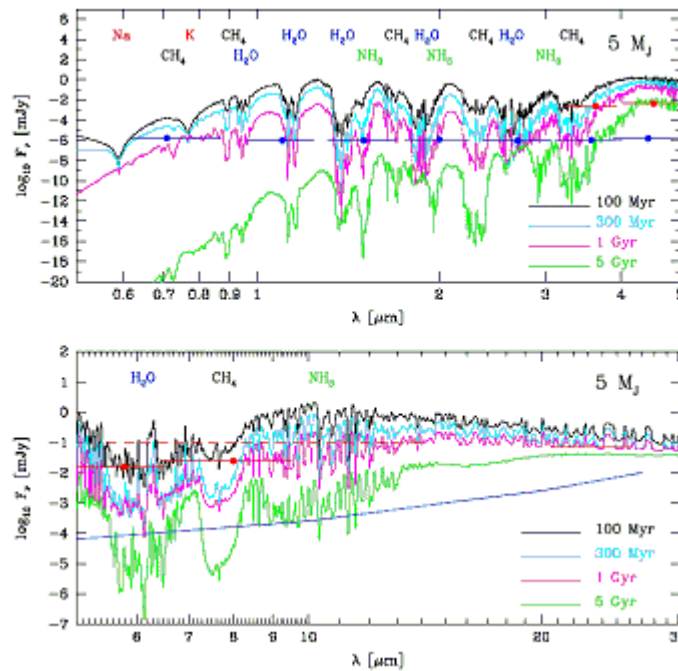
JWST must have the coronagraphic sensitivity to detect a Jupiter analog out to  $\sim 30$  pc. As shown in Figure 6-4, this would be a broadband detection, taking advantage of both the long-wavelength end of the  $5 \mu\text{m}$  excess emission as well as the more typically blackbody emission at much longer wavelengths.



**Figure 6-4. Detectability of a Jovian-sized Exoplanet**

Note: Model spectra of isolated giant planets of 7 and 2  $M_{JUP}$ , 10 parsecs from Earth, five billion years old star from (Sudarsky et al. 2003), compared with JWST mid-infrared imaging sensitivities. The 2  $M_{JUP}$  object is detectable in several wavelengths; the 7  $M_{JUP}$  object is bright enough that spectra at  $R \sim 3000$  can be collected, enabling atmospheric structure and composition to be inferred.

JWST must make spectra at 5 to 27  $\mu m$  of  $>1 M_{JUP}$  objects within 10 parsecs. This broad wavelength range is near, but generally longward, of the Planck function peak (Figure 6-4). JWST spectra of the nearest systems can provide detailed insights into the nature of giant planet atmospheres, including abundances of ammonia and methane that are key indicators of atmospheric mixing and temperature profiles. Figure 6-5 shows that many spectral features are available for filter and spectroscopic study in the near and mid infrared of objects with masses between that of Jupiter and brown dwarfs.



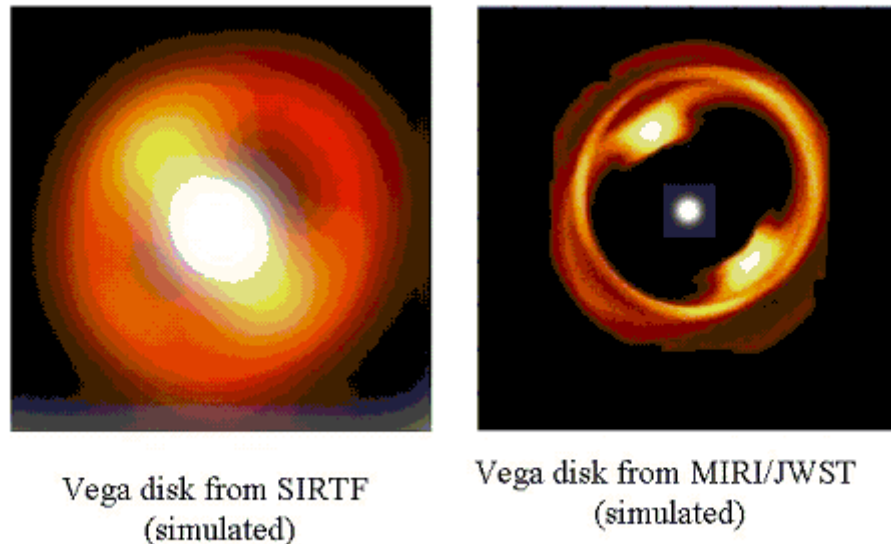
**Figure 6-5. Guide to Spectral Features in Extrasolar Giant Planets**

Note: A 5  $M_{JUP}$  planet of varying ages is shown, with major species responsible for the spectral features labeled. Rough Spitzer (red) and JWST (blue) sensitivities are shown. Detection of these features in the near- and mid-infrared samples different regions of the atmosphere, allowing abundance to be separated from temperature profile and the effects of clouds (Burrows et al 2003).

### 6.1.2.2 Circumstellar Disks

JWST must resolve the details of nearby disk structures to determine the dynamical effects of planets and through spectra, the radial and even azimuthal distribution of major elements and their molecular or mineralogical carriers. JWST will also be able to constrain the radial temperature distribution on the surface, and within, disks.

Models of debris disk evolution in the presence of planets can be tested to distances of 40 to 50 pc. The spatial resolution available with Spitzer on the Vega disk (8 parsecs away), sufficient to determine first order effects of a giant planet on disk morphology, can be achieved by JWST in the mid infrared on disks some 40 parsecs from the Earth. On Vega itself, details on the perturbations in disk structure on 10 AU scales are discernable (Figure 6-6).



**Figure 6-6. Seeing the Effect of Planets on Dust Disks**

Note: Simulation of the effect of a planet on the Vega dust disk, using the dynamical model of Wilner et al. (2001). On the left is the model structure if seen with Spitzer's resolution at Vega's actual distance from Earth, or equivalently, with JWST's resolution if Vega were at 40 parsecs. On the right is a simulated JWST view of the model for Vega's actual distance from Earth.

### 6.1.2.3 Comets

Each comet is a unique remnant from the Solar System formation, its current composition and physical properties provide a constraint on the conditions in the solar nebula 4.6 billion years ago. Comets were the building blocks of the giant planets' cores. Low resolution infrared spectroscopy of cometary dust will uncover mineralogical signatures, which can be compared with those seen in protostellar and planetary debris disks around nearby young stars and solar analogs, and potentially reveal the isotopic ratios of some major elements.

Observations of comets with JWST will enable investigations of the chemical composition of cometary ice and dust with unprecedented sensitivity. Near- and mid-IR spectroscopy of cometary comae can be used to measure abundances of H<sub>2</sub>O, CO, CO<sub>2</sub>, and CH<sub>3</sub>OH in even relatively faint comets. The near-IR spectrometer's resolution should be high enough to measure the ratio of ortho and para H<sub>2</sub>O separately (Crovisier et al. 1997), possibly providing a probe of the comet's formation temperature (Mumma et al. 1987). Likewise, mid-IR spectroscopy can probe the mineralogy of cometary dust grains in virtually any comet passing through the inner Solar System. Finally, JWST's ability to image cometary nuclei at both mid and near-infrared wavelengths with high spatial resolution and sensitivity will allow high accuracy measurements

of sizes and albedos for essentially every observable comet, including those of both dynamical classes (short- and long-period). The results from cometary programs can be combined with those from programs investigating circumstellar disks and star formation regions to build a complete picture of planetary system formation and evolution.

JWST will measure the CO<sub>2</sub> abundance in any comet that comes within ~3 AU of the Earth and Sun and is still within the solar elongation angle constraints of the observatory (i.e., elongation angles between 85 and 135°). Depending on the circumstances, JWST can measure CO<sub>2</sub> emission in either the  $\nu_3$  band near 4.3  $\mu\text{m}$  or the  $\nu_2$  band near 15  $\mu\text{m}$ , both of which are exceptionally strong. Only JWST can effectively probe CO<sub>2</sub>; strong absorption in the terrestrial atmosphere prevents ground-based infrared observations, and the molecule does not have a permanent electric dipole moment, and so it does not emit in the radio.

Although we now have several means of measuring H<sub>2</sub>O and CO directly in comets from ground-based observatories, JWST will do so with much higher sensitivity. From the ground, H<sub>2</sub>O can only be observed in intrinsically faint intercombination bands; atmospheric absorption prevents observations in the much stronger fundamental bands. By observing H<sub>2</sub>O in the fundamental bands with a resolving power of ~100, JWST will not only detect much lower activity levels, but will also measure the ortho-to-para ratio (OPR) in many comets. Since the OPR is not modified in cometary comae, nor is its value expected to change from the warming of a nucleus during its short-duration trek through the inner Solar System, the OPR probably reflects the comet's temperature at the time of its formation. Cosmic rays may alter the OPR in the outer layers of comets in the Oort cloud; in that case the OPR may provide an unambiguous method for identifying the comet as dynamically new.

Measuring the HDO/H<sub>2</sub>O value in long- and short-period comets (or the latter's larger cousins in the Kuiper Belt) will supplement the D/H determinations made to date. This is key to understanding the relationship between water in comets and that on the Earth and Mars. JWST's high sensitivity spectral observations in the near- and mid-IR will enable this measurement in certain comets and the largest KBOs.

Similarly, the CO abundance can be probed directly from ground-based facilities, but JWST would again provide at least an order of magnitude improvement in sensitivity. Thermal emission from a warm telescope and the Earth's atmosphere limits the capabilities of even large ground-based infrared facilities. Radio observations of CO are much less sensitive than IR observations because of CO's small electric dipole moment.

Spitzer will have neither the spectral resolution nor the wavelength range required for gas-phase cometary studies. In addition, JWST's spatial resolution will permit great advances over Spitzer for the study of solid phase materials in both comets and circumstellar material. By the time JWST is launched, several comet missions may have occurred, and the results from those missions will undoubtedly influence the observing program.

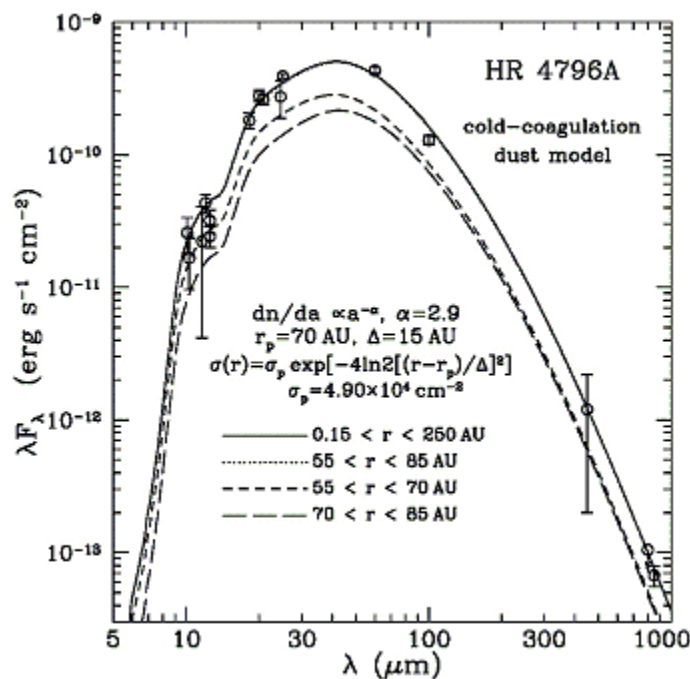
### 6.1.3 Observatory Capabilities

#### 6.1.3.1 Extra-Solar Giant Planets

JWST narrow-band imaging coronagraphy will discover extrasolar planets obtain preliminary information on their temperature and, using the semi-major axis and the stellar age, their mass. Widely separated companions or isolated giant planets will be analyzed for metallicity indicators with JWST spectroscopy at  $R \sim 1000$ . Detailed information on the surface gravity and, through modeling, their mass and atmospheric temperature-pressure profiles can be obtained at spectral resolutions as coarse as  $R \sim 100$ , but more reliably from spectra with  $R \sim 1000$ .

#### 6.1.3.2 Circumstellar Disks

Spectroscopic resolution of 1000 in the near and mid-infrared are required to resolve the spectral signatures of key ices and silicates in the disks, and coronagraphic capability is necessary to block the light of the central star and observe planets perturbing the disk structures. Mid-infrared is of particular interest for the disk since the grain temperatures in the disks are generally low enough that the peak emission will be in the mid-infrared (Figure 6-7); dust that is warm enough to radiate at much warmer temperatures contributes little to the total infrared signature.



**Figure 6-7. Models and Observations of a Dusty Disk**

Note: Models of the spectral energy distribution the dusty disk around HR4796 are compared with observations. Emissions from different ranges of semi-major axis are plotted. Most of the action in terms of peak emission and interesting spectral features lies beyond  $5 \mu\text{m}$ , as is the case for most dust disks (Li and Lunine 2003).

### 6.1.3.3 Comets

To observe an appropriate sample of comets, JWST must be able to track an object that moves with respect to the background stars at a minimum rate of 0.030 arcsec/sec. Even with the Solar and anti-Solar exclusion angles, this capability allows tracking of long-period or hyperbolic comets such as Halley and Hale-Bopp to within about 2 AU of the sun, at which point they are very active. It also permits studies of short-period comets such as Wirtanen, Encke, and Hartley 2.

As most comets tend to be relatively bright, the exposure times will be short. A slight discrepancy between the true rate and the JWST tracking rate is not an issue. However, for distant cometary nuclei (and KBOs, see §7.2.3 below), the exposure times are longer. In that case, differences between the true and implemented rates can cause the object to drift across pixels, smearing the image and degrading the signal. The track rate of the object must be followed to an accuracy of a small fraction of a pixel in a typical imaging frame time, e.g. 0.005 arcsec in 1000 sec (for very faint and typically slow objects).

In virtually all cases for cometary (indeed most solar system) observations, the object motion during the typical exposure time is well approximated by a linear track. For example, this is the only mode that Spitzer is using for moving target tracking. JWST is not required to track a curved trajectory.

There are no special requirements on guide star availability, field of regard, length of track with a single guide star, or stability of pointing. It is recognized that some reduction of image quality may occur during moving object tracking.

Peaking up on a comet (or other moving target) will not differ from peaking up on a fixed target. In those cases, where one is measuring abundances in a more extended comae (or on a resolved planetary disk), similar techniques would be used as for an extended source such as a galaxy. There are no *a priori* roll-angle limitations.

Solid bodies (Kuiper Belt Objects) are not expected to have spectral lines, but warmer objects (comets) will have extremely rich spectra. With the sensitivity of JWST, there may be thousands of overlapping lines. JWST spectroscopy of comets requires  $R \sim 1000$ , or  $R \sim 3000$  for some isotopic ratios.

Cometary molecules have spectral lines across the entire JWST spectral range. Primary vibration-rotation lines are concentrated in the 1-5  $\mu\text{m}$  band. We also require observations in the 15  $\mu\text{m}$  window to study  $\text{CO}_2$ .

## 6.2 EARLY EVOLUTION OF PLANETARY SYSTEMS

### 6.2.1 Questions

#### 6.2.1.1 What is the Source of Water and Organics for Planets in Habitable Zones?

A preponderance of geochemical evidence suggests that Earth's water did not come from locally formed planetesimals at 1 AU. However, the source of water is uncertain. Asteroids are a

dynamically plausible source and could be isotopically consistent if chondrites are a typical sample of that region of the primordial asteroid belt that supplied water to the Earth. However, problems remain with this model, in particular, it is difficult for the Earth to retain the water during the high velocity impact of very large bodies that dynamical simulations indicate would have been the principle sources of water. Cometary HDO/H<sub>2</sub>O values measured in three long-period comets are twice that of Earth's ocean water. Also, the D/H ratio of short period comets and their presumed Kuiper Belt progenitors remains unknown. By measuring isotopic ratios in comets and larger Kuiper Belt bodies, JWST can solve this part of the puzzle, removing a major uncertainty in the source of water for our own planet.

Similarly, the source of the early abundance of Martian water is uncertain; as on Earth, it could be local, asteroidal, or cometary. The continued search for extant Martian water inventories is of relevance to this problem as well.

Comets remain a highly plausible source of Earth's organics, and the inventory of organics derived from high sensitivity infrared spectra will be of value in constraining this part of the story. JWST measurements of the composition and structure in protoplanetary disks around other stars will extend the quantification of the source of water and organics to putative habitable worlds around stars other than the Sun.

#### **6.2.1.2 How are Systems Cleared of Small Bodies?**

Studies of active gas-dust disks, and remnant disks from JWST, with Spitzer disk studies as a foundation, will better quantify how solid debris and gas is cleared from such disks. JWST will address the timing, the role of the planets, and the amount of remnant gas and dust during what would be the early history of planetary systems, following their formation. Measurement of the isotopic, elemental and molecular abundances in icy bodies in the outer Solar System, as well as in large bodies such as Titan, will provide a body of chemical data that will allow us to determine the relationship of these various bodies to a set of putative primitive reservoirs, constrained as well by the extrasolar disk observations.

Also, understanding the physical characteristics, including chemistry, of KBOs as a function of subclass (Plutino, Scattered Disk, Classical) provides important clues to the dynamical evolution of the Kuiper Belt. JWST will particularly important for mid-infrared measurements of a reasonably large sample of these objects.

### **6.2.2 Observations**

#### **6.2.2.1 Comets**

Comets, through collisions with the terrestrial planets, might represent a significant source of volatile material in the inner Solar System. A substantial portion of the Earth's oceans and organic material were probably provided by cometary bombardment, tying comets directly to the origin of life. JWST's spatial and spectral resolution will allow studies of gas phase processes in the inner comae of active comets, revealing their composition.

### 6.2.2.2 Kuiper Belt Objects

JWST must obtain near-IR and mid-IR measurements of the brightness of KBOs to separate size and albedo effects, thereby constraining physical size. Spectra will be obtained of the larger and closer KBOs to identify major molecules and isotopic ratios for the largest objects, including Triton, Pluto, Charon, Quaoar, and Varuna. In this context, Neptune's satellite Triton is a captured Kuiper Belt Object, as indicated by Triton's retrograde orbit and Neptune's dilapidated remnant regular satellite system.

JWST must obtain high resolution, near-infrared spectra of bodies hundreds of kilometers in size, typical of those detected in surveys of the Kuiper Belt, to determine the presence of various ices, including deuterated species and other isotopic bands, as shown in the simulation of Figure 6-8. Isotopic ratios, including  $^{13}\text{CO}/^{12}\text{CO}$ ,  $\text{C}^{18}\text{O}/\text{C}^{16}\text{O}$ ,  $^{13}\text{CH}_4/^{12}\text{CH}_4$ ,  $\text{CH}_3\text{D}/\text{CH}_4$ ,  $^{15}\text{NH}_3/^{14}\text{NH}_3$ ,  $\text{HDO}/\text{H}_2\text{O}$ , may be detectable on Pluto (the largest KBO) depending on their abundances. This information will reveal the relationship of comets and the bound water in carbonaceous chondrites to the ices on the surfaces of the largest Kuiper Belt Objects.

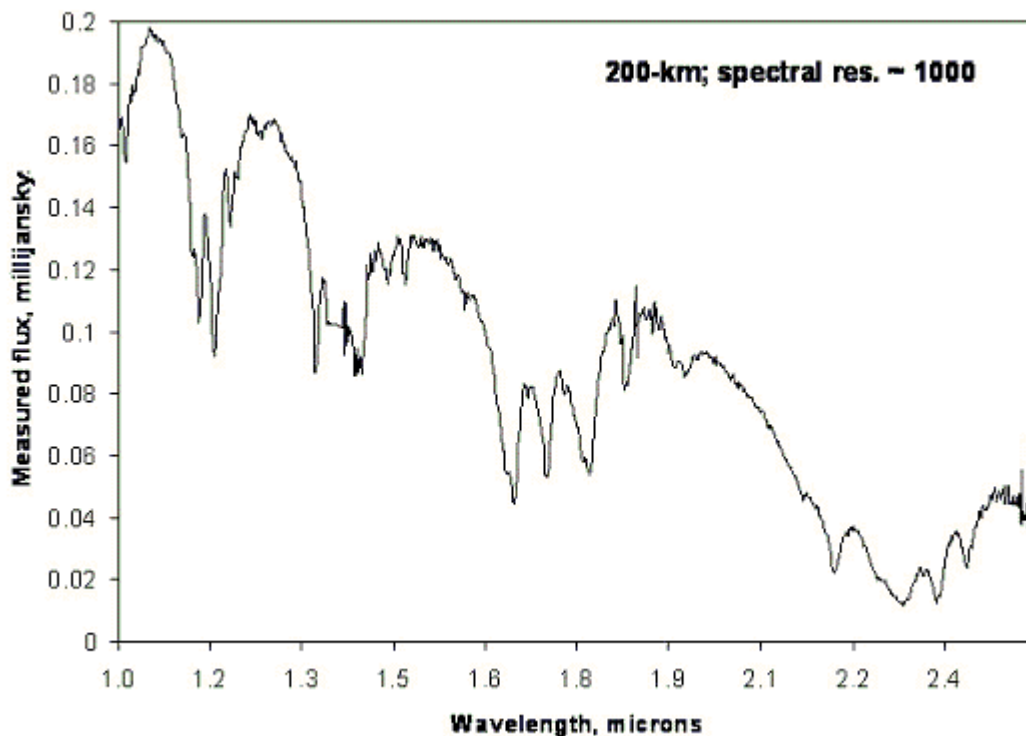
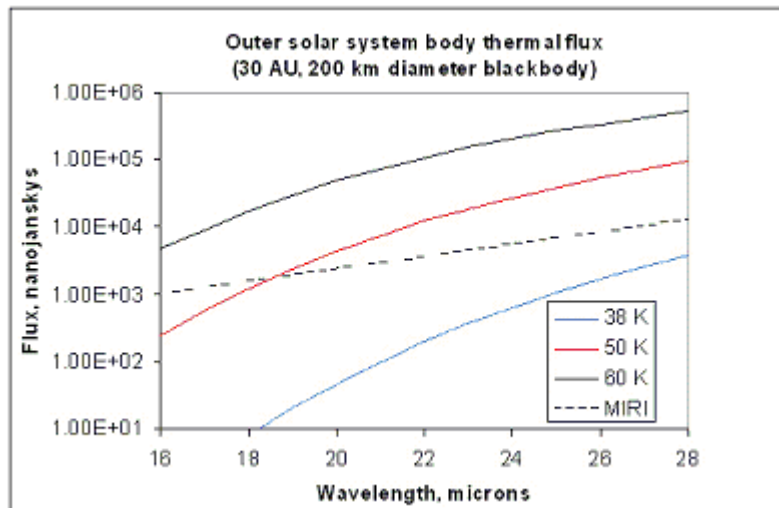


Figure 6-8. Simulated KBO Spectrum

Note: The simulated the spectrum of a Kuiper Belt Object with a 200-km radius using a spectrum borrowed from that of Pluto. This spectrum is well above JWST’s sensitivity limit. Many or all of the major ices can be identified with absorption features out to 2  $\mu\text{m}$ .

JWST must detect the mid-infrared signature of Kuiper Belt Objects down to typical discovery sizes (hundreds of kilometers) in order to provide a separate determination of size and brightness, when combined with the optical brightness (Figure 6-9). These observations will provide information on the dynamical state of the Kuiper Belt, in particular whether the size distribution suggests a population built by accretion or by collisional grinding. They will provide an indication of the color variation and possible existence of multiple classes of Kuiper Belt objects (e.g. volatile-rich versus quiescent).



**Figure 6-9. Sensitivity of MIRI for KBOs**

Note: Blackbody flux versus wavelength in the mid-infrared for a 200-km radius Kuiper Belt object at three different possible temperatures, sited some 30 AU from JWST. Also shown as a dashed line is the JWST sensitivity limit for a broadband detection (SNR=10,  $10^4$  second integration).

**6.2.3 Observatory Capabilities**

**6.2.3.1 Extra-Solar Giant Planets**

As in section 6.1 above.

**6.2.3.2 Circumstellar Disks**

As in section 6.1 above.

### 6.2.3.3 Kuiper Belt

Target tracking requirements for Kuiper Belt objects are equal to, or less than, those of comets (discussed above in section 6.1.3). The additional requirements for these objects include large aperture because of the faintness of the bodies, spectroscopic capability for the larger and closer ones, and accurate tracking rates because of the long exposure times. Current 10-meter ground-based systems can achieve R-band  $\sim 28$  over full night integration, and recent HST-WFPC-2 searches are comparable. JWST will enable deeper searches for new objects and higher sensitivity characterization of known ones. Photometry of near-infrared reflected light and of thermal emission at 20  $\mu\text{m}$  will constrain simultaneously the albedo and the radius. Spectroscopic studies at resolutions of  $\sim 1000$ , comparable to the best ground-based studies of Pluto and Triton, will be extended to smaller and more distant bodies. If spectral resolution 3000 is implemented then the highest resolution spectroscopic studies, reserved for Triton and Pluto, will seek isotopic ratios in the water ice and other components, as well as surface temperature monitoring through the nitrogen overtone band. While the near infrared is the most familiar territory in this regard because of ground-based spectra of Triton and Pluto, rotational features will appear in the mid-infrared and will provide compositional and isotopic data not attainable in the near-IR overtone region. Mid-infrared detection leading to the separation of the size and albedo of such objects is of fundamental importance and requires the high mid-infrared sensitivity that JWST provides.

## 6.3 TRANSITION TO LIFE

### 6.3.1 Questions

#### 6.3.1.1 **What are the Planetary Evolutionary Pathways by which Habitability is Established or Lost?**

How do planets come to be habitable? The role of the giant planets in dynamical stability, timing of terrestrial planet formation, and supply of volatiles, has been discussed above. Do the properties of the precursor gas-dust disks, and their co-evolution with forming planets, also determine habitability through planetary system architecture and planetary masses? Are there young systems that seem, in terms of disk architecture or presence of giant planets, to be on a trajectory to nurture the development of habitable planets?

#### 6.3.1.2 **Does our Solar System Harbor Evidence for Steps on these Pathways?**

Mars is a world on which water once flowed, lakes stood, oceans might have come and gone, and perhaps with them life. Why did Mars become cold and dry? Was loss of water a consequence of the atmospheric erosion of carbon dioxide and other greenhouse gases, or the irreversible production of carbonates on the surface from water and atmospheric carbon dioxide? Where are the carbonates? Are there patches where clues to the evolutionary drying of Mars can be found? What was the original inventory of water on Mars?

Saturn's moon Titan is a Mercury-sized world rich in organic molecules, endowed with a dense atmosphere, and possessed of a water-ice and rock interior that could, under other circumstances, supply abundant water for life. Could Titan have been habitable earlier in its history? How does

the surface-atmosphere exchange of mass and energy work on an organic-rich but abiotic world, which is subject to weak solar forcing but strong seasonal variations? Are there surface and lower atmospheric changes on Titan that might be missed by Cassini in its four-year tour?

Prior to the origin of life, chemical processes may have led to a substantial level of complexity. Depending on the nature of the prebiotic environment, available building blocks may have included amino and hydroxy acids, purines, pyrimidines, assorted sugars and fatty acids. These could have combined to form polymers of largely random sequence and mixed stereochemistry (handedness). Unfortunately, remote detection of stereochemical orientation is almost impossible, while amino acid detection might be possible were sufficient concentrations to exist on optically exposed surfaces.

### **6.3.2 Observations**

JWST will observe Titan to establish a long-time baseline of atmospheric and surface changes connected with the 2004-2008 Cassini mission survey, creating a potentially long (10 year +) baseline of space borne near-infrared observations of Titan's surface and atmosphere. We will use near-infrared spectrometry with spectral resolution a factor of six better than on Cassini to determine the types of organic species present on the surface. Thus, while Cassini gets better spatial resolution, JWST will achieve higher spectral resolution over the mid-latitudes regions of Titan. We will determine whether surface changes or secular atmospheric changes are in evidence over a decadal timescale.

## 6.4 SUMMARY

**Table 6-2. JWST Measurements for the Planetary Systems Theme**

Observation	Primary Instruments	Magnitude or flux	Target Density
Extra-Solar Giant Planets			
Isolated objects	NIRCam	AB = 30 mag	Short list
Bound planets (coronagraphy)*	NIRCam	AB = 18 mag	Short list
In-depth study	FGS-TF NIRSpec MIRI	AB = 27 mag, R = 100 AB = 23 mag, R = 3000 AB = 23 mag	Short list Short list Short list
Circumstellar disks	MIRI	AB = 24 mag	Short list
Comets	NIRSpec MIRI	AB = 23 mag, R = 3000 AB = 23 mag	10/yr
Kuiper Belt Objects	NIRCam NIRSpec MIRI	AB = 30 mag AB = 23 mag, R = 3000 AB = 23 mag	1 arcmin <sup>-2</sup>
Satellites	NIRSPEC MIRI	AB = 23 mag, R = 3000 AB = 23 mag	Short list

\* Note: the bound planet brightness is based on a 2-Jupiter mass planet 5 AU from an M star 10 parsecs from Earth.

**Table 6-3. Required Capabilities for the Planetary Systems Theme**

	Exoplanets	Circum-stellar disks	KBOs, comet nuclei	Comet comae	Satellites
NIRCam					
Broad-band	X				
Sensitivity	X	X	X		X
FOV				X	
PSF					
Dynamic range	X	X			
Coronagraphy	X	X			
NIRSpec					
Multi-obj FOV				X	
R = 100					
R = 1000	X	X	X	X	X
R = 3000	X	X	X	X	X
Sensitivity	X	X	X		X
High Contrast		X			
MIRI					
Broad-band	X	X	X	X	X
Sensitivity	X		X		
FOV		X		X	
PSF					
R = 3000 IFU	X	X	X	X	X
Coronagraphy	X	X		X	
FGS-TF					
Short $\lambda$ TF			X	X	X
Long $\lambda$ TF	X	X	X		
Sensitivity					
FOV					
Coronagraphy					
Observatory					
PSF $\lambda < 1 \mu\text{m}$		X			
Stable PSF	X	X			
Stable Image	X		X	X	X
CVZ					
Moving Targets			X	X	X

## **7.0 REQUIREMENTS RATIONALE**

To achieve its science requirements, JWST must be a large cold telescope, with a wide field of view, exceptional angular resolution and sensitivity, and wide wavelength coverage in both imaging and spectroscopy. The scientific objectives impose observational requirements that can only be met with JWST; no other existing or planned ground-based telescope or space telescope mission can meet these requirements. This section provides the rationale for the science requirements listed in the next section.

### **7.1 WAVELENGTH RANGE**

The wavelength range of JWST (0.6 to 27  $\mu\text{m}$ ) is set by the scientific objectives, as limited by hardware feasibility and by competing facilities. The first galaxies to form are predicted to be bright in the ultraviolet, but the redshift caused by the expansion of the Universe could place their shortest observed wavelengths as long as 2.5  $\mu\text{m}$ . Shorter wavelength coverage is required for those objects that are not as far away. Longer wavelength coverage is required to determine whether the candidate first light objects are in fact at high redshift and devoid of heavy elements. It is also required to observe the formation of stars and planetary systems, because these objects are intrinsically cool, and are hidden by dust.

The short wavelength capabilities of JWST are limited by mirror technology. It would be difficult and costly to produce mirror segments of the required accuracy to provide good images at shorter wavelengths, and indeed it is a difficult challenge to meet the image quality requirements even for the selected wavelength range. The short wavelength cutoff is also set by the mirror coatings, since the optimal coating for longer infrared wavelengths is gold. Gold does not reflect well at wavelengths less than 0.6  $\mu\text{m}$ . In addition, the pointing control system requirements would be much more difficult to meet if the telescope were required to produce the image quality that a large perfect mirror would enable. With a larger telescope working at longer wavelengths, JWST's image quality is qualitatively comparable to that of HST.

The selection of the JWST detector technology also provides limited sensitivity beyond 0.6 and 27  $\mu\text{m}$ . The long wavelength cutoff is also set by the rise of stray light from the sunshield, which limits the gains of JWST sensitivity over Spitzer. Extension of the wavelength range beyond that selected would have been very costly and would not be competitive with other facilities.

### **7.2 IMAGING AND CORONAGRAPHY**

Imaging is fundamental to astronomy, and all of JWST's scientific objectives require images. JWST imaging must resolve small, faint objects and do accurate photometry in a variety of broad and narrow filter bandpasses. Imaging and photometry in broad-band filters (R~5) are required to detect the spectral breaks used to identify high-redshift objects. Broad-band imaging of galaxies at moderate redshift is essential to trace the morphological construction of the Hubble Sequence. Photometry will identify the low-mass end of the stellar initial mass function, and narrow-band imaging will determine the constituents and evolutionary state of proto-planetary systems. JWST's imaging and photometry of Kuiper Belt Objects will lead to their classification and history.

Studies of the history of planetary systems, starting with the proto-planetary disk phase, must contend with the blinding glare from the parent star. High-contrast imaging, or coronagraphy, is required to detect disks and the mature planets they eventually produce.

### **7.3 SPECTROSCOPY**

Choosing the appropriate spectral resolution is essential to identifying and studying high redshift objects, particularly since some are expected to emit half their light in the single Lyman  $\alpha$  spectrum line. It is also essential to the other themes, as it shows relative motions between and within galaxies, and reveals the chemical constituents and physical conditions of all the objects. Spectroscopy with modest spectral resolution ( $R \sim 100$ ) is required for ultimate sensitivity on faint objects with wide spectral features, including galaxies, dust grains, and planetary atmospheres. Higher spectral resolution ( $R \sim 1000$ ) is required to learn about the physical and chemical properties of all objects bright enough to be detected spectroscopically. Even higher resolution ( $R \sim 3000$ ) is required to investigate motions of galaxies, complex molecular emissions from objects like comets, planets, exoplanets, young stars, and star forming regions. The instrument designs will produce spectral resolutions that are functions of wavelength, and must be carefully optimized for the scientific programs.

### **7.4 NEAR-INFRARED CAMERA**

JWST requires a Near-Infrared Camera (NIRCam), covering the 0.6 to 5  $\mu\text{m}$  spectral range. NIRCam must provide a wide field of view, a selection of multiple bandwidths defined by filters, and coronagraphic capabilities. It must also provide the ability to sense the wavefront errors of the observatory, in order to produce the required spatial resolution. NIRCam must have a coronagraphic capability in order to detect extra-solar giant planets.

### **7.5 NEAR-INFRARED SPECTROGRAPH**

JWST requires a Near-infrared Spectrograph (NIRSpec), covering the 0.6 to 5  $\mu\text{m}$  spectral range. NIRSpec must have the ability to produce at least 100 simultaneous spectra from objects selected over a wide field of view with spectral resolutions of  $R \sim 100$  and 1000.

As a goal, NIRSpec may provide a spectral resolution  $R \sim 3000$ , and an Integral Field Unit for imaging spectroscopy at  $R \sim 3000$ .

### **7.6 MID-INFRARED INSTRUMENT**

JWST requires a Mid-InfraRed Instrument (MIRI), covering the 5 to 27  $\mu\text{m}$  spectral range. The MIRI must produce both images and spectra, with resolution up to  $R \sim 3000$ . The spectrometer will be an integral field unit, to improve mapping speed and reduce the need for precise pointing. MIRI shall provide broad band filters for imaging and a coronagraphic capability. As a goal, the long wavelength spectroscopy provided by MIRI may extend to 29  $\mu\text{m}$ . If this goal is adopted, it would enable observation of the molecular hydrogen line at 28.3  $\mu\text{m}$ , which is highly desirable even if there were reduced sensitivity.

## 7.7 TUNABLE FILTERS

JWST requires a Tunable Filter capability (FGS-TF), covering the 1.0 to 4.8  $\mu\text{m}$  spectral range. The FGS-TF must provide high spatial resolution imaging with a spectral resolution of  $R \sim 100$ . The FGS-TF is required for surveys of emission-line objects at particular wavelengths, and must provide a wide field of view. The FGS-TF must include a coronagraphic capability to image circumstellar disks and extra-solar giant planets. As a goal, the wavelength range of the FGS-TF may be extended to shorter wavelengths. The implementation that has been chosen provides this tunable filter capability in the Fine Guidance Sensor.

## 7.8 MINIMUM COLLECTING AREA

JWST requires a telescope with a minimum collecting area of 25 square meters. As the observing time to reach a particular point source sensitivity scales inversely as the square of the area (and the 4<sup>th</sup> power of the diameter), JWST science capability is highly sensitive to this parameter.

## 7.9 IMAGE QUALITY

### 7.9.1 Strehl Ratio

Image quality drives sensitivity, the ability to resolve closely separated sources, and the ability to determine the morphology of the detected objects. It is critical to the scientific success of JWST, but is limited by optical technology. A cost-benefit analysis has determined that a diffraction-limited telescope at wavelengths 2  $\mu\text{m}$  and longer will enable the scientific measurements required by JWST. A diffraction-limited telescope is defined as having a Strehl ratio greater than or equal to 0.8. To enable diffraction-limited imaging and efficient spectroscopy at all wavelengths greater than 2  $\mu\text{m}$ , we require the telescope to be diffraction-limited at 2  $\mu\text{m}$  over the fields of view of NIRCcam and NIRSpec and at 5 micron over the field of view of MIRI. Optimum utilization of diffraction limited imaging further requires that the image be appropriately sampled by the focal plane detector. NIRCcam should Nyquist sample the PSF at 2.0  $\mu\text{m}$  in its short wavelength channel. In the longer wavelength NIRCcam channel the PSF should be Nyquist sampled at  $\sim 4.0 \mu\text{m}$ , while for MIRI the PSF should be Nyquist sampled at  $\sim 7.0 \mu\text{m}$ .

JWST will achieve this image quality using image-based wave front sensing and control (WFS/C) of the primary mirror to ensure that the wavefront error is acceptable. JWST will also use a fine guidance sensor (FGS) in the focal plane to maintain pointing during observations. The WFS/C and FGS are derived requirements.

### 7.9.2 Encircled Energy

JWST is specified to be diffraction-limited at 2.0  $\mu\text{m}$ . At shorter wavelengths where scattering due to mid-frequency mirror figure errors drives imaging performance, image quality has been constrained by means of an encircled energy specification. The minimum encircled energy within a radius of 0.15 arcsec is required to be  $> 74$  percent at 1  $\mu\text{m}$ .

When the image quality is good (that is, when the Strehl ratio is high), most of the light from a star falls into a pattern that is close to the ideal diffraction pattern. When the image quality is less than diffraction limited, the observational capabilities are determined by the details of the map of wavefront error across the telescope pupil. Simulations show that the encircled energy is a good proxy for the ability to determine the colors of distant galaxies from JWST observations, since many of them are very faint and have angular sizes comparable to the telescope beam size. Because the galaxies have unknown shapes and are very faint, we will not be able to use mathematical fitting or deconvolution to correct for the effects of telescope beam shape. The sensitivity of the JWST at short wavelengths is critically dependent on this image quality, and degrades extremely rapidly as the wavefront error increases. At 0.7  $\mu\text{m}$ , the requirement allows the image quality to be far worse than diffraction limited. To aid in the interpretation of short-wavelength images, we must develop the next generation of object finding, object deblending, and object fitting algorithms that incorporate the actual JWST PSF, including its PSF wings and mid-frequency errors.

### **7.9.3 Point-Spread Function Symmetry**

The point spread function (PSF) of JWST images must be azimuthally symmetric. Several observing programs use the detailed shapes of very small galaxies (characterized by an ellipticity) to derive essential characteristics of galaxy formation. In addition, the relative orientations and ellipticities of faint galaxies depend on the gravitational lens effects of intervening matter. In order to understand, measure, and interpret the galaxy shapes, the PSF must be azimuthally symmetric and stable for long exposures.

### **7.9.4 Point-Spread Function Stability**

The JWST image shape must remain constant between wavefront adjustments. If it is not, then photometric accuracy would be degraded, and it would be impossible to determine whether distant galaxies are point objects, round or elliptical. As some regions will be observed in repeated deep exposures over the course of weeks, it is necessary that the effects of instability be controlled, measured, and calibrated in the data processing. Changes greater than 2% in the image shape would limit the accuracy of the scientific results.

## **7.10 SENSITIVITY**

Sensitivity to faint sources is essential to enable JWST to identify them and to determine their colors, redshifts, and chemical and physical conditions. The sensitivity which JWST will achieve is governed primarily by telescope size, optical transmission, image quality, stray light, and detector performance. The requirements have been set through extensive cost-benefit studies, performance budgets, and risk analyses.

The optical transmission requirements have been developed as the maximum reasonably achievable with suitable performance margin. The main contributing factor is the reflectivity of the selected mirror coatings, with small contributions from contamination and micrometeoroid damage. The optical transmission requirements are derived from the sensitivity requirements.

The system-level sensitivity requirements of the JWST have been developed as the best that is reasonably achievable with suitable performance margin. Although stated as requirements for 10,000 or 100,000 second exposures, it will be possible for JWST to detect objects fainter than the values tabulated, using longer exposures.

## **7.11 STRAY LIGHT**

### **7.11.1 Diffuse Stray Light**

The diffuse stray light from the telescope and sunshield must be less intense than the minimum zodiacal light brightness at wavelengths less than 10  $\mu\text{m}$ . Diffuse stray light limits sensitivity by adding noise to the observations. At short wavelengths, self-generated stray light from warm parts of the observatory is not expected to be important. However, at long wavelengths, the dominant sources are expected to be the warm parts of the sunshield and telescope support tower, which emit towards the telescope. In the context of the selected mission design, this is not a driving requirement, because the requirements to provide radiatively cooled near-IR detectors are much more difficult to meet. The stray light limits the sensitivity of the MIRI imaging at longer wavelengths, and requires high detector readout rates to avoid saturation. With the selected mission configuration, it would be exponentially difficult to extend this stray light requirement to wavelengths longer than 10  $\mu\text{m}$ , but any improvements would still translate directly into improved MIRI sensitivity.

### **7.11.2 Local Stray Light and Dynamic Range**

The previous section discusses requirements on diffuse stray light from sources far from the line of sight, i.e. greater than 1 degree away. This section discusses bright sources within the field of view.

These bright sources have two kinds of effects: they produce a diffuse glow that diminishes rapidly with distance from the source, and they produce recognizable features and structures that may hide other targets in certain parts of the field.

The JWST has the sensitivity to detect objects around 32<sup>nd</sup> magnitude in extremely long exposures, but the typical field of view also contains stars of 16<sup>th</sup> or 17<sup>th</sup> magnitude, which are over a million times brighter. Understanding and control of stray light at this level is required for success in the deep exposures planned for JWST. This degree of performance will contribute to coronagraphic capability as well.

The JWST requirements for local stray light are no greater than 2% of the light from point sources shall appear in the form of stray light or ghosts in an SI focal plane. In particular, no ghost images shall have a point spread function smaller than 1 arcsecond in diameter, unless the ghost image is spatially coincident with the observed image, or its integrated flux is less than 1% of the observed image flux.

## 7.12 LONG EXPOSURES

The ultimate scientific potential of the JWST depends critically on the ability to observe survey fields for very long times with multiple field orientations, with possible total exposure times of many weeks.

For reasons of operations planning and efficiency, the Observatory must be able to observe any fixed field at a fixed roll angle for a continuous period of 10 days. This is essential for operations planning and efficiency, and may be required to obtain good scientific results. Due to systematic errors that might be introduced with changes of roll angle, we must be able to reach the needed cosmological sensitivity without a change of roll angle. Ten day cumulative exposures on a single field are required for deep observations.

Ultimate sensitivity with long exposures will require careful analysis and calibration of the instruments in flight, and the development of ground software to compensate for anomalous effects in the data. Complex observing strategies will be required to randomize the various errors and allow them to average out on long exposures. These exposures may last for many weeks, possibly divided into shorter exposures spread out over months. There is no requirement for them to be accomplished all at once, and there are good reasons to spread them out so that time-dependent phenomena like supernovae and calibration drifts can be recognized. The goal is that the signal to noise ratio continues to improve as the square root of observing time, out to the longest practical times. It is recognized that no practical hardware requirement can be written and no test program can be devised at present to guarantee that these long exposures achieve the desired sensitivity, but the observatory should be designed and analyzed with this goal in mind.

Phenomena that could interfere with achieving this goal include uncontrolled drifts in the calibration parameters of the telescope and detectors, changes of the point spread function and stray light, aging of the detectors due to cosmic ray damage, uncontrolled temperature fluctuations, failures of individual microshutters or detector pixels, and changes of the telescope transmission function with contamination and micrometeoroid etching.

JWST must develop operational sequences and data reduction and calibration software to enable these long exposures. The software must be able to register and combine exposures made under different conditions of field offset, orientation, field distortion, and overall scale.

## 7.13 PHOTOMETRIC ACCURACY AND CALIBRATION

In order to interpret the observations made by JWST, they must be converted to physical units and compared to theory or other observations. The observations must be repeatable and calibratable. Calibration of JWST imaging will be done by standard procedures which use dark imaging, flat fields, standard stars, and self-calibration techniques. Calibration of spectroscopy will be more difficult, due to the complexity of the microshutter array or integral field units, and the dependence on the position of the object within a slit.

The requirement for calibrated data implies knowledge of instrument and filter transmission functions, the ability to observe objects already calibrated by other equipment, and the ability to

produce detailed models for deviations from ideal behavior. This work will be done by the instrument teams, the integration and test teams, and the science operations center.

#### **7.14 FIELD OF VIEW**

A wide field of view is essential, but costly. The earliest observable objects are expected to be rare in survey images dominated by nearby galaxies. The widest possible field of view combined with the longest possible exposure times and highest sensitivity is essential in the search. However, a wide field of view is costly because of the number of sensors required, so JWST must be designed to achieve the maximum affordable field consistent with the required angular resolution. For surveys of unresolved point sources with a Nyquist-sampled telescope, the performance metric is  $D^2 N_{\text{pix}}$ , where  $D$  is the telescope diameter and  $N_{\text{pix}}$  is the number of detector pixels. There is a very high premium on maintaining the field of view of the cameras, and an observatory with 1/4 the number of pixels would be comparable in survey efficiency to an observatory with half the mirror diameter.

#### **7.15 ANGULAR RESOLUTION**

Angular resolution, governed primarily by the telescope size, is required to separate the small, faint images of the first light objects from the overlapping images of abundant nearby galaxies. In most cases the instruments are required and designed to be Nyquist limited near the middle of the spectral range, i.e. the pixel size =  $\lambda/2D$ , where  $D$  is the circumscribed telescope diameter. This allows for a relatively wide field of view but requires sub-pixel rastering, or dithering, of the pointing to achieve accurate imaging and photometry for shorter wavelengths, and to achieve the maximum angular resolution allowed by the hardware. Similar considerations apply to the spectrometers, particularly because their effective entrance slits are much larger than the telescope Nyquist limit. The requirement of a 25 m<sup>2</sup> collecting area and the shape of the primary mirror fix the telescope diffraction limited angular resolution.

#### **7.16 POINTING ACCURACY**

Pointing accuracy, resolution, and stability are related to the angular resolution and sensitivity requirements. The detailed JWST pointing requirements are derived from the telescope optical error budget (Ref. #) and the pointing error budget (Ref. #). Initial pointing accuracy must be sufficient to enable the required guide stars to be located. After guide star acquisition, the accuracy and resolution must be sufficient to permit the sub-pixel dithering needed for accurate image reconstruction at wavelengths that are not Nyquist sampled. In addition, pointing must be able to place selected targets on the entrance slits of the spectrometers, and on the black spots or nulling spots of the coronagraphs.

After data acquisition, pointing must be known well enough to compare observations with those from other observatories. This places requirements on the knowledge of the optical field distortion of the telescope and instruments, and on the accuracy of the guide star catalog. For those objects that merit precise position determination, the knowledge of the telescope internal coordinate system must not be the limiting factor.

These requirements lead to specifications on knowledge of the field distortion, of the order of 5 milliarcsec. Being able to accurately account for optical field distortion is also critical for allowing ~100 or more objects to be placed accurately on their respective slits during NIRSpec multi-object spectroscopic observations. Flight calibration of the field distortion is required to confirm predictions based on the optical design and to measure parameters that may change after launch.

### **7.17 FIELD OF REGARD**

JWST must be able to see regions 5 degrees in radius around the Ecliptic poles continuously all year, in order to build up an ultra-deep exposure through repeated observations, while searching for variable objects like supernovae. The Ecliptic poles are relatively dark in zodiacal light and interstellar dust emission, and will be prime targets for cosmological studies. A continuous viewing zone will also allow repeated observations of reference stars to verify observatory performance and photometric stability.

The JWST must be able to observe any object in the sky at some time during the year. Although cosmologically interesting targets have a relatively uniform distribution across the sky, JWST must be able to observe nearby objects that are not distributed evenly, and fields that have been previously observed with other observatories.

The JWST must also be able to observe in any specific direction for a large fraction of the year. This permits the build-up of long observations, permits flexibility in scheduling, and permits exploratory observations for long observing programs that will be implemented later.

In order to take full advantage of the required field of regard, JWST must be able to observe any point in it at any allowable roll angle, with a probability of acquiring a guide star of at least 95% under nominal conditions. This will ensure that most of the required targets will be observable. Some of the observing programs consist of mapping relatively large contiguous areas of the sky. The inability to find a guide star would leave a hole in the map. Scheduling workarounds if an observation cannot be made as expected would be costly, would reduce observing efficiency, and might not succeed.

### **7.18 MOVING OBJECT TRACKING**

Comets in the inner solar system move at rates up to 0.03 arcsec/s relative to the guide stars. Outer solar system objects move more slowly, but are much fainter and require long exposure times. Precise trajectories are needed at the angular resolution of the observatory to avoid loss of angular resolution and to enable placing the moving targets on specific NIRSpec slits or the MIRI spectrometer field of view. It is recognized that the image quality requirements will have to be relaxed slightly while tracking moving objects, and that exposure times (or the track paths) will be limited by the path of a the guide star. There are no requirements to track accelerating objects (curved trajectories), to track objects continuously as the guide stars cross sensor chip boundaries, to have special guide star availability, or to observe in any special orientations.

## 7.19 MISSION LIFETIME

The JWST must operate with all science instruments for at least 5 years after completion of commissioning. Long lifetime is required to discover and monitor high redshift supernovae, because their characteristic decay time of 2 months is stretched (by the cosmic redshift) to 2 years or more. Long life is also required to enable follow-up observations to be planned after initial discoveries. The observed scientific turnaround time to analyze data, publish them, and propose new observations is 3 years. Finally, even observing the sky as fast as it can go, JWST can cover only a tiny fraction in its lifetime. The science that JWST will do is almost directly proportional to the total observing time; or the product of mission lifetime and observing efficiency. In order to exploit the full scientific potential of the mission, a lifetime of 10 years is desired. Although we do not require mission assurance to guarantee a lifetime greater than five years, JWST must maintain the possibility of a mission lifetime of ten years.

## 7.20 OBSERVING EFFICIENCY

JWST is a precious resource and its scientific output and effective sensitivity is expected to depend strongly on its operational efficiency. Increasing the observing efficiency by 10% is comparable in effect to increasing the number of pixels by 10% or increasing the telescope diameter by 5%. Efficient operation is required, despite the many ways that observing time can be lost.

These ways include: initial on-orbit checkout, calibration time, recovery from transient events, repointing, acquisition of guide stars, instrument set-up, firing the jets, safe hold due to anomalies, data losses in transmission, wavefront sensing observations, cosmic ray hits on the detectors, and energetic Solar flares.

To meet the level 1 requirement for observing time, JWST must achieve 70% observation efficiency. It is expected that parallel instrument calibration will be required to achieve this.

## 7.21 DATA RATE

The JWST instrument sensitivity is limited by cosmic rays, which can not be avoided. As a result, detectors must be read frequently and the data sent to the ground for cosmic ray recognition and removal. The data rate must be sufficient so that cosmic ray effects can be mitigated in ground data processing, without degrading the sensitivity significantly. The decision has been made that no on-board processing will be provided to perform cosmic ray detection and removal. This is scientifically conservative and minimizes the cost of flight hardware, but requires a relatively high data rate and long telemetry contact times.

The data rate to the ground is set in the Level 1 requirement. The received data must not be corrupted, which leads to requirements which limit the error rate.

## 7.22 TARGETS OF OPPORTUNITY

On rare occasions it will be necessary to interrupt the JWST observing plan to make special observations. Examples might include supernovae or gamma ray burst sources (now thought to

be related). JWST must provide the capability to make special observations with priority determined by the scientific importance and time scales determined by the natural phenomena. We do not require a response of less than one day.

### **7.23 PUBLIC ACCESS**

The science requirements in this document define the measurement capabilities of the telescope, but do not require that particular observations be made. Allocation of the observing time on the JWST will be made to the astronomical community according to the policies described in the JWST Program Plan.

## 8.0 SCIENCE REQUIREMENTS

This section contains the JWST science requirements. The rationale for the requirements in each sub-section below is provided in the corresponding sub-section of section 7. All JWST requirements which are not referenced herein, are derived from the requirement that JWST be capable of making astronomical observations, as specified in the Program Plan and by SR-1.

The full traceability of each science requirement by theme and instrument capability is summarized below in Table 8.1. For reference the flowdown of Science Requirements to the JWST Program Plan (PP; Level 1 requirements), the Mission Requirements document and the ISIM Requirements document is also provided in Appendix B.

**Table 8.1: Science Requirement Traceability Matrix**

	First Light	Assembly of Galaxies	Birth of Stars and Protostars	Planetary Systems and the Origins of Life	Science Requirements
<b>NIRCam</b>					
Broad-band	✓	✓	✓	✓	SR-1, SR-2, SR-4, SR-5
Sensitivity	✓	✓	✓	✓	SR-10, SR-16, SR-17
Photometric Accuracy	✓	✓	✓	✓	SR-20
FOV	✓	✓	✓	✓	SR-21
PSF	✓	✓			SR-11, SR-22, SR-14, SR-15
PSF $\lambda < 1 \mu\text{m}$	✓	✓	✓	✓	SR-13
Coronagraphy			✓	✓	SR-3
<b>NIRSpec</b>					
Multi-obj FOV	✓	✓	✓	✓	SR-7, SR-21
R = 100	✓	✓	✓		SR-4, SR-6
R = 1000	✓	✓	✓	✓	SR-4, SR-6
R = 3000		✓	✓	✓	SR-4, SR-6
Sensitivity	✓	✓	✓	✓	SR-10, SR-16
Photometric Accuracy	✓	✓	✓	✓	SR-20
PSF	✓	✓			SR-11, SR-22, SR-14, SR-15
PSF $\lambda < 1 \mu\text{m}$	✓	✓	✓	✓	SR-13
High Contrast			✓	✓	
<b>MIRI</b>					
Broad-band	✓	✓	✓	✓	SR1, SR2, SR4, SR-8, SR-12
Sensitivity	✓	✓	✓	✓	SR-10, SR-16, SR-17
Photometric Accuracy	✓	✓	✓	✓	SR-20
FOV	✓	✓	✓	✓	SR-21
PSF	✓	✓	✓	✓	SR-12, SR-23

R = 3000 IFU	✓	✓	✓	✓	SR-4
Coronagraphy			✓	✓	SR-3
<b>FGS-TF</b>					
Short $\lambda$ TF	✓	✓	✓	✓	SR-9
Long $\lambda$ TF	✓	✓	✓	✓	SR-9
Sensitivity	✓	✓			SR-10, SR-16, SR-17
Photometric Accuracy	✓	✓	✓	✓	SR-20
FOV	✓	✓			SR-21
PSF	✓	✓			SR-11, SR-22, SR-14, SR-15
Coronagraphy			✓	✓	SR-3
<b>Observatory</b>					
Stable PSF	✓	✓	✓	✓	SR-11, SR-12, SR-13, SR-14, SR-15
Field of Regard	✓	✓	✓	✓	SR-26, SR-27, SR-28, SR-30
Stable Image	✓	✓	✓	✓	SR-19
Pointing Precision	✓	✓	✓	✓	SR-24
Precision Offsets	✓	✓	✓	✓	SR-25
CVZ	✓				SR-29
Moving Targets				✓	SR-31
Mission Lifetime	✓	✓	✓	✓	SR-32, SR-33
Observing Efficiency	✓	✓	✓	✓	SR-34
Data Rate	✓	✓	✓	✓	SR-35
Targets of Opportunity	✓	✓	✓	✓	SR-36
Public Access	✓	✓	✓	✓	SR-37

## 8.1 WAVELENGTH RANGE

SR-1 JWST shall be capable of making astronomical observations at wavelengths from 0.6 to 27 micrometers.

## 8.2 IMAGING AND CORONAGRAPHY

SR-2 JWST shall provide imagery with spectral resolution in the range  $3 < R < 200$ , including broadband ( $3 < R < 7$ ) imaging with a minimum of 16 discrete filter bandpasses, continuously distributed over a wavelength range 0.6 to 27 micrometers.

SR-3 JWST shall have coronagraphic imaging capability over the wavelength ranges 2 to 27 micrometers

## 8.3 SPECTROSCOPY

SR-4 JWST shall provide spectroscopy with spectral resolution in the range  $50 < R < 5000$  over a wavelength range 0.6 to 27 micrometers

#### 8.4 NEAR-INFRARED CAMERA

SR-5 JWST shall have a near-infrared camera (NIRCam) capable of operating over the wavelength range 0.6 to 5 micrometers and producing images with spectral resolution less than 100.

#### 8.5 NEAR-INFRARED SPECTROGRAPH

SR-6 JWST shall have a near-infrared spectrograph (NIRSpec) operating over the wavelength range 0.6 to 5 micrometers and producing spectra with spectral resolutions of approximately 100 and 1000

SR-7 NIRSpec shall be capable of obtaining simultaneous spectra of more than 100 objects.

As a goal, NIRSpec may also provide a spectral resolution of order 3000.

#### 8.6 MID-INFRARED INSTRUMENT

SR-8 JWST shall have a mid-infrared instrument (MIRI) capable of operating over the wavelengths range 5 to 27 micrometers and producing both images with spectral resolution less than 100, and spectra with spectral resolution  $R \sim 2000$ .

As a goal, the transmission of the optical telescope element and the long wavelength spectroscopy provided by MIRI may extend to 29 micrometers.

#### 8.7 MEDIUM AND NARROW-BAND IMAGING

SR-9 JWST shall have tunable filters capable of operating over the wavelength range 1.2 to 4.8 micrometers and producing images with spectral resolution of order 100.

As a goal, the tunable filters of FGS-TF may extend shorter of 1.2 micrometers and up to 5 micrometers.

SR-38 JWST shall provide a medium band imaging capability ( $8 < R < 20$ ) over the wavelength range 1.0 micrometers to 5 micrometers, and a limited narrowband capability ( $20 < R < 200$ ) comprised of discrete filters.

#### 8.8 PRIMARY MIRROR AREA

SR-10 JWST shall have a primary mirror whose unobscured light collecting area is no less than 25 square meters.

## 8.9 IMAGE QUALITY

### 8.9.1 Strehl Ratio

- SR-11 JWST shall be capable of diffraction-limited imaging at 2.0 micrometers wavelength, defined as having a Strehl Ratio greater than or equal to 0.8.
- SR-39 JWST shall shall be capable of diffraction-limited imaging at 4.0 micrometers wavelength, defined as having a Strehl Ratio greater than or equal to 0.8.
- SR-12 JWST shall be capable of diffraction-limited imaging at 5.6 micrometers wavelength, defined as having a Strehl Ratio greater than or equal to 0.8.

### 8.9.2 Encircled Energy

- SR-13 The total encircled energy of an image of a point source over the FOV of the Near-Infrared Camera shall be greater than 74 percent within a circle of 0.15 arc-second radius at a wavelength of 1 micrometers, and shall remain so over a period of 24 hours without intervention by ground command

### 8.9.3 Point Spread Function Symmetry

- SR-14 The maximum effective anisotropy in NIRCcam for a PSF at a wavelength of 2  $\mu\text{m}$  weighted with a circular Gaussian with a 1-sigma radius ranging from 0.1" to 1" shall not exceed 1% when averaged in quadrature over the field of view and over the time interval between successive wavefront sensing measurements.

### 8.9.4 Point Spread Function Stability

- SR-15 The rms variation of each component of the effective anisotropy of the JWST PSF at 2  $\mu\text{m}$ , weighted with a circular Gaussian of widths between 0.1" and 1", shall not exceed 0.1% over the field of view in NIRCcam at any given time.

## 8.10 SENSITIVITY

- SR-16 “Sensitivity” is defined to be the brightness of a point source detected with a signal-to-noise ratio of 10 in a 10,000 second integration, or 100,000 seconds for a NIRSpec exposure. Everywhere within the field of view of an instrument, which has the appropriate wavelength coverage and spectral resolution, the sensitivity for science observations of targets at the North Ecliptic Pole shall be at least that specified in the following table. (See Appendix A of the MRD for reference parameters) (MR-51).

**Table 8-1. Required Sensitivity Values**

Wave-length (μm)	Instrument / Mode	Sensitivity
1.1	NIRCam	$1.21 \times 10^{-34} \text{ Wm}^{-2}\text{Hz}^{-1}$ SN=10 in 10,000 s or less and R=4 bandwidth
2	NIRCam	$1.04 \times 10^{-34} \text{ Wm}^{-2}\text{Hz}^{-1}$ SN=10 in 10,000 s or less and R=4 bandwidth
3.5	FGS-TF	$3.68 \times 10^{-33} \text{ Wm}^{-2}\text{Hz}^{-1}$ SN=10 in 10,000 s or less and R=100 bandwidth
3.0	NIRSpec/ Low Res	$1.2 \times 10^{-33} \text{ Wm}^{-2}\text{Hz}^{-1}$ SN=10 in 10,000 s or less and R=100 bandwidth
2.0	NIRSpec/ Med Res	$5.2 \times 10^{-22} \text{ Wm}^{-2}$ SN=10 in 100,000 s or less
10	MIRI/ Broad-Band	$7.0 \times 10^{-33} \text{ Wm}^{-2}\text{Hz}^{-1}$ SN=10 in 10,000 s or less and R=5 bandwidth
21	MIRI/ Broad-Band	$7.3 \times 10^{-32} \text{ Wm}^{-2}\text{Hz}^{-1}$ SN=10 in 10,000 s or less and R=4.2 bandwidth
9.2	MIRI/ Spectrometer	$1.0 \times 10^{-20} \text{ Wm}^{-2}$ SN=10 in 10,000 s or less and R=2400 bandwidth
22.5	MIRI/ Spectrometer	$5.6 \times 10^{-20} \text{ Wm}^{-2}$ SN=10 in 10,000 s or less and R=1200 bandwidth

## 8.11 STRAY LIGHT

### 8.11.1 Diffuse Stray Light

SR-17 JWST shall be capable of Zodiacal light background limited imaging over the wavelength range 0.6 to 10 micrometers.

### 8.11.2 Local Stray Light

SR-18 No greater than 2% of the light from point sources shall appear in the form of stray light or ghosts in an SI focal plane. In particular, no ghost images shall have a point spread function smaller than 1 arcsecond in diameter, unless the ghost image is spatially coincident with the observed image, or its integrated flux is less than 1% of the observed image flux.

## 8.12 LONG EXPOSURES

SR-19 The observatory shall be capable of providing the same orientation of its FOV over at least 10 consecutive days for observations of any available fixed target.

**8.13 PHOTOMETRIC ACCURACY AND CALIBRATION**

SR-20 JWST shall be capable of achieving data calibration into physical units with absolute accuracies shown in the following table.

**Table 8-2. Required Calibration Accuracies**

<b>Absolute Calibration Accuracy</b>	<b>Flux (%)</b>	<b>Flux (%)</b>	<b>Flux (%)</b>	<b>Wavelength (% resolution element)</b>
	Imagery	Coronagraphic Imagery	Spectroscopy	Spectroscopy
NIRCam	5	5	NA	NA
NIRSpec	NA	NA	10	12.5
MIRI	5	15	15	10
FGS-TF	5	10%	NA	10%

**8.14 FIELD OF VIEW**

SR-21 The minimum effective science field of view in square arcminutes shall be 9.4 for NIRCam, 9.0 for NIRSpec, 3.5 for MIRI and 4.8 for FGS-TF.

**8.15 ANGULAR RESOLUTION**

SR-22 NIRCam shall Nyquist sample the diffraction limit of the OTE at 2.0 micrometers.

SR-40 NIRCam shall Nyquist sample the diffraction limit of the OTE at 4.0 micrometers.

SR-23 MIRI imaging shall Nyquist sample the diffraction limit of the OTE at no more than 7.0 micrometers.

**8.16 POINTING ACCURACY**

SR-24 JWST shall be capable of pointing with an absolute astrometric accuracy of 1 arcsec RMS.

SR-25 JWST shall be capable of relative offsets with an accuracy of 5 milliarcsec RMS.

**8.17 FIELD OF REGARD**

SR-26 Over an interval of one sidereal year, the Observatory shall be capable of observing anywhere within the celestial sphere.

- SR-27 The Observatory shall be capable of observing at least 35% of the celestial sphere at any time.
- SR-28 The Observatory shall be capable of observing targets in 50% of the celestial sphere for at least 60 consecutive days per year.
- SR-29 The Observatory shall be capable of observing any point on the celestial sphere within 5 degrees of the ecliptic poles at any time.
- SR-30 The Observatory shall have greater than 95% probability of acquiring a guide star and maintaining pointing stability on any fixed target for any valid attitude within the Field of Regard.

### **8.18 MOVING OBJECT TRACKING**

- SR-31 When requested the Observatory shall track targets which exhibit any angular velocity in the range of 30 mas/s over a total motion 30 arcsec with respect to the guide star (**TBR**).

### **8.19 MISSION LIFETIME**

- SR-32 The JWST science mission lifetime, after commissioning, shall be a minimum of 5 years.

As a goal, the JWST science mission lifetime shall be 10 years. In support of this goal, JWST shall meet the following requirement:

- SR-33 Propellant shall be sized for 10 years of operation after launch.

### **8.20 OBSERVING EFFICIENCY**

- SR-34 After commissioning, the ratio of prime exposure time on scientific targets to total elapsed time shall be greater than 70%.

### **8.21 DATA RATE**

- SR-35 During a normal operations contact, the JWST shall be capable of downlinking, and the Ground Segment capable of capturing and processing 229 Gigabits of science data, compressed from 458 Gigabits.

### **8.22 TARGETS OF OPPORTUNITY**

- SR-36 JWST shall be capable of observing targets of opportunity on timescales less than or equal to 2 days.

### **8.23 PUBLIC ACCESS**

SR-37 JWST shall enable a general observer and archival research program with observing time allocated according to peer review of science proposals.

**APPENDIX A. ABBREVIATIONS AND ACRONYMS**

<b>Abbreviation/ Acronym</b>	<b>Definition</b>
$\mu\text{m}$	Micrometer
$\text{\AA}$	Angstrom, 0.1 nm, $10^{-4} \mu\text{m}$
AB magnitude	$31.4 - 2.5 \log_{10}(F_{\nu}/1 \text{ nJy})$ , where $F_{\nu}$ is flux density
$\text{AB}_{1400}$	AB magnitude at 1400 $\text{\AA}$
ACS	Advanced Camera for Surveys
AGN	Active Galactic Nucleus
ASWG	Ad Hoc Science Working Group
AU	Astronomical Unit, mean distance of Earth from Sun
$A_{\nu}$	Absorption (in magnitudes) at V band
BLR	Broad Line Region
CDM	Cold Dark Matter
CMB	Cosmic Microwave Background
COBE	Cosmic Background Explorer
DRM	Design Reference Mission
EW	equivalent width
FWHM	Full width at half maximum
$F_{\nu}$	spectral flux density, e.g. $\text{W/m}^2 \text{ Hz}$
GO	General Observer
GOODS	Great Observatories Origins Deep Survey
GRB	Gamma Ray Burst
HDF	Hubble Deep Field
HR diagram	Hertzsprung-Russell diagram, in which each point shows the temperature or color of a star, and its luminosity
IFU	Integral Field Unit
IGM	intergalactic medium
IMF	Initial mass function, the distribution of stellar masses when they are formed
IRAS	Infrared Astronomical Satellite
ISO	Infrared Space Observatory
ISOCAM	Camera on ISO satellite

ISWG	Interim Science Working Group
JWST	James Webb Space Telescope
Jy	Jansky, $10^{-26}$ W/m <sup>2</sup> Hz
KBO	Kuiper Belt Object
$L_{\odot}$	luminosity of the Sun
$L_{AB}$	AB magnitude in L photometric filter band
$L_{bol}$	bolometric luminosity
Ly $\alpha$ , Lyman $\alpha$	Hydrogen emission line from n=2 to 1 at 0.1216 micrometers wavelength
$M_{\odot}$ , $M_{sun}$	mass of the Sun
$M_B$	absolute magnitude in the rest-frame B band
$M_{AB}$	AB Magnitude, See Above
MIRI	Mid Infrared Instrument
MRD	Mission Requirements Document
$M_{sun}$ $M_{Jup}$	mass of the Sun, mass of Jupiter
Myr	Million years
N(z)	number density of objects as function of redshift z
NIR	near infrared, 1-5 micrometers, $\mu$ m
NIRCam	Near Infrared Camera
NIRSpec	Near Infrared Spectrograph
nm	Nanometer, $10^{-3}$ $\mu$ m
OPR	ortho-to-para ratio
PAH	polycyclic aromatic hydrocarbons, a large class of abundant interstellar molecules
pc	Parsec, distance at which 1 AU subtends 1 arcsec, about 3.26 light years
Pop I	Population I stars, recently formed in the spiral disks of galaxies from material enriched by earlier generations
Pop II	Population II stars, formed much earlier before spiral disks
Pop III	Primordial population III stars, formed of original hydrogen and helium
Proplyd	Protoplanetary disk
QSO	Quasi-Stellar Object
R	spectral resolution, $\lambda/\Delta\lambda$
rms	Root mean square
S/N, SNR	signal to noise ratio
SED	Spectral Energy Distribution

SFR	Star Formation Rate
SMC	Small Magellanic Cloud
SN, SNe	Supernova, Supernovae
SRD	Scientific Requirements Document
SWG	Science Working Group
UDF	Ultra Deep Field
ULIRG	Ultra-Luminous Infrared Galaxy
UV	ultraviolet
WMAP	Wilkinson Microwave Anisotropy Probe
y dex	A factor of $10^y$
YSO	Young Stellar Object
z	redshift
Z	metallicity
$Z_{\text{sun}}, Z_{\odot}$	metallicity of the Sun
$\Lambda$ CDM	Lambda Cold Dark Matter (a form of cosmology with acceleration)
$\lambda_{3727}$	wavelength of 3727 Angstroms
$\sigma$	Standard deviation
$\tau$	Optical depth (dimensionless)

## APPENDIX B: SCIENCE TO MISSION REQUIREMENT TRACEABILITY SUMMARY

A-3

CHECK WITH JWST DATABASE AT:

<https://ngst1.hst.nasa.gov/>

TO VERIFY THAT THIS IS THE CORRECT VERSION PRIOR TO USE.

SR No.	Science Requirement	MRD	Program Plan	ISIM
SR-1	JWST shall be capable of making observations at wavelengths from 0.6 - 27 microns	107	PP 5.1.1	ISIM-152
SR-2	JWST shall provide imagery with spectral resolution in the range $3 < R < 200$ , including broadband ( $3 < R < 7$ ) imaging with a minimum of 16 discrete filter bandpasses, uniformly distributed over a wavelength range 0.6 to 27 micrometers.	185	PP 5.5.1	ISIM-252
SR-3	JWST shall have coronagraphic imaging capability over the wavelength ranges 2 to 27 micrometers		PP 5.1.1.3	ISIM-852
SR-4	JWST shall provide spectroscopy with spectral resolution in the range $50 < R < 5000$ over a wavelength range 0.6-27 microns	186	PP 5.5.2	ISIM-253
SR-5	JWST shall have a near-infrared camera (NIRCam) capable of operating over the wavelength range 0.6 to 5 micrometers and producing images with spectral resolution less than 100.		PP 5.5.1	ISIM-153 ISIM-367
SR-6	JWST shall have a near-infrared spectrograph (NIRSpec) operating over the wavelength range 0.6 to 5 micrometers and producing spectra with spectral resolutions of approximately.		PP 5.5.2	ISIM-450
SR-7	NIRSpec shall be capable of obtaining simultaneous spectra of more than 100 objects.		PP 5.1.1.2	ISIM-1247 ISIM-451
SR-8	JWST shall have a mid-infrared instrument (MIRI) capable of operating over the wavelengths range 5 to 27 micrometers and producing both images with spectral resolution less than 100, and spectra with spectral resolution $R \sim 1000$ .		PP 5.5.3	ISIM-153 ISIM 495-497
SR-9	JWST shall have tunable filters capable of operating over the wavelength range 1.0 to 4.8 micrometers, and producing images with spectral resolution of order 100.		PP 5.1.1.1 PP 5.1.1.3	ISIM-153 ISIM-586
SR-38	JWST shall provide a medium band imaging capability ( $8 < R < 20$ ) over the wavelength range 1.0 micrometers to 5 micrometers, and a limited narrowband capability ( $20 < R < 200$ ) comprised of discrete filters.		PP 5.5.1 PP 5.5.3	
SR-10	The JWST Optical Telescope Element shall have a primary mirror whose unobscured light collecting area is no less than 25 square meters	198	PP 5.4.1	
SR-11	JWST shall be capable of diffraction-limited imaging at 2.0 micrometers wavelength, defined as having a Strehl Ratio greater than or equal to 0.8.	110	PP 5.4.2	
SR-39	JWST shall be capable of diffraction-limited imaging at 4.0 micrometers wavelength, defined as having a Strehl Ratio greater than or equal to 0.8.	110		

SR-12	JWST shall be capable of diffraction-limited imaging at 5.0 micrometers wavelength, defined as having a Strehl Ratio greater than or equal to 0.8.	116		ISIM-851
SR-13	The total encircled energy of an image of a point source over the FOV of the Near-Infrared Camera shall be greater than 74 percent within a circle of 0.15 arc-second radius at a wavelength of 1 micrometers, and shall remain so over a period of 24 hours without intervention by ground command.	111,113, 114,115	PP 5.4.3	ISIM-847
SR-14	The maximum effective anisotropy in NIRCcam for a PSF at a wavelength of 2 micrometers weighted with a circular Gaussian with a 1-sigma radius ranging from 0.1" to 1" shall not exceed 1% when averaged in quadrature over the field of view and over the the time interval between successive wavefront sensing measurements.	118		ISIM-233
SR-15	The rms variation of each component of the effective anisotropy of the JWST PSF at 2 micrometers, weighted with a circular Gaussian of widths between 0.1" and 1", shall not exceed 0.1% over the field of view in NIRCcam at any given time	119		ISIM-234
SR-16	“Sensitivity” is defined to be the brightness of a point source detected with a signal-to-noise ratio of 10 in a 100,000 second integration. Everywhere within the field of view of an instrument, which has the appropriate wavelength coverage and spectral resolution, the sensitivity for science observations of targets at the North Ecliptic Pole shall be at least that specified in the following table.	51	PP 5.4.4	ISIM-153
SR-17	JWST shall be capable of Zodiacal light background limited imaging over the wavelength range 0.6 to 10 micrometers.	121, 122, 125	PP 5.4.4	
SR-18	No greater than 2% of the light from point sources shall appear in the form of stray light or ghosts in an SI focal plane. In particular, no ghost images shall have a point spread function smaller than 1 arcsecond in diameter, unless the ghost image is spatially coincident with the observed image, or its integrated flux is less than 1% of the observed image flux.			
SR-19	The observatory shall be capable of providing the same orientation of its FOV over at least 10 consecutive days for observations of any available fixed target.	177		

SR-20	JWST shall be capable of achieving data calibration into physical units with absolute accuracies as defined.	110, 111, 113, 114, 155, 308, 313, 315, MRDXX		ISIM-1290
SR-21	The minimum effective science field of view in square arcminutes shall be 9.4 for NIRCcam, 9.0 for NIRSpec, 3.5 for MIRI and 5.4 for FGS-TF.	369, 370		ISIM-235
SR-22	NIRCcam shall Nyquist sample the diffraction limit of the OTE at 2.0 micrometers.			Add to ISIM <b>(TBD)</b>
SR-40	NIRCcam shall Nyquist sample the diffraction limit of the OTE at 4.0 micrometers.			Add to ISIM <b>(TBD)</b>
SR-23	MIRI imaging shall Nyquist sample the diffraction limit of the OTE at 6.4 micrometers.			Add to ISIM <b>(TBD)</b>
SR-24	JWST shall be capable of pointing with an absolute astrometric accuracy of 1 arcsec RMS.	173, 175		ISIM-256
SR-25	JWST shall be capable of relative offsets with an accuracy of 5 milliarcsec RMS.	174, 182		ISIM-257
SR-26	Over an interval of one sidereal year, the Observatory shall be capable of observing anywhere within the celestial sphere.	103		
SR-27	The Observatory shall be capable of observing at least 35% of the celestial sphere at any time	104		
SR-28	The Observatory shall be capable of observing targets in 50% of the celestial sphere for at least 60 consecutive days per year.	105		
SR-29	The Observatory shall be capable of observing any point on the celestial sphere within 5 degrees of the ecliptic poles at any time.	106		
SR-30	The Observatory shall have greater than 95% probability of acquiring a guide star and maintaining pointing stability on any fixed target for any valid attitude within the Field of Regard.	171		ISIM-264
SR-31	The Observatory shall be capable of tracking targets which exhibit any angular velocity in the range of 30 mas/s over a total motion 30 arcsec with respect to the guide star.	371, 372		
SR-32	The JWST science mission lifetime, after commissioning, shall be a minimum of 5 years.	44	PP 5.6	ISIM-349
SR-33	Propellant shall be sized for 10 years of operation after launch.	48		ISIM-542
SR-34	After commissioning, the ratio of prime exposure time on scientific targets to total elapsed time shall be greater than 70%.	102		ISIM-916

SR-35	During a normal operations contact, the JWST shall be capable of downlinking, and the Ground Segment capable of capturing and processing 229 Gigabits of science data, compressed from 458 Gigabits.	76		ISIM-111
SR-36	JWST shall be capable of observing targets of opportunity on timescales less than or equal to 2 days.			
SR-37	JWST shall enable a general observer and archival research program with observing time allocated according to peer review of science proposals.	344		

